Metadiscourse and Voice Construction in Discussion Sections in BA Theses by Chinese University Students Majoring in English

Jian-E Peng¹ and Yuting Zheng¹

Abstract
Voice is considered essential in academic writing, and metadiscourse is an important device contributing to voice. This study explores the use of metadiscourse and voice construction in Bachelor of Arts (BA) theses written at the onset and final stages by university undergraduates majoring in English in China. A corpus consisting of the discussion sections in the first and final versions of 35 BA theses was built, annotated, and analyzed. Two academics from this university were then invited to evaluate 10 pairs of the texts and specify textual elements that conveyed voice and to provide further comments in a follow-up interview. Results showed that the students used significantly more evidentials, hedges, and boosters in the final versions. The reviewers perceived minor growth in voice strength from the sample texts, and they commented that both content-related features and metadiscourse contributed to voice. This study highlights the importance of cultivating undergraduates’ awareness of voice construction and the use of metadiscourse in academic writing.

Keywords
voice, interactive metadiscourse, interactional metadiscourse, BA theses, undergraduates majoring in English

Introduction
Constructing an authorial voice, which is essential in academic writing (Morton & Storch, 2019), could be challenging to writers who use English as an additional or foreign language (EAL or EFL), and absence of voice has been noted as the common problem among these writers (Flowerdew, 2001; Hyland, 2016). Voice is viewed to represent a writer’s points of view, visibility in the text, and identity (Hyland & Sancho-Guinda, 2012; Tardy, 2016), and also to reflect the interaction between the reader and the text (Tardy, 2012). A graduation thesis is a typical genre expected to convey the writer’s voice; however, due to the “arcane conventions of academic discourse” (Hyland, 2016, p. 62) and the high-stakes status of graduation theses (Petrić, 2012), EAL writers may find thesis writing a daunting task.

In China, and probably in other similar EFL contexts (see also Altınmakas & Bayyurt, 2019), EFL undergraduate students who major in English are mandated to complete a thesis in English as a prerequisite to obtaining their Bachelor of Arts (BA) degree. This means that they have to possess skills in both research and English academic writing, which are usually insufficiently trained in BA programs in EFL contexts (see Zhang & Zhan, 2020). Voice-related problems such as lack of original ideas have been found prevailing in BA theses in China (Sheng & Zhou, 2011). In their systematic examination on educational and immediate contextual factors influencing Chinese students’ voice development, Zhang and Zhan (2020) pointed out that there has been a lack of both awareness and emphasis of voice in English writing instruction, and voice is absent in any scoring rubric in high-stakes English tests in China. Many teachers are even not familiar with the concept of voice (Zhang & Zhan, 2020). However, Zhang and Zhan (2020) noted that voice is now receiving much recognition in English language teaching at the tertiary level and young generations may be more apt at learning to develop voice in English writing. Therefore, empirical evidence is highly needed for understanding EFL undergraduates’ voice construction, especially in their English BA theses.

¹Shantou University, China

Corresponding Author:
Jian-E Peng, Professor, Department of Foreign Languages and Literature, College of Liberal Arts, Shantou University, 243 Daxue Road, Shantou 515063, Guangdong, China.
Email: pengjiane@stu.edu.cn
The construction of an authorial voice has been accorded increased importance in the research on essay writing (Çandarlı et al., 2015; Yoon, 2017; Zhao, 2013, 2017) and academic research writing (Matsuda & Tardy, 2007; Morton & Storch, 2019; Peng, 2019). While how voice is represented textually remains a question open to discussion (Morton & Storch, 2019), the trend in this field has been associating textual voice with the deployment of metadiscourse, often based on Hyland’s (2005) categories of interactive metadiscourse and interactional metadiscourse. Interactive metadiscourse indicates the writer’s awareness of readers’ needs and is used to organize a text to facilitate readers’ comprehension of the text, while interactional metadiscourse manifests the writer’s presence and involvement with both the text and readers. Although many studies have only examined the interactional dimension (Yoon, 2017; Zhao, 2013), the interactive dimension was considered here for two reasons. First, a purpose of this study is to systematically explore the use of metadiscourse in BA theses, which is an uncharted area. More importantly, interactive metadiscourse is an important device producing textual coherence and thus can reflect voice because, as Thompson (2012) noted, textual coherence can manifest the presence of voice.

Many studies have revealed that metadiscourse to different extents contributes to authorial presence (Çandarlı et al., 2015) or voice strength in undergraduates’ argumentative writing (Yoon, 2017; Zhao, 2017) or doctoral theses (Morton & Storch, 2019; Thompson, 2012). However, little is known about the use of metadiscourse in EFL undergraduates’ BA theses and voice constructed therein. The discussion section, in particular, should convey authorial voice as in this section the writer needs to establish links between the current findings and the literature, and claim research significance (Geng & Wharton, 2016). In addition, due to the fact that BA theses involve high-stakes evaluation, students are likely to engage in the process of writing and revising their theses. As academic writing is “always a situated practice” (Hyland, 2016, p. 63), the revision process may present students with “apprenticeship opportunities” during which changes in metadiscourse use may contribute to changes in voice. While voice development in academic writing has been recently explored (Morton & Storch, 2019), documented findings in this regard are still scarce.

To address the issues identified above, this study aims to explore the use of metadiscourse and voice construction in the discussion sections in the first and final drafts of BA theses written by Chinese first language (L1) students majoring in English and to discern the differences in the use of metadiscourse, if any, in relation to voice changes from the onset to the final stages of their thesis writing. Inquiry into metadiscourse use and its relation to voice construction in BA theses on a developmental timeline hopefully can offer pedagogical implications for assisting undergraduates’ thesis writing in China and other similar EFL contexts.

**Review of Literature on Voice and Metadiscourse**

**Research on Voice in Academic Writing**

Voice is an elusive and multifaceted construct, which has been conceptualized in different ways, ranging from being one’s personal and individual features in writing to showing one’s identity within certain disciplinary and social circumstances (Tardy, 2016; Yoon, 2017). This study concurs with Matsuda’s (2001) definition of voice as “the amalgamative effect of the use of discursive and non-discursive features that language users choose, deliberately or otherwise, from socially available yet ever-changing repertoires” (p. 40). More specifically, voice in this study is viewed as the writer’s textual presence. It should be noted that voice, in line with Matsuda’s (2001) definition, is usually perceived holistically from discursive and nondiscursive sources, intentionally or involuntarily. Besides, voice is both individual and social, meaning that the writer always constructs the unique self while conforming to the conventions and practice within a specific academic community. In addition, voice is also viewed as the textual effect on the reader, which is generated from the interaction between the writer, the reader, and the text (Matsuda & Tardy, 2007). Matsuda and Tardy (2007) argued that voice is jointly constructed by writers and readers because ultimately it is the readers who sense the effect of voice. Hence, voice research can benefit from an inclusion of the reader’s perspective.

While voice is widely accepted as indispensable in academic writing, recent empirical attention has been attached to issues as to what discursive devices contribute to voice and how to measure voice strength. As mentioned above, voice does not seem to rise from the use of any single textual feature; instead, Tardy and Matsuda (2009) found in their study that editorial board members of journals constructed a writer’s identity based on a wide range of features such as strange use of jargon or tentative style. Researchers have often subscribed to Hyland’s (2008) text-oriented voice model to examine voice elements, which is further introduced in the next section. Based on Hyland’s (2008) categories of metadiscourse, Zhao (2013) developed an analytic rubric for measuring voice in argumentative writing. Through factor analysis and also qualitative methods, such as think-aloud and interviews, the rubric was validated to capture three dimensions of voice: “(1) the presence and clarity of ideas in the content; (2) the manner of the presentation of ideas; and (3) the writer and reader presence” (Zhao, 2013, p. 201).

The analytic voice rubric established by Zhao (2013) has inspired ensuing studies into voice and its relationship with text quality. In her analysis of 200 Test of English as a Foreign Language (TOEFL) essays, Zhao (2017) found that, while the three dimensions of voice were correlated with holistic essay scores, the presence and clarity of ideas mostly predicted essay quality. This proposition was endorsed by Stock and Eik-Nes (2016) who cautioned that overemphasis...
on linguistic features may blind us from noticing “content-related features that might be more relevant in the construction of voice” (p. 89). Yoon (2017) drew on this rubric and reported, in 219 argumentative essays written by Greek L1 EFL students, weak to moderate correlations between textual voice elements and holistic voice strength, but weak correlation between holistic voice strength and essay quality. These findings collectively suggest that a strong voice should not be simplistically assumed to lead to high quality of essays.

While researching voice in essay writing by means of rigorous measurement has achieved a breakthrough, research on voice in graduation theses has mostly adopted holistic measures. For instance, Morton and Storch (2019) investigated voice development in two sets of comparable texts written by three PhD students in the first year and toward the end of their candidature. They adopted the reader’s perspective and invited five PhD supervisors to evaluate the writer’s voice conveyed in the selected texts based on a rubric adapted from Tardy (2012). Morton and Storch’s (2019) study innovatively elicited qualitative findings by asking the supervisors to mark with a highlighting pen any fragments of the text that gave them a sense of the writer’s voice and to attend a follow-up interview. This study concluded that the supervisors’ impressions of voice varied across their language backgrounds, disciplinary specialties, personal histories, and preferences. The less conclusive results compared with those obtained through measuring voice in essays may be because thesis writing entails disciplinary knowledge and methodological expertise besides manipulating language to express ideas. In other words, academic conventions such as the use of citation, knowledge display, and knowledge creation (Geng & Wharton, 2016) may complicate the measurement of voice in a thesis, not to mention that voice is multifaceted in its own right.

**Voice as Interaction Through Metadiscourse**

While voice can be expressed by various discursive means, Hyland (2008) proposed a text-oriented approach to capturing voice. This approach stresses the essence of interaction in voice construction, taking voice as the way writers “position themselves and their work in relation to other members of their group” (Hyland, 2008, p. 6). Hyland (2005) established an interpersonal model of metadiscourse, which distinguishes between interactive and interactional metadiscourse. Interactive metadiscourse concerns linguistic resources for organizing the text and for guiding the reader to comprehend the text, whereas interactional metadiscourse concerns the ways in which writers position their stance and attitudes, and involve readers into the text.

Of the two types of metadiscourse, interactional metadiscourse has received more research attention, particularly since Hyland (2008) explicitly proposed an interactional model in which voice is realized by two subtypes of interactional metadiscourse: stance and engagement. Stance is writer-oriented and refers to “the ways writers present themselves and convey their judgements, opinions, and commitments” (Hyland, 2008, p. 7) by means of four types of linguistic resources: hedges, boosters, attitude markers, and self-mention. Engagement tends to be reader-oriented, which functions to acknowledge the presence of the readers and align them toward the writer’s arguments or opinions by means of five textual elements: reader mention, directives, questions, knowledge reference, and personal asides. Hyland’s (2008) interactional model of voice has provided a solid theoretical foundation for the research of voice in academic writing. In their study, comparing the results and discussion chapters of master’s theses written in English and in Spanish, respectively, by English L1 and Spanish L1 students, Lee and Casal (2014) reported that the English texts contained significantly higher occurrences of metadiscourse, specifically, hedges, boosters, and attitude markers; the Spanish texts consisted of more self-mentions and engagement markers, which was attributed to the importance attached to in-groupness and politeness strategies in Spanish culture. Burneikaitė (2008) also identified significantly more frequencies of emphatic markers (a notion related to boosters) in master’s theses in linguistics written by English L1 writers than in those by Lithuanian-speaking EAL writers.

Arguably, interactive metadiscourse also reflects the writer’s voice. Interactive metadiscourse includes transitions, frame markers, endophoric markers, evidentials, and code glosses (Hyland, 2005; presented later). These resources can serve as signposts, guiding the readers through the text by anticipating their needs and expectations. Thompson (2012) rightly pointed out that metadiscourse can be used to augment textual coherence, which can “help cement the voice of authority” (p. 125). This insight found empirical support in Morton and Storch’s (2019) study where a PhD supervisor believed that coherence shows the writer’s control over their material and is “thus an important criterion in assessing authorial voice” (p. 19). Hyland and Tse (2004) also explicitly postulated that “all metadiscourse is interpersonal in that it takes account of the reader’s knowledge, textual experiences, and processing needs” (p. 161). Li and Wharton (2012) stressed that writers with a broad repertoire of metadiscourse can construct an intentional stance and voice in text. Therefore, it is fair to state that interactive metadiscourse, with its inherent interpersonal nature, builds connections between the writer, the reader, and the text, and also conveys the writer’s voice. Compared with interactional metadiscourse, interactive metadiscourse has been under-explored (Cao & Hu, 2014), let alone its association with voice. Lee and Casal’s (2014) study reported that the English thesis extracts, compared with the Spanish extracts, consisted of significantly more instances of transitions, endophoric markers, and evidentials, which assist readers’ textual navigation and are typical in English rhetorical conventions. In Burneikaitė’s (2008) study, the Lithuanian students employed significantly more text-connectives (a notion
related to transitions). This result was attributed to “an (over-)excessive emphasis on connectives in academic writing instruction at secondary or undergraduate level” (Burneikaitė, 2008, p. 43).

The view that metadiscourse can function as resources for constructing voice has also been articulated in other studies besides Hyland (2008). For instance, Helms-Park and Stapleton (2003) developed a Voice Intensity Rating Scale for measuring voice in argumentative writing. This scale contains four components, of which assertiveness can be established through the use of hedges and boosters, self-identification through self-mentions, and reiteration of central point arguably through code glosses. The last component, authorial presence and autonomy of thought, is evaluated primarily based on readers’ impression of the written text (Zhao & Llosa, 2008). Such impression can potentially derive from readers’ sense of being involved into the writer’s stance or attitudes and being guided through the discourse, which may be created by metadiscourse (Hyland, 2008). This being said, the roles of metadiscourse markers in conveying voice may not be on a par. Ádel (2006) pointed out that “the interpersonal function is, in fact, fundamental to the overall category” (p. 17). It is likely that interactional metadiscourse markers (e.g., hedges, boosters, and attitude marker) better demystify voice as they naturally reflect the writer’s stance and attitudes.

In brief, important issues regarding metadiscourse in relation to voice construction and voice improvement in BA thesis writing have remained unaddressed. Many EAL undergraduate students are inexperienced in both empirical research and English academic writing whose rhetoric conventions can be different from the L1 writing (Hirvela & Belcher, 2001). Efforts in this light can generate important implications for supporting undergraduates in their thesis writing, preparing them for further academic pursuits, and enhancing their interpersonal skills in written communication. This study was guided by the following research questions:

**Research Question 1 (RQ1):** To what extent do the first and final versions of the discussion section of the BA theses differ in the use of interactive and interactional metadiscourse?

**Research Question 2 (RQ2):** What are the metadiscourse elements, if any, that give the reviewers a sense of voice?

**Research Question 3 (RQ3):** Do the reviewers perceive any differences in voice construction in the sample pairs of the texts? If yes, in what ways?

## Method

**Corpus Materials and Informants**

The corpus materials included 35 sets of discussion sections taken from the first and the final versions of 35 BA theses written by English majors studying in a university in Southern China. English majors in this university have to complete a graduation thesis of about 6,000 to 8,000 words at the end of their final year. Students are required to, within a period of about 6 months, submit at least two drafts to their supervisors and revise the drafts according to the supervisor’s feedback before they are approved to submit the final version. Therefore, students’ engaging in process-oriented writing is mandated to improve the quality of the theses. The 35 sets of discussion sections were all taken from theses in applied linguistics with the consent of their writers. The decision to focus on applied linguistics was motivated by the generally shared argument that “academic writing varies systematically across disciplines” (Li & Wharton, 2012, p. 346). While the discussion section is often a necessary component in evidence-based research in applied linguistics, in the fields of literary or translation research, considerable structural variations can exist and there may not be a distinct discussion section in the theses, as was found in many theses archived in China National Knowledge Infrastructure (CNKI), which is the largest database of academic works in China.

We kept only the discussion section in each thesis, with other parts removed. As in some theses results/findings and discussion were blended into one section, we retained only the scripts related to discussion according to the function of a discussion section, that is, discuss the meaning and significance of the results/findings, as specified in Bitchener (2010). While this may affect text coherence, we considered it applicable in the current context because, as we found, the undergraduate writers basically followed the results-discussion pattern. These texts formed a corpus of 57,372 words, with 25,869 words in the first-draft subcorpus and 31,503 words in the final-draft subcorpus.

To triangulate the corpus analysis of metadiscourse in relation to voice construction, we also incorporated the reader’s perspective. Two native Chinese-speaking faculty members in the university where the corpus materials were collected were invited as informants, who were given the pseudonyms Jane and Mary. They both held a PhD degree, obtained overseas, and were experienced in publishing research articles in prestigious international journals such as those indexed in the Social Sciences Citation Index (SSCI). They were supervisors of the English majors, but not the supervisors of the theses analyzed in this study. They were invited to first review 10 of the 35 pairs of the discussion sections, rate the voice strength of the sample texts based on a rubric used in Morton and Storch (2019), and then to attend a follow-up interview.

**Instruments**

The instruments used in this study consist of three parts. The first is the UAM CorpusTool Version 3.3 (O’Donnell, 2019), which is a freeware annotation tool. This tool was used to facilitate the coding of metadiscourse in the discussion...
sections included in the corpus and it could return with the frequencies of the coded types of metadiscourse.

The second instrument is a rubric adapted from Morton and Storch (2019) to elicit the reviewers’ impressions of the sample discussion extracts. The rubric evaluated the extracts in terms of central focus, style, language, the writer’s credibility, and authorial presence. Voice strength examined in this study was based on the reviewers’ ratings on a 5-point scale in response to the statement “The writer’s authorial presence in this paper is” anchored, respectively, from weak/not noticeable (score = 1) to strong/very noticeable (score = 5).

An interview protocol was used to elicit the reviewers’ comments on their impression of voice construction and voice development in the English majors’ theses and factors contributing to their impression. This protocol is shown in Appendix B.

**Procedures**

This study was conducted in two phases. In the first phase, the discussion sections in the corpus were analyzed and coded using UAM CorpusTool (O’Donnell, 2019). The data were coded by the second author and, to enhance the reliability of the coding, five pairs of the texts (14.3% of the data) were randomly selected and independently coded by the first author. Inter-coder reliability calculated for each subtype of interactive and interactional metadiscourse ranged between 90% and 93%, which were “in the 90% range” specified by Miles and Huberman (1994, p. 64). Upon completing the coding, normalized frequencies per 1,000 words of the two main types of metadiscourse and their subtypes, that is, (frequency of each type/number of words in the text) * 1,000, were calculated by the UAM CorpusTool and obtained for subsequent statistical analyses.

In the second phase, the two informants were invited to review 10 sample pairs of texts, each being labeled with a number plus “a” or “b” (e.g., Text 1a, Text 1b). They were informed that the labels of “a” and “b” in the file name were randomly used without suggesting the sequence of a pair of texts. The reviewers were asked to highlight any words, expressions, or other textual chunks that gave them a sense of the writer’s voice (see Morton & Storch, 2019). Following this, the metadiscourse elements highlighted by the reviewers were identified and categorized according to Hyland’s (2005) framework, and their frequencies in the first and final versions of each pair, respectively, were then calculated. The reasons for presenting only 10 pairs of texts to the reviewers were twofold. First, language teachers nowadays are already “steeped in work and responsibility” (Hobbs & Kubanyiova, 2008, p. 503), which was the case for the two reviewers at the time this study was conducted, so it was impractical to request them to review 35 pairs of texts. In addition, as the purpose of the second phase was to complement the corpus-based analysis with qualitative findings from the reader’s perspective, the 10 pairs of texts were considered suitable for fulfilling this purpose. The reviewers were also asked to rate each text’s voice strength based on the rubric adapted from Morton and Storch (2019). They were then interviewed individually for about 20 to 30 min. The interviews were recorded with their consent and transcribed verbatim.

**Analytic Framework of Metadiscourse**

The corpus was analyzed based on Hyland’s (2005) interpersonal model of metadiscourse, which distinguishes between interactive metadiscourse and interactional metadiscourse. Interactive metadiscourse includes five subtypes: transitions, frame markers, endophoric markers, code glosses, and evidentials. Transitions (therefore) function to signal logical connections within the text and thus enhance cohesion. Frame markers are used to organize texts in terms of sequences (second), topics (with regard to), discourse stages (thus far), and discourse goals (aim to). Endophoric markers are used to refer to other parts of the text (as mentioned above). Code glosses function to exemplify or rephrase propositional meanings (in other words). Evidentials are used to introduce information from other sources into the text through citations (according to).

While interactive metadiscourse is used to maneuver textual elements to meet readers’ expectations of the unfolding text, which enhances textual coherence and thereby reflects writers’ presence, interactional metadiscourse is used to directly address the readers. Five subtypes of interactional metadiscourse were examined in this study: hedges, boosters, attitude markers, self-mentions, and engagement markers. Hedges (perhaps) are the devices used to mitigate assertiveness and open dialogic space, whereas boosters (demonstrate) are used to emphasize writers’ certainty and to close down alternatives. Attitude markers are used to convey writers’ attitudes toward propositions, which include attitude verbs (prefer), adverbs (interestingly), and adjectives (important). Self-mentions refer to the use of first-person pronouns and possessive structures to indicate author presence in the text (I). Engagement markers are devices used to explicitly address the reader (you) or pull them into the discourse (note; see Appendix A for the metadiscourse markers examined in this study).

**Data Analysis**

Data analysis includes several steps. Based on the framework presented above, tokens of interactive and interactional metadiscourse and their subtypes were identified and annotated. Several practices were adopted. First, the identification of self-mentions was based on Hyland’s (2005) framework, also considered a broad approach (Alotaibi, 2018; Hyland, 2017), which sidesteps a distinction made by Ädel (2006) between references to the world of discourse and those to the real world, which represents a narrow approach. While differentiating self-mentions in the two “worlds” (Ädel, 2006)
surely allows a focused lens on the text per se and on “the current writer and reader qua writer and reader” (p. 183), in practice, self-mentions referencing to the world of discourse and real world are not always clear-cut (Alotaibi, 2018) and may depend on genres (Alotaibi, 2018) and disciplines (Lee & Casal, 2014). Therefore, to avoid the risk of “eliminating much of what makes metadiscourse a powerful analytic tool” (Hyland, 2017, p. 27), Hyland’s (2005) broad approach was adopted in this study.

Following Cao and Hu’s (2014) analysis, intra-sentential transitions (e.g., because, since) were not counted as transitions because they basically fulfill syntactic functions instead of acting as “metadiscoursal logical markers” (p. 19). Besides, the often-used pair of phrases, on the one hand and on the other hand, were coded as frame markers signaling spatial sequences when they were used to introduce sequences, but coded as transitions when they were used to show logical connections (Cao & Hu, 2014). Modal verbs like could and would were counted as hedges only when they were used to qualify degrees of certainty, but were excluded when used “in a subjunctive mood or past tense that had nothing to do with hedging” (Zhao, 2013, p. 205). In addition, a distinction was made between self-mention we and engagement marker inclusive we, the former referring to authorship, whereas the latter functioning to capture readers’ attention or involve readers in the discourse (Hyland, 2005).

To address the first research question, we performed dependent $t$ tests to compare between the first and final versions the normalized frequencies of interactive metadiscourse and interactional metadiscourse, respectively. The assumptions that sample means and the differences between sample means were normally distributed were met (Field, 2009). However, the data of the subtypes of each metadiscourse violated the assumption of normal distribution. Hence, Wilcoxon signed-rank test, which is the nonparametric equivalent of dependent $t$ test (Field, 2009), was performed to test the differences in each subtype of metadiscourse between the first and the final drafts.

To answer the second research question, the metadiscourse elements highlighted by the reviewers were categorized according to Hyland’s (2005) framework and their frequencies in the first and final versions of each pair, respectively, were then calculated. As the purpose here was to complement the corpus-based analytic findings from the reader’s perspective, inferential statistics was not processed.

The third research question was addressed by drawing on several types of data. The voice strength of each text was obtained by averaging the two reviewers’ ratings. We then examined the voice strength of each text and the frequencies of metadiscourse markers highlighted by the reviewers as reflecting the writer’s presence for the purpose of spotting association or compatibility, if any, between the two types of information. In addition, qualitative content analysis (Bryman, 2012) was performed on the interview transcripts. This involved repeatedly reading and searching in the reviewers’ comments for themes regarding voice improvement in the sample texts and possible underlying reasons. The themes reported in this article concern the roles of metadiscourse and other aspects that were perceived by the reviewers to account for differences in voice between the first and final versions of the texts. In addition, a theme regarding supervisors’ role in guiding students’ voice construction also inductively emerged from the interview data.

### Results

#### Use of Metadiscourse in the First and Final Versions of the Texts

For the dependent $t$ tests and Wilcoxon test conducted in this study, effect size (ES) was obtained, referencing Field (2009), by calculating $r$ in $t$ test ($r = \sqrt{t^2 / t^2 + df}$) and $r$ in Wilcoxon test ($r = Z / \sqrt{n}$). The cutoff values of $r$, suggesting a small, medium, and large effect are 0.1, 0.3, and 0.5, respectively (Field, 2009).

The results of dependent $t$ tests showed that both interactive and interactional metadiscourse occurred more frequently in the final-draft subcorpus ($M = 19.39, SD = 7.02$; $M = 19.40, SD = 7.22$, respectively) than in the first-draft subcorpus ($M = 16.61, SD = 8.06$; $M = 13.24, SD = 7.44$, respectively). These differences were statistically significant ($t = 2.37, p = .02, r = .29$ for interactive metadiscourse; $t = 4.42, p < .001, r = .53$, both $M = 7.44, SD = 19.39$, large ES, for interactional metadiscourse).

Detailed differences in the subtypes of interactive and interactional metadiscourse between the two subcorpora were further examined by running a series of Wilcoxon signed-rank tests. Table 1 shows the descriptive statistics and the test results.

As seen in Table 1, the final-version subcorpus contained significantly more evidentials ($z = 2.37, p = .02, r = .29$), more hedges ($z = -2.22, p = .03, r = .27$), and more boosters ($z = -2.37, p = .02, r = .28$) than the first-version subcorpus. Except for the three subtypes, no significant differences in other subtypes were identified between the two subcorpora. The increased use of evidentials, hedges, and boosters in the two versions is illustrated in the following three pairs of examples, with these markers in the second version shown in bold:

1a) The result of the research shows that the more frequently students use English language learning strategies, the higher TEM-Fours scores they achieve. Therefore, English learners should pay attention to the use of English language learning strategies. (Text 12, first version)

1b) This study indicates that the more frequently the participants used English language learning strategies, the higher TEM-Fours scores they achieved.
This result is consistent with the research results in Li (2007) and Yang (2007). Li also reported that ... (Text 12, final version)

As seen in Examples 1a and 1b, the writer in the final version employed the evidentials in the form of citation to align his or her result with that of previous studies, which was not found in the first draft. The use of the evidential to represent an idea from another source in the final version can convey the writer’s voice through his or her alluding to the disciplinary literature.

(2a) SLLE influenced students’ intended effort. (Text 18, first version)

(2b) The results indicate that students’ English learning experience in schools may greatly influence their intended effort. (Text 18, final version)

In contrast to Example 2a, Example 2b contains the hedge may, which indicated the writer’s awareness of calculating the weight given to his or her assertion (Hyland, 2005). The current finding adds evidence to Hyland’s (2008) finding that hedges are “the most frequent feature of writer perspective” (p. 12). Zhao’s (2013) factor analysis result also showed that the use of hedges is one unique dimension of voice.

(3a) FD/FI cognitive style affects not only the listening comprehension level, but also the use of listening strategy. (Text 4, first version)

(3b) The learners’ different cognitive style (FD/FI) was also found to impact their use of listening strategies. (Text 4, final version)

In Examples 3a and 3b, the use of the booster found in the final version indicated the writer’s commitment to the proposition, through which the readers are encouraged to take on board his or her opinion.

**Metadiscourse Perceived to Convey Voice**

The metadiscourse elements highlighted by the reviewers as giving them a sense of voice were categorized and calculated. The reviewers’ ratings of each text’s voice strength were averaged. Table 2 shows the frequencies of the highlighted metadiscourse markers in each text and its corresponding voice strength.

The first result found in Table 2 is that the metadiscourse markers giving the reviewers a sense of voice included all subtypes. Of these subtypes, hedges, transitions, and evidentials were the top three in terms of frequencies (i.e., 85, 83, and 65, respectively). Coincidently, as reported above, two of these three subtypes, that is, hedges and evidentials, were found to occur significantly more frequently in the final-version subcorpus than in the first-version subcorpus.

Table 2 also shows that metadiscourse markers exhibited more instances in the final version of eight pairs of the texts and showed fewer instances in the final version of only two pairs (i.e., Text 2 and Text 10). This result was also largely consistent with the Wilcoxon signed-rank test results reported previously, which together may imply a rough tendency toward an increased use of metadiscourse in the final version. In other words, the increased use of metadiscourse played a part in capturing the reviewers’ attention when they interpreted the writer’s voice in the text.

It can also be observed in Table 2 that the pairs of Texts 6, 1, and 5 ranked the top three in the increased use of metadiscourse (i.e., 28 − 5 = 23, 41 − 22 = 19, and 22 − 11 = 11, respectively), and these three pairs happened to rank top three in voice strength growth (3.25 − 1.75 = 1.50, 3 − 1.5 = 1.5, and 4.25 − 3 = 1.25, respectively). Similar changes in

### Table 1. Descriptive Statistics and Wilcoxon Signed-Rank Test Results.

| Subtype            | First version | Final version | z  | p   | r   |
|--------------------|--------------|---------------|----|-----|-----|
|                   | M            | SD            | M  | SD  |     |
| Interactive        |              |               |    |     |     |
| Transitions        | 7.52         | 4.88          | 7.81| 3.85| −0.14| .89 | .02|
| Frame markers      | 1.96         | 2.51          | 1.86| 1.91| −0.74| .46 | .09|
| Endophoric markers | 1.10         | 1.27          | 1.30| 1.16| −0.97| .33 | .12|
| Evidentials        | 4.05         | 3.53          | 5.09| 3.39| −2.22*| .03 | .27|
| Code glosses       | 2.20         | 2.55          | 3.23| 2.64| −1.75| .08 | .21|
| Interactional      |              |               |    |     |     |
| Hedges             | 8.46         | 5.46          | 13.00| 5.53| −4.42**| .00 | .53|
| Boosters           | 1.80         | 2.10          | 2.96| 2.96| −2.37*| .02 | .28|
| Attitude markers   | 1.09         | 1.32          | 1.17| 1.32| −0.54| .59 | .06|
| Self-mentions      | 0.13         | 0.42          | 0.14| 0.84| −0.67| .50 | .08|
| Engagement markers | 1.76         | 2.85          | 2.19| 1.94| −1.79| .07 | .21|

*p < .05. **p < .01.
the same direction could be observed in the pairs of Texts 2, 4, and 7. Although these anecdotal findings certainly could not amount to any conclusive evidence, it may be inferred that authorial voice perceived by the reviewers may, at least partially, be attributed to the use of metadiscourse. The following two pairs of examples illustrate the metadiscourse markers not found in the first drafts of Text 4 and Text 1 but were highlighted by the reviewers in their final versions.

(4a) Also, ought-to L2 self as an extrinsic motivation shapes ideal L2 self directly because the attitudes to the importance of L2 from learners’ parents, peers, boss form an opinion field. (Text 4, first version)

(4b) Also, ought-to L2 self as an extrinsic motivation was found to shape ideal L2 self directly. This is probably because others’ attitudes toward the importance of L2 influence could influence learners’ expectations of their L2-related self-image in the future. (Text 4, final version)

The proposition in Example 4b was more sufficiently elaborated than in Example 4a, and the writer added two hedges probably, could when offering his or her interpretation. The use of these hedging devices indicated the writer’s recognition of other alternative opinions and cautions to allow space for negotiation (Hyland, 2005). Notably, this uncertainty made in staking a claim did not seem to erase the writer’s voice but instead enabled the writer to enhance his or her presence in the text by inviting other voices in the proposition.

(5a) The result of this study that father’s employments and income place stronger influence on the motivation reflect nowadays’ income gap between genders in China. (Text 1, first version)

(5b) In addition, the present study found that father’s employment and income had stronger influence on learner’s motivation than mothers’ did. This result may reflect the income gap between genders nowadays in China. (Text 1, final version)

Similarly, in Examples 5a and 5b, the original sentence was revised into two sentences in the final version, and the transition in addition and hedge may were added. The use of the transition here indicated the writer’s awareness of readers’ expectations and his or her efforts in guiding the readers through the text. Hence, this transition could function to improve the information flow and thereby textual coherence, which resulted in a clearer presence of the writer as the text organizer.

**Voice Construction in the First and Final Versions of the Sample Texts**

During the interviews, both Jane and Mary commented that overall a certain extent of improvement in voice construction
could be perceived when comparing each pair of the texts, but emphasized that much variations existed across different pairs. This was corroborated by the figures in Table 2, where six pairs registered increased voice strength (Texts 1, 3, 4, 5, 6, and 7), two no such increase (Texts 9 and 10), and two a drop in voice strength (Texts 2 and 8). Based on the last column in Table 2, the voice strength of the first and final versions of the 10 texts could be averaged, respectively, and the resulted values were 3 and 3.225, which indicated a minor increase in voice strength.

When asked to consider the role of metadiscourse in constructing voice, both Jane and Mary acknowledged that metadiscourse plays important roles. For instance, “signaling words” could “bring the writer’s voice to the readers” (Jane); words like to summarize could “draw the readers’ attention, that is, I am going to summarize my above-mentioned opinion” (Mary). Hedging devices such as may, possibly could “soften the writer’s argument” and when combined with verbs, “the writer’s voice could become clearer” (Jane).

Besides metadiscourse, three content-related aspects were considered by the reviewers as factors accounting for differences in voice construction between the two versions of the sample texts, or lack thereof. The first aspect is clearer connections between the present study and previous studies presented in the final version. When commenting on the pair of Text 1, Jane noted,

I think in the final draft the writer at least mentioned many consistencies between his/her findings and those in the literature, which was not mentioned in the first draft. The first draft reads like he/she was reporting what others had done, as in the Literature Review section, which showed unclear connections between [others’ studies] and his/her study.

Jane’s opinion was exactly shared by Mary, who also commented on the final version of Text 1, “He or she also built connections with previous studies, saying what previous studies have found, and my research confirmed their findings, etc.”

The second aspect contributing to voice are the explicit statements of the significance of the present study found in the final version. Jane emphasized, “How to highlight the significance? How to highlight the findings? What are the significant points? If the writer could make revisions in these lights, his or her voice could be projected.” Mary also addressed, “In term of authorial voice, I think he or she highlighted his or her findings, so [the voice] was stronger.”

The reviewers also mentioned that more elaborations or explanations on the writer’s propositions found in the final version contributed to a stronger voice. For instance, when comparing the pair of Text 5, Jane noted that the writer’s explanations on his or her propositions in the final version were much clearer than in the first draft. She added, “If the authorial voice was stronger, it was mainly due to some revisions, such as giving clearer explanations.” Coincidently, Mary also mentioned that the final draft of Text 5 was “overall written in a clearer way” and “the writer’s explanations on the present findings were clearer.”

Notably, near the end of the interview, both reviewers acknowledged the important role that supervisors could play in guiding students to convey voice in their BA thesis writing, as Jane noted,

I think if the teacher consciously does this thing and tells students that “in your final version, I hope you can express your own voice as a researcher,” better effect may be achieved. Otherwise, if [voice construction] was carried out unconsciously in their writing, maybe there is no much effect.

However, based on the sample texts they reviewed, they doubted that voice was not stressed in the supervision process, as Mary mentioned,

When I was highlighting [the textual elements], my feeling was that voice was not the concern of the student writers’ supervisors.

It can be seen that with their professional training and experience in academic writing, both Mary and Jane regarded voice construction important, but for students to successfully construct their voice in thesis writing, explicit guidance from their supervisors is indispensable.

**Discussion**

This study has compared undergraduate English majors’ use of metadiscourse and voice construction in the first and final versions of the discussion sections of 35 BA theses. Premised on the large body of literature on the contribution of metadiscourse to voice construction in academic writing (Hyland, 2008; Li & Wharton, 2012; Thompson, 2012; Zhao, 2013), this study first identified significant increase in the use of interactive and interactional metadiscourse from the first to the final versions of the 35 extracts. Among the subtypes, evidentials, hedges, and boosters were used significantly more frequently in the final-version subcorpus. Evidentials refer to the metalinguistic resources that incorporate “a community-based literature” (Hyland, 2005, p. 51) to support writers’ arguments. The increased occurrence of evidentials may imply that, at the later stage, the undergraduates became more familiar with using citations to render their findings and conclusions more cogent, which is a typical feature in academic writing. The increase in the use of hedges and boosters somewhat confirmed the salient roles of these two subtypes reported in Yoon (2017). Yoon (2017) found that hedges and boosters significantly correlated with holistic voice strength while hedges significantly correlated with holistic essay quality. Despite the different genres examined in Yoon’s (2017) study and the present study, these results imply that hedges and boosters are more accessible to undergraduates when they seek progress in
academic writing. Zhao (2017) also identified that “the writer and reader presence” dimension of voice, which includes hedges and boosters, was significantly correlated with TOEFL essay scores.

From the reader’s perspective, the metadiscourse highlighted by the reviewers as signs of voice in the sample texts included all the subtypes, of which hedges, transitions, and evidentials were the top three with distinctive frequencies. Interestingly, the corpus-based analysis also revealed that hedges and evidentials occurred significantly more frequently in the final-version subcorpus. Considered jointly, these results may be explained by the fact that undergraduates are familiar with transitions, which are commonly emphasized in writing instruction at undergraduate level (Burneikaitė, 2008; Li & Wharton, 2012), but they have less experience in using hedges and evidentials related to citations, which characterize conventions in academic writing (Groom, 2000; Hyland & Sancho-Guinda, 2012). For instance, according to Davis (2013), even international postgraduate students often encounter linguistic and cultural barriers in source use. Therefore, in this study, the student writers’ notable progress in using hedges and evidentials may be attributed to their practice during the process of thesis writing. The current finding also confirmed the close relationship between hedges and voice construction (Yoon, 2007). It may imply that experienced academics, such as the reviewers in this study, tend to sense textual voice from the writer’s discursively expressed cautions in staking claims (Nelson & Castelló, 2012), textual coherence created with the use of transitions (Thompson, 2012), and citations or allusion to reliable sources (Groom, 2000). In particular, the reviewers seemed to attach particular attention to interactive metadiscourse markers (i.e., transitions and evidentials) when discerning voice. These results were expected as transitions can “create textual cohesion by signaling logical links between propositions” (Cao & Hu, 2014, p. 19), and manifest the writer’s awareness of the reader’s expectations. Their resulted coherence can augment the writer’s voice. In addition, evidentials used for citing sources, when masterly used, can function to manipulate the degree of authorial voice (see Groom, 2000; Peng, 2019). In sum, the current findings provided support to the role of at least some subtypes of interactive metadiscourse in conveying voice.

It should be noted that, from the reviewers’ perspectives, while voice can be constructed by skilled use of metadiscourse, three aspects were all important: clear connections built between the present study and previous studies, explicit statements of research significance, and elaborations or explanations on the writer’s propositions. These aspects largely overlapped with the ideational dimension of voice established by Zhao (2013) and corresponded to the content-related features emphasized by Stock and Eik-Nes (2016). Although the reviewers did not explicitly associate voice construction with text quality, it appeared that the above aspects considered important in voice construction were also the benchmark in their evaluation of students’ theses. These findings also echoed Zhao’s (2017) quantitative result that only the ideational dimension of voice significantly predicted text quality. Taken together, it could be argued that “clear presentation of a unique and sophisticated idea” underscored by Zhao (2017) in the writing of argumentative essays may also apply to undergraduates’ thesis writing in this context.

In addition, the reviewers’ comment on voice not being a concern of supervisors coincidently corroborated Zhang and Zhan’s (2020) perspective that voice has not been a focus in English writing instruction in China. This could be because supervisors themselves lacked an awareness of voice (Zhang & Zhan, 2020). It may also be explained by the situations in many undergraduate programs for English majors in China. According to Sheng and Zhou (2011), teachers often have to attend to grammar and word choice in writing instruction due to students’ low language proficiency. Understandably, in a long piece of writing like thesis, mistakes in language forms and problems in structures may be pervasive, which could occupy supervisors’ major attention. That said, the above comment was only the reviewers’ speculation and compelling evidence collected from supervisors of BA theses is needed in future research.

Another attempt of this study was to compare differences in the use of metadiscourse and voice construction between the first and the final versions of undergraduates’ English BA theses. Based on the results of corpus-based analysis and “the reader-based approach” (Stock & Eik-Nes, 2016, p. 89), we could only conclude that while minor increase in voice strength might have occurred, variations existed across the student writers. There were some signs that the increase in the use of metadiscourse matched the increase in voice strength evaluated by the reviewers; however, due to the small sample size, these were strictly restricted to anecdotal observations that require far more empirical evidence before any conclusions could be made. Moreover, the two reviewers’ comments on voice improvement in the sample texts indicated that they tended to grasp voice in a holistic way, capitalizing on features at different levels ranging from word choice to explicit connections established between the current and previous studies. This was also reported by Morton and Storch (2019) when they described how five PhD supervisors evaluated voice in doctoral theses. Collectively, these results pointed to the elusive and ambiguous nature of voice often discussed in the field (Morton & Storch, 2019; Xu & Zhang, 2019).

Some cautions should be considered when interpreting the current findings. First, this study involved only a small size of texts due to our focus on research-based theses in applied linguistics. While this focused lens has allowed us to explore patterns of metadiscourse use and voice construction, similar inquiry into BA theses in other fields such as literary and translation research was not attempted in this study. The second aspect is that we only compared the textual aspects
contributing to differences in voice construction in the first and the final versions of the BA theses, but have not explored what happened in between or, more specifically, other factors internal or external to the student writers. This was due to the focus set on the corpus-based analysis and to the fact that all the writers of these BA theses were graduated from the university at the time of the research. Even being approached, it was unlikely that they would recall accurately the details of their thesis writing process. Possibly due to similar reasons, Morton and Storch (2019) also did not examine things happening to the three PhD candidates between the first year and toward the end of their candidature when comparing their texts written in the two stages. Hence, longitudinal studies are certainly imperative in future research to harness insights in this direction.

Notwithstanding its limitations, this study may indicate some pedagogic implications. The result that undergraduates used significantly more such types of metadiscourse as evidentials, boosters, and hedges in their final drafts suggests the accessibility of metadiscourse even to the beginner writers of empirical research although these writers may not necessarily use these metadiscourse markers by intentional choice. Therefore, second language (L2) writing instructors and supervisors need to raise students’ awareness of the interactional functions of metadiscourse and their roles in constructing voice. This could be realized by first giving explicit instructions. Teachers could simplify the theoretical underpinnings of Hyland’s (2005) model of metadiscourse and use plain language to explain each type of metadiscourse and its functions along with its exemplary use in academic writing.

Upon equipping students with the knowledge of metadiscourse markers, teachers could provide sample research papers and ask students to identify all metadiscourse markers in those papers. Articles published in prestigious journals or articles by proficient English-speaking writers in a learner corpus such as the British Academic Written English (BAWE) corpus (Nesi & Gardner, 2012) would be useful materials. Several rounds of practice would be needed before students could digest the delicate discursive effects of the metadiscourse markers. Teachers could then supply students with different versions that juxtapose the presence and absence of metadiscourse markers in places where they are necessary and ask students to comment on different textual effects. In addition, students should be given opportunities to revise their drafts, during which peer review and teacher review should be carried out to enhance their understanding of and ability to use metadiscourse skillfully in academic writing.

Another important implication is that students should receive systematic training in the ideational dimension of voice (Zhao, 2013), such as connecting to the literature, citing reliable sources, and discursively spelling out research significance. These aspects, as found in this study, carry much weight in constructing an authorial voice in academic writing.

Conclusion

This study has explored metadiscourse use and voice construction in a group of undergraduate English majors’ BA theses written up at two points in time: the onset and final stages of their thesis writing. Based on the analysis of the corpus comprising the discussion sections in the first and final versions of 35 BA theses, we found that the student writers used significantly more evidentials, hedges, and boosters in the final versions. From the two invited reviewers’ perspectives, minor growth in voice strength could be discerned, particularly manifested in the use of transitions, hedges, and evidentials. However, while the roles of metadiscourse in voice construction were acknowledged, content wise, connections to previous studies, explicit statement of research significance, and elaborations or explanations on the writer’s propositions were believed to play more significant roles in projecting an authorial voice.

The novelty of this study lies in its particular lens placed on voice construction and metadiscourse use in EFL undergraduates’ thesis writing, an under-explored area that pertains to EAL student writers worldwide. Undergraduate students usually have received only basic, if at all, training in conducting empirical research and academic writing. Their academic experiences that culminate with the high-stakes BA thesis writing could exert an important impact on their future academic development. A large body of research has shown that voice construction in academic writing is a challenge faced by EAL writers or international students (Davis, 2013; Flowerdew, 2001; Hirvela & Belcher, 2001). Matsuda (2001) argued that Japanese students’ difficulties in constructing voice mainly stem from their lack of familiarity with strategies for voice construction in English. Previous studies (Burneikaite, 2008; Lee & Casal, 2014) have also shown that Spanish L1 and Lithuanian EAL student writers in their theses employed metadiscourse in ways different from their English L1 counterparts, partly due to different rhetoric conventions. This study has further shed light on how metadiscourse and other content-related aspects could contribute to voice construction in undergraduates’ BA theses written in English. Findings of this study have also alerted both language teachers and thesis supervisors of EAL students to the importance of deliberately guiding students to construct voice by means of various discursive devices. In doing so, teachers and supervisors should, in the first place, possess the awareness, knowledge, and strategies of voice construction, which are crucial to effectively implementing voice pedagogy (Zhang & Zhan, 2020) and enhancing students’ skills in managing interaction in written communication.
Appendix A

Metadiscourse Markers Examined in This Study

Transitions
- accordingly
- additionally
- also
- although
- all then
- and
- as a result
- at last
- at the same time
- at the meanwhile
- besides
- but
- contrary to
- consequently
- further
- furthermore
- hence
- however
- in addition
- in contrast
- in the meantime
- leads to
- moreover
- nevertheless
- on the contrary
- on the other hand
- similarly
- since
- so
- still
- therefore
- thus
- what is more
- yet

Endophoric markers
- above examples
- as mentioned above
- as previously introduced
- from what was reported
- from the table
- as what have been discussed
- in this example
- in these examples
- (in) this section
- (in) this study
- (see) Figure X
- (in) the interview
- the ensuing close analysis
- X above

Evidentials
- (date)/(name)
- as X suggests
- according to X
- (various forms of citations)

Code glosses
- for example
- for instance
- I mean
- i.e.
- it means
- in other words
- or X
- specifically
- such as
- that is
- this means
- that is to say
- (that) means
- which means

Hedges
- about
- apparent
- apparently
- appeared
- appears
- at this time
- could
- could be
- from an objective view
- from one’s perspective
- from this perspective
- generally speaking
- indicated
- in general
- in most cases
- in terms of this study
- in this way
- in this sense
- in this situation
- in this regard
- in one’s opinion
- in this case
- largely
- likely (to)
- mainly
- may (be)
- maybe
- may not
- might
- might be
- mostly
- nearly
- often
- on average
- on this study
- perhaps
- possibly
- probably
- quite
- relatively
- seems
- seemed to
- should
- should not
- slightly
- simply
- sometimes
- somewhat
- suggest(ed/s)
- tend(ed) to
- to a certain extent
- to a large extent
- to some extent
- would (be)
- wholly speaking

Boosters
- actually
- always
- assumed
believed
demonstrate(d)
displays
find
found
implied
implicating
know
reflects
reveal(ed)
showed / shown /shows
important
more importantly
indispensable
interestingly
necessary
meaningful and helpful
noteworthy
not surprised
paradoxically
promising
reasonable
reasonably
sensible
surprisingly
understandable
unexpected
vital
worth noting
worth to note
Self-mentions
I
we
my (study)
the researcher
Engagement markers
compare(d)
considering
comparing with
don’t/didn’t/does not
have to
lead to
must
need to
note
regard
should (be)
should (be)
we/us (inclusive)

Appendix B

Interview Protocol

1. What are your impressions on the writer’s presence in undergraduate English majors’ theses? How important do you think authorial voice is in an undergraduate’s thesis?
2. Based on Texts 1a and 1b, in what ways do you think that the writer’s voice has become stronger? (Repeat this question for the remaining pairs until saturation point was reached).
3. Generally, did you perceive any development of authorial voice when comparing the pairs of the 10 texts?
4. In the teaching of second language writing, in what ways can we teach undergraduate English majors to construct their voice in academic writing?

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Jian-E Peng https://orcid.org/0000-0003-3985-4233

References

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

Alotaibi, H. S. (2018). Metadiscourse in dissertation acknowledgments: Exploration of gender differences in EFL texts. Educational Sciences: Theory & Practice, 18(4), 899–916. https://doi.org/10.12738/estp.2018.4.0247

Altınmakas, D., & Bayyurt, Y. (2019). An exploratory study on factors influencing undergraduate students’ academic writing practices in Turkey. Journal of English for Academic Purposes, 37, 88–103. https://doi.org/10.1016/j.jeap.2018.11.006

Bitchener, J. (2010). Writing an applied linguistics thesis or dissertation: A guide to presenting empirical research. Palgrave Macmillan.

Bryman, A. (2012). Social research methods (4th ed.). Oxford University Press.

Burneikaité, N. (2008). Metadiscourse in linguistics master’s theses in English L1 and L2. Kalbotyra, 59(3), 38–47. https://doi.org/10.15388/Klbt.2008.7591

Çandarlı, D., Bayyurt, Y., & Martı, L. (2015). Authorial presence in L1 and L2 novice academic writing: Cross-linguistic and cross-cultural perspectives. Journal of English for Academic Purposes, 20, 192–202. https://doi.org/10.1016/j.jeap.2015.10.001

Cao, F., & Hu, G. (2014). Interactive metadiscourse in research articles: A comparative study of paradigmatic and disciplinary influences. Journal of Pragmatics, 66, 15–31. https://doi.org/10.1016/j.pragma.2014.02.007

Davis, M. (2013). The development of source use by international postgraduate students. Journal of English for Academic Purposes, 12(2), 125–135. https://doi.org/10.1016/j.jeap.2012.11.008

Field, A. (2009). Discovering statistics using SPSS (3rd ed.). Sage.

Flowerdew, J. (2001). Attitudes of journal editors to non-native speaker contribution. TESOL Quarterly, 35(1), 121–150. https://doi.org/10.2307/3587862

Geng, Y., & Wharton, S. (2016). Evaluative language in discussion sections of doctoral theses: Similarities and differences between L1 Chinese and L1 English writers. Journal of English for Academic Purposes, 22, 80–91. https://doi.org/10.1016/j.jeap.2016.01.001
Groom, N. (2000). Attribution and averral revisited: Three perspectives on manifest intertextuality in academic writing. In P. Thompson (Ed.), Patterns and perspectives: Insights into EAP writing practice (pp. 14–25). Centre for Applied Language Studies.

Helms-Park, R., & Stapleton, P. (2003). Questioning the importance of individualized voice in undergraduate L2 argumentative writing: An empirical study with pedagogical implications. *Journal of Second Language Writing, 12*, 245–265. https://doi.org/10.1016/j.jslw.2003.08.001

Hirvela, A., & Belcher, D. (2001). Coming back to voice: The multiple voices and identities of mature multilingual writers. *Journal of Second Language Writing, 10*(1–2), 83–106. https://doi.org/10.1016/S1060-3743(00)00033-2

Hobbs, V., & Kubanyiova, M. (2008). The challenges of researching language teachers: What research manuals don’t tell us. *Language Teaching Research, 12*(4), 495–513. https://doi.org/10.1177/1362168808097162

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

Hyland, K. (2008). Disciplinary voices: Intersections in research writing. *English Text Construction, 1*(1), 5–22. https://doi.org/10.1075/etc.1.1.03hyl

Hyland, K. (2016). Academic publishing and the myth of linguistic injustice. *Journal of Second Language Writing, 31*, 58–69. https://doi.org/10.1016/j.jslw.2016.01.005

Hyland, K. (2017). Metadiscourse: What is it and where is it going? *Journal of Pragmatics, 113*, 16–29. https://doi.org/10.1016/j.pragma.2017.03.007

Hyland, K., & Sancho-Guinda, C., (Eds.). (2012). Stance and voice in written academic genre. Palgrave MacMillan.

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics, 25*(2), 156–177. https://doi.org/10.1093/applin/25.2.156

Lee, J. J., & Casal, J. E. (2014). Metadiscourse in results and discussion chapters: A cross-linguistic analysis of English and Spanish thesis writers in engineering. *System, 46*, 39–54. https://doi.org/10.1016/j.system.2014.07.009

Li, T., & Wharton, S. (2012). Metadiscourse repertoire of L1 Mandarin undergraduates writing in English: A cross-contextual, cross-disciplinary study. *Journal of English for Academic Purposes, 11*(4), 345–356. https://doi.org/10.1016/j.jeap.2012.07.004

Matsuda, P. K. (2001). Voice in Japanese written discourse: Implications for second language writing. *Journal of Second Language Writing, 10*(1–2), 35–53. https://doi.org/10.1016/S1060-3743(00)00036-9

Matsuda, P. K., & Tardy, C. M. (2007). Voice in academic writing: The rhetorical construction of author identity in blind manuscript review. *English for Specific Purposes, 26*(2), 235–249. https://doi.org/10.1016/j.esp.2006.10.001

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Sage.

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

Nelson, N., & Castelló, M. (2012). Academic writing and authorial voice. In M. Castelló & C. Donahue (Eds.), *University writing: Selves and texts in academic societies* (pp. 33–52). Emerald.

Nesi, H., & Gardner, S. (2012). Genres across the disciplines: Student writing in higher education. Cambridge University Press.

O’Donnell, M. (2019). *UAM Corpus tool: Text annotation for the 21st century*. http://corpus tool.com/index.html

Peng, J. (2019). Authorial voice constructed in citation in literature reviews of doctoral theses: Variations across training contexts. *Journal of English for Academic Purposes, 37*, 11–21. https://doi.org/10.1016/j.jeap.2018.11.001

Petrič, B. (2012). Legitimate textual borrowing: Direct quotation in L2 student writing. *Journal of Second Language Writing, 21*(2), 102–117. https://doi.org/10.1016/j.jslw.2012.03.005

Sheng, G., & Zhou, R. (2011). Yingyu zhuanye benke biye lunwen cunzai wenti de fenxi he duice [Analysis of problems and countermeasures in graduation theses of English majors]. *Shanghai Haiyang Daxue Xuebao*, 20*(4), 636–640.

Stock, I., & Eik-Nes, N. L. (2016). Voice features in academic texts: A review of empirical studies. *Journal of English for Academic Purposes, 24*, 89–99. https://doi.org/10.1016/j.jeap.2015.12.006

Tardy, C. M. (2012). Current conceptions of voice. In K. Hyland & C. Sancho-Guinda (Eds.), *Stance and voice in written academic genre* (pp. 34–48). Palgrave MacMillan.

Tardy, C. M. (2016). Voice and identity. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 349–364). Walter de Gruyter.

Tardy, C. M., & Matsuda, P. K. (2009). The construction of author voice by editorial board members. *Written Communication, 26*(1), 32–52. https://doi.org/10.1177/0741088308327269

Thompson, P. (2012). Achieving a voice of authority in PhD theses. In K. Hyland & C. Sancho-Guinda (Eds.), *Stance and voice in written academic genre* (pp. 119–133). Palgrave MacMillan.

Xu, L., & Zhang, L. J. (2019). L2 doctoral students’ experiences in thesis writing in an English-medium university in New Zealand. *Journal of English for Academic Purposes, 41*, Article 100779. https://doi.org/10.1016/j.jeap.2019.100779

Yoon, H.-J. (2017). Textual voice elements and voice strength in EFL argumentative writing. *Assessing Writing, 32*, 72–84. https://doi.org/10.1016/j.asw.2017.02.002

Zhang, F., & Zhan, J. (2020). Understanding voice in Chinese thesis writing in an English-medium university in New Zealand. *English Text Construction, Article 100844*. https://doi.org/10.1016/j.jslw.2020.100844

Zhao, C. G. (2013). Measuring authorial voice strength in L2 argumentative writing: The development and validation of an analytic rubric. *Language Testing, 30*(2), 201–230. https://doi.org/10.1177/0141938812456965

Zhao, C. G. (2017). Voice in timed L2 argumentative essay writing. *Assessing Writing, 31*, 73–83. https://doi.org/10.1016/j.asw.2016.08.004

Zhao, C. G., & Llosa, L. (2008). Voice in high-stakes L1 academic writing assessment: Implications for L2 writing instruction. *Assessing Writing, 13*, 153–170. https://doi.org/10.1016/j.asw.2008.10.003