Citations and the Nature of Cited Sources: A Cross-Disciplinary and Cross-Linguistic Study

Guihua Wang¹ and Guangwei Hu²

Abstract
Extant scholarship on citation has examined a limited number of citational features, adopted disciplinary and ethnolinguistic perspectives disjunctively, and paid little systematic attention to the nature of cited sources. Drawing on appraisal theory, the present study investigated the nature of cited sources, namely personalization (i.e., whether humans are foregrounded as a cited source) and identification (i.e., whether and how the cited sources are identified), to understand their dialogic functionality in knowledge making. We analyzed citations in a corpus of 84 research articles sampled from two disciplines and two languages. Greater citation-based dialogic contraction was found in the medical articles than in the applied linguistic articles, whereas the cross-linguistic contrasts revealed a mixed picture. The differences are explained in terms of divergent epistemologies, cultural beliefs, discursive practices, institutional settings, and co-patterning of different citation features.

Keywords
citation, personalization, identification, appraisal, dialogic functionality

Introduction
Drawing on Bakhtin’s (1981) notion of dialogue, Kristeva (1986) made the insightful observation that “Any text is constructed as a mosaic of quotations; any text is the absorption and transformation of another” (p. 37). Academic texts have the most explicit form of intertextual borrowing, namely citations, which constitute a polyphonic phenomenon par excellence and serve the dialogic functions of creating a conversation with previous and anticipated speakers/writers (Fairclough, 1992; Fløttum et al., 2006; Hyland, 2002; Swales, 2014). A multifunctional, sophisticated, and fundamental academic practice, citations provide cogent evidence of knowledge accumulation and an interactive avenue for academic communication. Rhetorically, they function to credit forerunners, acknowledge intellectual debt, demonstrate familiarity with a knowledge domain, locate one’s work in a continuous and collective endeavor, and support authorial claims (Gilbert, 1977; Hyland, 1999; Latour, 1987).

Given these important functions, citation practice has attracted growing attention in information science, history and sociology of science, and applied linguistics (White, 2004). In our home discipline (i.e., applied linguistics), scholarship on citation has concentrated primarily on citation density (Fløttum et al., 2006; Hyland, 1999, 2002; Thompson & Tribble, 2001), writer stance enacted in reporting verbs, tenses, and reporting structures (Charles, 2006; Davidse & Vandelanotte, 2011; Hawes & Thomas, 1997; Hyland, 2002; Kwon et al., 2018), ways of incorporating cited propositions and authors (Hu & Wang, 2014; Hyland, 1999; Peng, 2019; Swales, 1990), motivations and functions of citing behavior (Petric & Harwood, 2013; Schembri, 2009), and appropriate use of sources (Harwood & Petric, 2012; Pecorari, 2006; Shi, 2011; Wette, 2017). Previous research, however, has not systematically investigated how features of cited sources, such as personalization (i.e., whether humans are foregrounded as cited sources) and identification (i.e., whether and how the cited sources are identified), serve dialogic functions and achieve rhetoric persuasion. A few extant studies that did examine the nature of cited sources adopted either a cross-disciplinary (i.e., comparing the use of source features between disciplines) or a cross-linguistic perspective (i.e., comparing the use of source features in one language with another) disjunctively. These problems are addressed in the present study.

¹China West Normal University, Nanchong, P.R. China
²The Hong Kong Polytechnic University, Hunghom, Hong Kong

Corresponding Author:
Guihua Wang, College of Foreign Language Education, China West Normal University, Room 805, Comprehensive Building, 54 Lianchi RD, Shunqing, Nanchong 637000, P.R. China.
Email: samuel-w@qq.com

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Of the few studies that examined several source features from a cross-disciplinary perspective, Charles (2006) investigated the foregrounding of human elements in citing sentences. She found that the social science of politics and international relations deployed human subjects more frequently than the material sciences, which preferred non-human and / subjects instead. The difference was attributed to an epistemological contrast between the more personal, discursive approach to and the more objective, experiment-based construction of knowledge in the respective disciplines. Examining the age of cited sources, several studies (Bloch & Chi, 1995; Glänzel & Schoepflin, 1995; Pecorari, 2006) found that writers in sciences tended to use more recent citations than those in the humanities and social sciences due to differences in knowledge growth pattern, publishing speed, and discipline-specific expectations. Investigating the publication format of cited sources, Becher and Trowler (2001) showed that urban fields (i.e., a metaphorical characterization of sciences) favored articles, whereas rural fields (i.e., soft disciplines and hard, applied ones) favored book publications, due to the amount of space needed for reporting projects of different scales, publishing speed, and convenience in access to sources (Pecorari, 2006; Schembrì, 2009). Focusing on the status of cited sources, Moed and Garfield (2004) found that authoritative documents were cited less frequently in applied, technical, and engineering sciences compared with basic and pure sciences, because articles in the former disciplines aim “to provide a ‘blueprint’ for the reader to build apparatus or instruments which are intended to perform certain stated functions [and] do not need to use references to help to demonstrate their validity” (Gilbert, 1977, p. 117). Although these studies have shed much light on the features of cited sources, they examined those features largely in a discrete fashion and it is unknown whether those findings hold true across ethnolinguistic contexts (i.e., involving different cultural and linguistic backgrounds).

Another group of studies investigated a limited number of source features from a cross-linguistic perspective. Bloch and Chi (1995), for example, found that Chinese writers used older citations than their English counterparts because of their Confucian respect for the old and limited access to new publications in then China. Examining the language medium of cited sources, Biala (1986) found 97% of the cited sources were written in English in four cross-cultural psychology journals. Lillis et al. (2010) also found a privileging of English-medium sources in Hungary, Slovakia, Spain, and Portugal, due to a biased evaluation system anchored in who gets cited, where and by whom. Notably, these studies investigated a few features of cited sources separately and, with the rare exception of Bloch and Chi (1995), have not taken into consideration researchers’ disciplinary backgrounds.

In sum, most of the above-mentioned studies adopted cross-disciplinary or cross-linguistic perspectives in isolation, neglecting the fact that citation is a literacy practice situated in both small disciplinary cultures and big national cultures (Atkinson, 2004). In addition, the source features were examined in a disparate manner, without investigating their dialogic functions. To address these limitations, this study aimed to examine the joint and respective influences of disciplinary and ethnolinguistic backgrounds on the representation of cited sources in research articles, namely the personalization and identification of cited sources. Specifically, this study sought to answer the following questions:

1. Are there differences/similarities in the various forms of personalization and identification employed in research articles from a soft and a hard discipline?
2. Are there differences/similarities in the various forms of personalization and identification employed in research articles written in two languages?
3. Is there a discipline/language interaction on the various forms of personalization and identification employed in research articles?

To answer those questions, we examined the nature of cited sources from a double contrastive (i.e., cross-disciplinary and cross-linguistic) perspective (Fløttum et al., 2006) by adopting an appraisal-informed integrative framework as outlined below.

**Analytical Framework**

Synthesizing discrete descriptions of citation practices in previous research and focusing on the dialogic functionality of citations, Coffin (2009) proposed an integrative framework based on appraisal theory (Martin & White, 2005; White, 2003), an extension of systemic functional linguistics in the interpersonal dimension. The engagement system of appraisal theory is an “effective analytical tool for the analysis of stance-taking techniques and ways of establishing interpersonal relationships with readers” (Loghmani, 2020). The system takes a dialogical perspective and regards writing as a dialogue (Bakhtin, 1981) that involves preceding authors, intended audience, and the current writer who is hosting the conversation. Functionally speaking, the writer can capitalize on linguistic resources to accommodate anticipated alternatives, rejections, and criticism from the reader (i.e., to construct a dialogically expansive space) or to preempt possible alternatives and oppositions (i.e., to create a dialogically contractive space). The linguistic resources to realize such dialogic expansion and contraction in citations were characterized by Coffin (2009) into three broad dimensions: writer stance, textual integration of source, and the nature of source. With the first two dimensions reported elsewhere, this article focuses on the third dimension that classifies the nature of cited sources into personalization and identification (Figure 1).
Personalization

In Coffin’s (2009) framework, personalization zooms in on “whether the human dimension of the source is foregrounded or not” (p. 174) and consists of three types: human, non-human, and abstract human citations. In a human citation the cited proposition is attributed to one human being (example 1), a group of humans (example 2), or an institute consisting of humans (example 3). Rhetorically speaking, a proposition from a human source is inherently subjective and likely to invoke perceptions of human agency as a contamination of objectivity. In the three examples mentioned above, the animate locutions—Dickinson, Sinclair and Stubbs, and The South Korean Ministry of Education—foreground humans as sources of the cited propositions and frame the reported information as subjective, intuitive, and personal opinions/acts that are potentially debatable, contentious, and objectionable. In this way, a dialogic space is opened up for putative readers.

(1) Dickinson (1995: 165-174) argues that autonomy is both a learning attitude and an independent learning ability.

(2) Sinclair (1991) and Stubbs (1996) suggest that all lexical items have collocations.

(3) The South Korean Ministry of Education (1997) specified these words as objects of instruction by the end of the tenth grade.

Complicating the picture is that animate and inanimate words often co-occur as a cited source. As shown in examples (6) and (7), the abstract word theory and the inanimate expression empirical data are used as information sources to project objectivity of the reported information, only to be offset by the co-presence of human agents Hoey and Antón-Méndez.

(6) Hoey’s (2005) theory of Lexical Priming, which specifically addresses the phenomenon of collocation, argues that . . .

(7) Antón-Méndez (2010a)’s empirical data also account for . . .
Identification: Naming

The second category of source nature concerns how a cited source is identified and falls into three sub-categories, that is, naming, grouping, and status. Naming refers to whether the cited source’s publication information is fully specified. The publication details of a named citation are usually specified via canonical citational forms, viz. in-text parenthetical citations or bracketed numerical references pointing to footnote/endnote entries (examples 1–7). By contrast, the publication details of an unnamed citation are not clearly specified in standardized citation forms, a phenomenon called general reference in Charles (2006). Although the publication information is not fully specified, unnamed citations clearly signal that the referenced information is not the writer’s own view, as shown in examples (8) and (9), through the use of indefinite nouns/pronouns (some people), reporting verbs (argued and perceived), and/or quotation marks. For unnamed citations, we only included those that were clearly signaled, yet not in canonical citational forms, by the writer as re-use of a source from somewhere. Unacknowledged textual borrowing and appropriation, either legal or illegal, and the murky distinction between what is common knowledge that does not need to be cited and what is non-common knowledge that should be cited (see Shi, 2011) are not the main concerns of the present study. As regards rhetorical effects, named citations are often loaded with dialogically contractive force because a proposition from a fully documented, retrievable and verifiable source appears to be meticulous, evidence-based, and compelling than otherwise. Conversely, unnamed citations help to expand the dialogic space because an unspecified and unverifiable source tends to be untrustworthy, objectionable, and devalued in academia.

Identification: Status

Status concerns the relative standing of cited sources. Rhetorically, a proposition from an authoritative and credible source usually carries more persuasive force, leaves little room for reader refutation, and, therefore, is dialogically contractive. In contrast, a proposition from a less well-known source is more susceptible to readers’ questioning, hence dialogically expansive. The question is how to determine the status of a cited source. According to Coffin (2009), generally “a source is invested with high status when the human producer is fully acknowledged and publication details are supplied” (p. 177). Studies in information sciences have also tried to operationalize this concept. For example, Wu et al. (2012) operationalized source status in terms of author status and journal status, with the former determined by the ranking of the cited author’s affiliation and the latter based on the impact factor of the cited journal. Such attempts are far from being satisfactory, because it is not infrequent for a cited

Example (10): Our data demonstrating disparities in herpes zoster vaccine uptake, consistent with other national data, whereas MLC and MLS were reported to discriminate nonadjacent school levels only (e.g., Cooper, 1976; Yau, 1991).

Example (11): Numerous studies in several states have confirmed that GDL is associated with reductions in crashes involving young teens.

Example (12): From the psycholinguistic perspective, it is widely proposed that formulaic language offers processing advantages.
scholar to be an authority in a specialized area but affiliated with a less well-known university or for a seminal article to have been published in a journal without a high impact factor. Arguably, such operationalizations of academic status would be more problematic in cross-disciplinary and cross-linguistic comparisons given the existence of an audience effect—"whether sources are accorded higher or lower status will largely depend on the particular discourse community or institutional setting within which they are used or referenced" (Coffin, 2009, p. 177).

In brief, Coffin’s (2009) framework characterizes the multifaceted nature of cited sources coherently and comprehensively, mapping nuanced discursive differences to two rhetoric functions, viz. dialogic expansion and dialogic contraction. It provides an excellent tool for making cross-disciplinary and cross-linguistic comparisons of citation practices. In this study, we adopted Coffin’s analytical framework except for the subcategory of the status of a cited source due to the difficulty in its reliable operationalization discussed above.

Method

Corpus

To answer the research questions, we constructed a corpus comprising four parallel subcorpora (see Table 1). For cross-disciplinary comparisons, we chose applied linguistics and general medicine based on their traditional membership as a typical soft and a hard discipline, respectively (Fløttum et al., 2006; Hyland, 1999; Salager-Meyer, 1999). For cross-linguistic comparisons, Chinese and English were selected because the former has the most native speakers and the latter is a premier medium (Swales, 1987) of global academic communication.

Every effort was made to achieve maximum equivalence in terms of journal prestige, article type, the number of sampled articles, and publication time. Methodological equivalence in such crucial parameters is key to establishing tertium comparationis (Connor & Moreno, 2005), a solid basis for reliable and valid comparisons between disciplines and languages. Specifically, the selection of source journals was based on specialists’ nominations and journal citation reports: SCI (Science Citation Index); (b) SSCI (Social Sciences Citation Index); (c) CSCD (Chinese Science Citation Database); and (d) CSSCI (Chinese Social Sciences Citation Index). In this way, three leading journals were selected for each of the four subcorpora. From each journal seven empirical articles published in the same year were randomly sampled. For each article, we excluded the front matter (i.e., titles, authors, and abstracts/summaries), figures, tables, captions, explanatory footnotes, and back matter (i.e., acknowledgements, explanatory endnotes, author notes, references, and appendices), not only because these parts are conventionally excluded from analysis in previous studies (e.g., Fløttum et al., 2006; Hyland, 1999) but also because citations, the foci of our study, are sparse in those parts. As summarized in Table 2, our finalized corpus totaled over 353,000 words and comprised 84 research articles sampled from 12 journals published in two languages and two disciplines.

Data Coding

All the 84 articles were imported into UAM CorpusTool (version 2.8.7) and analyzed with Coffin’s (2009) framework presented earlier. In analyzing citations, we excluded: (a) medicine or medical equipment producers attributed to in parentheses, (b) mentions of commonly known instruments unless they were specially acknowledged by the writer, (c) internal references that point to other parts of the same article, and (d) testimonies from interviewees.

We counted the number of citations in the following ways: (a) when a cited proposition was attributed to a single source, it was counted as one citation; (b) where two or more independent sources were cited for one proposition, it was still counted as a single citation; (c) when a single sentence contained multiple sources cited for distinct propositions, multiple citations were counted; (d) if the referencing of a proposition motivated by apparently the same rhetorical function ran through several sentences, it was counted only as one citation unless the same source was presented more than once in parentheses or via superscript numbers; and (e) a second-hand citation (i.e., the cited in type) was counted together with the attributed primary source as a single citation.

To ensure coding reliability, the first author and a graduate student who was bilingual and familiar with appraisal theory was involved in two stages to code eight articles, that is, about 10% of the whole dataset. In the first intercoding stage, the two coders, after a training session, worked independently to identify the target features in four articles randomly selected from the four subcorpora. In the second stage, the two coders resolved discrepancies found in the first stage through discussion before starting to code another four articles independently. Due to rising convergence in understanding, the inter-coder agreements achieved in this stage were $\kappa=.78$ for personalization, $\kappa=.81$ for naming, and $\kappa=.68$ for grouping. These reliability indices are very good, indicating substantial or excellent inter-coder agreement (Landis & Koch, 1977). Given the acceptable inter-coder statistics, the first author proceeded to code all the remaining articles.

Data Analysis

To facilitate statistical comparisons among articles of varying lengths, frequencies of citation features were normalized by 1,000 words for all articles. The normalized frequency of a citation feature in each article was obtained by dividing the total occurrences of the citation feature by
SAGE Open

the total number of words in the article and then multiplying the quotient by 1,000 words. Unlike other studies that used zi (i.e., Chinese character) as the unit of standardization when quantifying Chinese written texts, we used ci (i.e., Chinese word) for that purpose. Our chosen unit is more justifiable because zi corresponds to morpheme in English (Wang, 1985; Zhang, 2007), but ci is a lexical unit in modern written Chinese equivalent to word in English. Thus, the adoption of zi as the unit of standardization inflates the results of normalization because a majority of Chinese words contain more than one character. To count the Chinese words in the subcorpora of Chinese articles, the lexical analysis program ICTCLAS (Institute of Computing Technology of the Chinese Academy of Sciences [ICTCLAS], 2011) that claimed an accuracy rate of 98.45% was used to automatically parse and count Chinese words. To check the accuracy of ICTCLAS’s output, one of us manually parsed the words in two Chinese articles, with each being randomly selected from the subcorpora of CAL and CMS. A comparison between the automatic and the manual parsing of the two articles showed that the program achieved an accuracy rate of 95% in the applied linguistics article and 91% in the medical article. Given the acceptable levels of accuracy, the frequency normalization in Chinese was based on the word counts yielded by the ICTCLAS program.

Several 2 (Chinese vs. English) × 2 (applied linguistics vs. medical sciences) between-subjects ANOVAs were run on the normalized frequency data to determine if there were statistically significant cross-disciplinary and cross-linguistic differences in the incidence of the target citational features, with the alpha set at .05 (two-tailed) for all the statistical tests. To exemplify the quantitative analyses, salient patterns of source use and prototypical examples were identified through close and iterative readings of all instances of the identified citations and presented after the statistical results in the following section.

Findings

Table 3 summarizes the descriptive statistics for the seven source sub-features by discipline and language. The mean frequencies of most features fluctuated markedly between the disciplines and/or languages. Table 4 presents the results of the two-way ANOVAs run on each sub-feature, indicating both significant disciplinary and ethnolinguistic effects on the source sub-features. The following sections present the results of the inferential tests and qualitative analyses by source sub-feature.

Personalization

The ANOVA results showed that both discipline and language had a significant main effect on the frequencies of all personalization features in the corpus. First, with respect to human citations, discipline had a significant main effect, $F(1, 80) = 61.688, p < .001, \eta^2 = .435$, with the applied linguistics articles ($M = 2.82, SD = 1.39$) using human citations about four times as frequently as the medical articles ($M = 0.77$, $SD = 1.13$).
Wang and Hu

The obtained effect size ($\eta^2_p = .44$) indicated that discipline accounted for 44% of the variance in the incidence of human citations. It exceeded the criterial value suggested by Cohen (1988) for a large effect (i.e., $\eta^2_p = .01$ for a small effect, $\eta^2_p = .06$ for a medium effect, and $\eta^2_p = .14$ for a large effect). Language had a much smaller but still significant main effect, $F(1, 80) = 4.212, p < .05, \eta^2_p = .050$, indicating that the Chinese articles ($M = 2.06, SD = 1.53$) used significantly more human citations than the English articles ($M = 1.53, SD = 1.64$). No significant discipline/language interaction effect was detected.

Second, similar patterns in the distributions of abstract human citations between the disciplines and languages were identified. Specifically, a significant main effect of discipline was found on the use of abstract human citations, $F(1, 80) = 35.368, p < .001, \eta^2_p = .307$, indicating that the applied linguistics articles ($M = 1.38, SD = 1.07$) used significantly more such citations than the medical articles ($M = 0.33, SD = 0.50$). The effect size was very large. The main effect of language was also significant but had a medium effect size, $F(1, 80) = 7.075, p < .05, \eta^2_p = .081$, with more abstract human citations found in the Chinese articles ($M = 1.09, SD = 1.06$) than in the English articles ($M = 0.62, SD = 0.85$).

The following extracts demonstrate how human and abstract human citations add a subjective flavor to the reported propositions/acts, opening up the space for reader involvement in knowledge negotiation. The researcher’s name 王奇民 (Wang Qimin) in example (14), the animate noun 学者 (scholars) in example (15), and the institute consisting of researchers The American College of Gastroenterology in example (16) all foreground the human dimension of the sources, infusing the reported information with subjectivity and negotiability. Likewise, in examples (17) and (18), the human post-modifier Brenner et al and pre-modifier Bardovi-Harlig和Dörnyei (1998) 的 (Bardovi-Harlig and Dörnyei’s) also display a strong presence of human agency, moderated by the objectification stemming from the use of the inanimate headwords 研究 (research). These examples illustrate that both human and abstract human sources can charge the cited proposition with subjectivity/negotiability and thus contribute to a dialogically expansive rhetoric.

(14) 王奇民 (2002) 调查了制约大学英语教学效果的
因素。（FLW1/CAL）

Wang Qimin (2002) investigated the factors that constrained the effects of college English instruction.

(15) 因此许多学者认为L2中也不存在这种反馈加工
（如Antón-Méndez 2010a）。（MFLS/CAL）

Therefore, many scholars think there is no such feedback processing in L2, either (e.g., Antón-Méndez 2010a).

### Table 3. Descriptive Statistics (citations per 1,000 words).

| Classification | Feature       | Applied linguistics | General medicine |
|----------------|---------------|---------------------|------------------|
|                |               | Chinese SD | English SD | Chinese SD | English SD |
| Personalization| Human         | 2.90 1.38   | 2.74 1.43   | 1.22 1.19   | 0.31 0.61   |
|                | Abstract human| 1.66 1.13   | 1.10 0.95   | 0.52 0.61   | 0.13 0.27   |
|                | Non-human     | 2.65 1.74   | 4.61 1.57   | 4.84 2.11   | 8.30 3.19   |
| Naming         | Named         | 7.07 3.35   | 8.18 2.74   | 6.17 1.95   | 8.67 3.02   |
|                | Unnamed       | 0.14 0.18   | 0.27 0.27   | 0.41 0.80   | 0.08 0.23   |
| Grouping       | Individual    | 5.53 3.12   | 5.53 2.21   | 4.37 1.95   | 5.26 2.58   |
|                | Collective    | 1.68 1.26   | 2.91 1.15   | 2.22 1.25   | 3.49 1.20   |

### Table 4. Results of Two-Way ANOVAs by Sub-Feature.

| Feature | Sub-feature | Discipline | Language | Discipline × language |
|---------|-------------|------------|----------|-----------------------|
|         |             | $F(1, 80)$ | $p$ Value | $\eta^2_p$ | $F(1, 80)$ | $p$ Value | $\eta^2_p$ | $F(1, 80)$ | $p$ Value | $\eta^2_p$ |
| Personality | Human       | 61.688     | .000      | .435      | 4.212     | .043      | .050      | 2.054     | .156      | .025      |
|           | Abstract human | 35.368     | .000      | .307      | 7.075     | .009      | .081      | 0.243     | .623      | .003      |
|           | Non-human   | 36.183     | .000      | .311      | 30.542    | .000      | .276      | 2.368     | .128      | .029      |
| Naming   | Named       | 0.106      | .746      | .001      | 8.630     | .004      | .097      | 1.276     | .262      | .016      |
|           | Unnamed     | 0.168      | .683      | .002      | 1.157     | .285      | .014      | 5.444     | .022      | .064      |
| Grouping | Individual  | 1.719      | .194      | .021      | 0.667     | .417      | .008      | 0.656     | .421      | .008      |
|          | Collective  | 4.400      | .039      | .052      | 22.251    | .000      | .218      | 0.008     | .929      | .000      |

$SD = 1.04$.
The American College of Gastroenterology recommends that screening in African American individuals begin at age 45 years because of . . . consecutively higher mortality in comparison with white individuals.22 (JAMA/EMS)

A study by Brenner et al.6 revealed that the same incidence of CRC and CRC mortality occurred in women 4 to 8 years later than in men. . . (JAMA/EMS)

According to Bardovi-Harlig and Dörnyei’s (1998) research, even when learners who collected and analyzed the data, and in example (20) the inanimate word rather than the researchers who collected and analyzed the data were used in the applied linguistics articles (M=3.65, SD=1.91) than in the medical articles (M=6.57, SD=3.19). A significant main effect was also found for language, F(1, 80)=30.542, p<.001, η_p²=.311, indicating that significantly fewer non-human citations were used in the English articles (M=6.45, SD=3.10) using more such grammatical competence in using target language. . . (MFL1/CAL)

Third, the ANOVA on non-human citations revealed a significant main effect of discipline, F(1, 80)=36.183, p<.001, η_p²=.311, indicating that significantly fewer non-human citations were used in the applied linguistics articles (M=3.63, SD=1.91) than in the medical articles (M=6.57, SD=3.19). A significant main effect was also found for language, F(1, 80)=30.542, p<.001, η_p²=.276, with the English articles (M=6.45, SD=3.10) using more such citations than their Chinese counterparts (M=3.75, SD=2.21). The effect sizes were large in both cases.

The following extracts showcase how non-human citations were typically used in our corpus to achieve dialogic contraction. Example (19) foregrounds the lifeless word data rather than the researchers who conducted the studies being backgrounded and relegated to less prominent positions in the brackets. In example (21), the same effect is achieved by directly reporting the cited content without mentioning the cited authors in the body of the citing sentence. All these nuanced discursive choices subsume human agency to highlight the non-human dimension of sources, facilitate the construction of factuality and objectiveness, and incline readers to read factually.

Observational, multicentre data suggest that this method might reduce the need for mechanical ventilation.15 (LANCET6/EMS)

However, many studies indicate that most people, even experienced orators, show different degrees of anxiety (McCroskey 1977: 78; Hahner, Scholoff & Salisch 2002: 12). (FLW2/CAL)

Lack of access to specialty care services at community-based health centers is a major problem, particularly for uninsured patients.18,19 (NEJM4/EMS)

Ortega (2003) provides strong confirmation that current research continues to employ these same two measures. . . (TQ1/EAL)

Shu Dingfang (2002) points out that foreign language teachers at the tertiary level should have knowledge of relevant languages, knowledge of language learning, and knowledge of pedagogy, all of which can only be acquired through learning and researching. (MFL7/CAL)

Injecting drug users (IDUs) have a much higher risk of morbidity and mortality than does the general population.1 (LANCET2/EMS)

The ANOVA on unnamed citations revealed no significant main effect for discipline or language. The discipline/language
interaction, however, was found to be significant with a moderate effect size, $F(1, 80) = 5.444$, $p < .05$, $\eta^2_p = .064$. The Chinese medical articles ($M = 0.41$, $SD = 0.80$) used unnamed citations most frequently, followed by the English applied linguistics articles ($M = 0.27$, $SD = 0.27$), the Chinese applied linguistics articles ($M = 0.14$, $SD = 0.18$), and the English medical articles ($M = 0.08$, $SD = 0.23$). Because the mean frequencies of unnamed citations were quite low across the sub-corpora and since the identified significant language effect on named citations was likely to be an artifact, the observed discipline/language effect should be treated with caution until it is confirmed by further evidence.

Unlike named citations, unnamed citations tend to open up the dialogic space and were realized in our corpus by the use of:

(a) indefinite pronouns/collective nouns (e.g., proponents, people, and researchers), abstract opinion nouns (e.g., opinions, assumption, views, and consensus), and names for theoretical approaches (e.g., cognitive linguistics, Marxists, and Aristotelians), as illustrated by examples (25) to (28); (b) reporting verbs in their passive forms in English or the Chinese passive sentence structure 被 to omit sources, such as examples (29) and (30); (c) reporting verbs in zero-subject sentences in Chinese as in example (31); and/or (d) quotations, as in example (33). Of note, zero-subject sentences are grammatically correct and typical of the topic-prominent Chinese language, whereas they are grammatically incorrect in the subject-prominent English language, where the subject is obligatory except in a few sentence structures (Li, 2010). As shown in example (31), the reporting verb 认为 (think) is used together with the adverb 目前 (currently) to pass off the reported idea as a general view among specialists rather than the writer’s own opinion. Despite varied surface forms, unnamed citations tend to be dialogically expansive in that they usually provide sweeping claims without the support of specified sources, increasing the debatability of the cited propositions.

(25) Proponents argue that providing this type of information directly to consumers may result in improved compliance with health-screening practices and more healthful lifestyle choices. (NEJM1/EMS)

(31) 目前认为热性惊厥导致神经损伤的机制中海马是一个重要环节... (AMUSTH3/CMS)

(26) 一般观点认为，神经损伤3～6个月内是康复的最佳时期. (CGP5/CMS)

The general opinion holds that the best period for nerves to recover is within three to six months after injury.

(27) There is thus a widespread assumption that students who have been admitted to a Swedish university ought, as a matter of course. ... (TQ3/EAL)

(28) CL [cognitive linguistics] holds that most idioms are motivated. ... (AL3/EAL)

(32) 一般认为，词汇知识包括广度知识和深度知识两个方面. (FLW3/CAL)

Generally think that vocabulary knowledge includes two aspects, namely breadth of knowledge and depth of knowledge.

(33) “费时低效、高分低能、聋哑英语”等问题引起了社会各界的广泛关注. (FLW1/CAL)

Problems like “high time investment but low efficiency, high test scores but low competence, and deaf-dumb English” have aroused concerns among people from all walks of life.

Note. The underlined parts in examples (31) and (32) are translated word-for-word to capture the zero-subject sentence structure in Chinese.

Grouping

The ANOVA on individual citations yielded no significant main effect or interaction of discipline and language. As regards collective citations, both discipline and language were found to have a significant main effect, though the interaction was non-significant. Specifically, a small but still significant effect was found for discipline, $F(1, 80) = 4.400$, $p < .05$, $\eta^2_p = .052$, with a higher instance of collective citations in the medical articles ($M = 2.86, SD = 1.37$) than in the applied linguistics articles ($M = 2.30, SD = 1.35$). A significant main effect was also found for language, $F(1, 80) = 22.251, p < .001$, $\eta^2_p = .218$, indicating that the English articles ($M = 3.20, SD = 1.20$) used collective citations more frequently than the Chinese counterparts ($M = 1.95, SD = 1.27$), with a large effect size.

The following extracts illustrate how individual citations may open up the dialogic space. In example (34), the belief held by a single scholar Kagan may sound high and dry for readers; similarly in example (35) the research act of one researcher 陶文好 (Tao Wenhao) appears idiosyncratic and lacks additional support. In comparison with a multiple-sourced proposition, the single-sourced proposition appears to be weakly supported and incurs less interpersonal cost for readers to challenge, leaving the dialogic space relatively open.

(34) However, Kagan (1992), in her synthesis of research on teacher education, believed that the effect of training is insignificant. ... (MLJ1/EAL)
By contrast, collective citations are typically charged with dialogically contractive force as demonstrated below. The reported information was discursively represented as a view shared by community members of objectivist linguistics in example (36), as a finding corroborated by a series of studies in examples (37) and (38), and as shared research acts in examples (39) and (40). In these examples, the citing writers invest heavily in the reported information by stacking and aligning with a group of sources to present a solid-looking claim that can only be challenged by a tough reader who finds the clustered citations perfunctory and misquoted (Latour, 1987). Thus, collective citations typically carry dialogically contractive force, positioning the audience to read compliantly.

Objectivist linguistic theory views idioms as arbitrary lexical units because their meaning cannot be computed or predicted on the basis ... (AL3/EAL)
Numerous studies in several states have confirmed that GDL is associated with reductions in crashes involving young teens.12,13,14,15 (JAMA5/EMS)
This study again confirmed the findings of researchers such as Chamot et al (1996), Thompson et al (1996), and Su Yuanlian et al (2003) that strategy training can improve listening comprehension. (FLW5/CAL)
More recently, scholars like Rimmer (2006, 2008), Ravid (2005), Ravid & Berman (2010), and Norris and Ortega (2009) have also problematized the application of subordination-based measures in the study of writing development. (TQ1/EAL)
Perhaps for that reason, a number of researchers (e.g., Hirsh & Nation, 1992; Hu & Nation, 2000; Laufer, 1989b) have endeavored to answer.... (TQ5/EAL)

Discussion
As reported above, the present study has located significant disciplinary and linguistic influences on the nature/features of cited sources. Cross-disciplinarily, in response to our first research question, the applied linguists used human and abstract human citations more frequently but non-human and collective citations less frequently than the medical scientists did. Cross-linguistically, with respect to our second research question, there was a higher incidence of human and abstract human citations but a lower incidence of non-human and collective citations in the Chinese articles than in the English articles. With respect to our third question, no significant discipline/language interaction effect was found for any of the source features. The observed differences are discussed below.

Disciplinary Influences
Our analyses revealed higher frequencies of dialogically expansive source features (i.e., human and abstract human citations) in the applied linguistics articles than in the medical articles, whereas the opposite patterns were found for dialogically contractive source features (i.e., non-human and collective citations). These results are consistent with our previous findings about greater citation-based dialogic expansion in applied linguistics versus greater dialogic contraction in general medicine, as reflected in citation density, writer stance, text integration, and author integration (Authors, 2014). These broad differences are linked to epistemologies—specialists’ epistemic perceptions of fundamental issues such as the scope, genesis, characteristics, structure, and growth pattern of knowledge (Audi, 1999). Such perceptions vary considerably across disciplines and are reflected in disciplinary citation practices.

Generally, hard disciplines, including general medicine, are dominated by positivist epistemologies (Hyland, 1999). The objects of study in these disciplines are “physical objects, biological systems reducible to physical ones, and processes involving these objects and systems” (Hu & Wang, 2014, p. 25). Knowledge making adopts a nomothetic approach and aims to discover objective laws on the basis of empirical facts obtained from observations or experiments utilizing vigorous research designs, standardized procedures, and accurate measurements (Hedges, 1987). Genuine knowledge is generally regarded as resulting from “the correct application of prescribed procedures” which allows nature to reveal “itself directly through scientific method” (Hyland, 1999, p. 355), whereas human intervention is often seen as a threat and therefore is undesirable and unwelcome. The results of knowledge inquiry typically take the form of formulas, theorems, and laws that are characterized by objectivity, verifiability, replicability, and strong predictability. Knowledge growth in hard disciplines is usually driven by a dominant paradigm and a convergent research agenda and, as a result, tends to be highly cumulative, linear, and centripetal (Becher & Trowler, 2001; Hedges, 1987).

Such epistemological orientations naturally call for a discursive style that emphasizes objectivity, impersonality, and certainty to reduce the dialogic space. The dialogic contraction can be achieved by the use of non-human and collective citations in the discipline of general medicine. Specifically, non-human citations, unlike human and abstract human...
citations that discursively foreground the human elements of sources, linguistically obscure human agency, “invest the sources with a greater degree of impersonality and objectivity” (Coffin, 2009, p. 185) so as to allow facts to speak for themselves without corrupting them “with personal judgement” (Hyland, 1999, p. 361), and position readers to read factually and compliantly. Similarly, collective citations befit the cumulative, linear, and centripetal pattern of knowledge growth in hard disciplines, where there are usually clusters of studies dealing with current central issues and following the same dominant research agenda and paradigm. In addition, stacking several studies in one citation represents the cited proposition as being mutually corroborated to boost its factuality, truthiness, and certainty. With a higher use of non-human and collective citations, the medicine articles in our corpus created dialogically contracted positions in a rhetoric effort to persuade readers sharing similar epistemological backgrounds and favoring such communication styles.

Soft disciplines, including applied linguistics, tend to be dominated by anti-positivist and anti-foundationalist epistemologies (Baert & Rubio, 2009). A basic assumption in soft disciplines is the existence of multiple true realities (Hu, 2018). The objects of study comprise human behaviors, experiences, actions, thoughts, feelings, etc. Thus, knowledge making in soft disciplines tends to involve numerous yet less amenable variables and holistic, critical and introspective thinking. This often results in knowledge claims taking the form of personal views, interpretations, and understandings, which are not easy to confirm or reject (Becher & Trowler, 2001; Hu, 2018). Accordingly, knowledge in soft disciplines tends to be idiographic, subjective, value-laden, contentious, and contingent on specific cultural-historical contexts (Becher & Trowler, 2001; Hyland, 1999; Latour & Woolgar, 1986). Moreover, knowledge growth in these disciplines tends to be reiterative, recursive, and heteroglossic with multiple discrete and competing research agenda/paradigms (Hyland, 2002; Meehl, 1978). As Becher and Trowler (2001) note, soft disciplines are loosely knit clusters of ideas without an articulated framework of development, and disciplinary communities are divergent and exhibit significant internal disagreement. Given such epistemological characteristics, discourse in soft disciplines tends to be subjective, idiosyncratic, argumentative, tentative, and dialogically expansive. This discourse style coheres well with a preference for human and abstract human citations that are more persuasive to readers who have heteroglossic orientations, recognize human involvement, and prefer to read critically and resistantly.

**Ethnolinguistic Influences**

As reported above, the Chinese articles had a higher incidence of human and abstract human citations but a lower density of non-human and collective citations than the English articles did. However, the results do not warrant the conclusion that there is greater citation-based dialogic expansion in the Chinese articles than in the English ones. This is because our previous study (Hu & Wang, 2014) on other citation features (including citation density, stance, author integration, and text integration) found strong overall dialogic contraction in the Chinese articles but strong dialogic expansion in the English articles. In view of our previous findings, the results of this study revealed a more complex picture of citation-based cross-linguistic differences in dialogic functionality.

The differences identified in this study can be explained by factors including co-patterning of citation features, preferences for sentence structures, and differences in cultural beliefs and (inter)national disciplinary contexts. As regards personalization, the dialogic expansion realized by a higher incidence of human and abstract human citations in the Chinese articles was counteracted by their co-patterning with endorse citations (i.e., citations where the citing writer explicitly expresses positive attitudes toward the cited sources). The following two extracts exemplify how the dialogic expansion created by human and abstract human citations is largely offset by dialogically contractive stance markers of “endorse” (renowned) in example (41) and “the most representative” in example (42). Of all the human and abstract human citations in the Chinese articles, about 40% of them co-occurred with dialogically contractive citations (mostly endorse citations), much higher than that in the English articles (32%). On the other hand, fewer human and abstract human citations in the Chinese articles (60%) than the English articles (68%) co-patterned with dialogically expansive citations (i.e., acknowledge and distance citations, where the citing writer projects, respectively, a neutral and a doubtful attitude to open up the dialogic space). Such co-patterning differences substantially reduced the dialogic expansion construed by human and abstract human citations in the Chinese articles.

(41) 著名学者McCroskey (1977: 78) 使用“交际畏惧”

**（communication apprehension）** 一词来描述。. . .

(FLW2/CAL)

_The renowned scholar McCroskey (1977: 78)_ used the term “communication apprehension” to _describe_. . . .

(42) 关于词汇深度知识的界定，Nation (1990) 的理论最具代表性。

(FLW3/CAL)

_As to defining the depth of lexical knowledge, Nation’s (1990) theory is the most representative._

In addition, the predilections for various personalization features may also be linked to cross-linguistic differences in using passive structures that are capable of removing the agent of human actions. The Chinese language is not an inflected language, and verbs in Chinese do not have past participial forms. The Chinese sentence structure consisting of _bei_ (彼), which “most resembles a typical Indo-European
passive” (Halliday, 2006, p. 349), is less often used than the English passive voice, which is pervasive in English, especially in academic writing. Thus, there would be a greater need to use a person as the subject in Chinese reporting sentences, but a higher likelihood of using inanimate words as subjects in English passive structures.

Moreover, on a macro level, a higher incidence of human and abstract human citations in the Chinese articles was probably linked to different cultural beliefs. One fundamental Chinese philosophical belief that is shared by Confucianism, Buddhism, and Taoism is ‘tianren heyi’ (天人合一), which means the harmonious oneness/wholeness of heaven (i.e., nature) and humanity (Ji, 2006). “All the basic and persistent theses of Chinese philosophy show a proclivity toward unity versus disunity, oneness versus scatteredness, identity versus difference, and continuity versus discontinuity . . . There is no real opposition in each kind of contrast” (Cheng, 2002, p. 367). Given such a belief, the agency of humans in knowledge construction is recognized and encouraged by Chinese researchers. In contrast, the Western world generally holds a dualistic and analytical view: There are “two entirely distinct ontological zones: that of human beings on the one hand; that of nonhumans on the other” (Latour, 1993, pp. 10, 11), with “a broad dichotomy, expressed at the level of epistemology or the theory of knowledge, between ‘subjectivism’ and ‘objectivism’” (Bourdieu, 1991, p. 11). Thus, the Chinese belief in oneness of nature and mankind meshes well with the use of human and abstract human citations that foreground the personal dimension of sources, whereas the Western dualistic view encourages the use of non-human citations to obliterate subjectivity from knowledge inquiry.

As regards collective citations, the higher frequency found in the English articles than in the Chinese articles is counter-intuitive. As Chinese culture is a typical collective culture that de-emphasizes individuals but values collectivism in accordance with Confucianism (Salager-Meyer, 1999), collective citations with dialogic contraction should have been preferred by the Chinese academics. However, two factors may have intervened. First, the number of studies on particular topics available for citing depends on the age and size of a discipline (Flottum et al., 2006). Because of wars and Cultural Revolution in the recent Chinese history, modern disciplines did not start to develop substantially until the Reform and Opening-up of China in the early 1980s. Thus, the age and size of most disciplines in the Chinese national context differ greatly from those in the English-medium international community. Granted, Chinese-speaking scholars have an improved access to English-medium research, but language barriers and limited access to international databases still constitute pervasive constraints for the Chinese research communities. In this light, the number of potential citable studies for Chinese scholars would be markedly smaller than those for their English-speaking counterparts. Second, English-speaking researchers may have a stronger motivation to cite multiple sources for one proposition to woo editors/reviewers of international journals, which tend to put a premium on joining the academic dialogue through relating to previous research. In brief, both the fewer potentially citable studies on relevant topics and the lower motivation for citation might have contributed to the lower frequency of collective citations found in the Chinese articles.

Conclusion

This study set out to explore the nature of cited sources within an appraisal-based framework and from a double contrastive perspective. It found several cross-disciplinary and cross-linguistic differences in the personalization and the identification of cited sources. Cross-disciplinarily, the applied linguistics articles were characterized by greater dialogic expansion because of the more frequent use of human and abstract human citations, whereas the medical articles exhibited greater dialogic contraction due to the more frequent use of non-human and collective citations. These differences are interpretable in terms of divergent and discipline-specific epistemological beliefs. Cross-linguistically, the Chinese articles used human and abstract human citations more frequently but non-human and collective citations less frequently than the English articles did. However, these ethnolinguistic differences do not indicate greater citation-based dialogic expansion in the Chinese articles but might have arisen from cross-linguistic differences in sentence structures, cultural beliefs, and (inter)national disciplinary contexts.

These findings provide several implications for citational instruction and practice. First, the disciplinarily valued and ethnolinguistically preferred ways of source use identified in this study may inform novice researchers’ citation practice in those disciplines and ethnolinguistic settings that tend to share broadly similar epistemologies, cultural beliefs, and discursive practices. The observed patterns of source use in general medicine and applied linguistics may respectively inform source use in related disciplines, for example, biology as a hard science and educational research as a soft discipline because of shared epistemological beliefs and discourse conventions. However, their applicability to disciplines (e.g., business and management) that are not closely related to applied linguistics and general medicine need to be verified in future research. The reported patterns of source use in English and Chinese can also inform English- and Chinese-speaking novice researchers’ source use in their respective ethnolinguistic settings. Whether the same patterns of source are extrapolatable to other language pairs calls for further research because each ethnolinguistic community may have unique philosophical assumptions, epistemological beliefs, and language practices. Second, our findings can be used to raise students’ and researchers’ awareness of disciplinary and ethnolinguistic influences on
source use. The observed discipline- and language-specific patterns of source use can be explicitly incorporated into pedagogical tasks that link source features to occluded disciplinary and socio-cultural beliefs, conventions and practices. Last but not the least, our empirical findings can be drawn on to construct a repertoire of lexico-grammatical resources for source use within a coherent framework of dialogical functionality. Such resources would be invaluable to novice researchers who are learning to assume a dialogically expansive or contractive voice strategically to engage their readers in a knowledge-making dialogue for optimal persuasiveness. To conclude, the use of cited sources should be taught as a discipline-specific and ethnonlinguistically situated literacy practice with a view to achieving effective dialogic functionality.

Acknowledgments
The authors are very thankful to Lilin Choo who has kindly helped to code our data in the intercoding stage.

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) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study is funded by Sichuan Federation of Social Science Associations, SC20WY033.

ORCID iD
Guihua Wang https://orcid.org/0000-0002-6501-0643

References
Atkinson, D. (2004). Contrasting rhetorics/contrasting cultures: Why contrastive rhetoric needs a better conceptualization of culture. Journal of English for Academic Purposes, 3, 277–289.
Audi, R. (1999). Epistemology: A contemporary introduction to the theory of knowledge. Routledge.
Baert, P., & Rubio, F. D. (2009). Philosophy of the social sciences. In B. S. Turner (Ed.), The new Blackwell companion to social theory (pp. 60–80). Wiley-Blackwell.
Bakhtin, M. M. (1981). The dialogic imagination: Four essays (C. Emerson & M. Holquist, Trans.). University of Texas Press.
Baldauf, R. B. (1986). Linguistic constraints on participation in psychology. American Psychologist, 41, 220–224.
Becher, T., & Trowler, P. R. (2001). Academic tribes and territories: Intellectual enquiry and the culture of disciplines (2nd ed.). SRHE and Open University Press.
Bloch, J., & Chi, L. (1995). A comparison of the use of citations in Chinese and English academic discourse. In D. D. Belcher & G. Braine (Eds.), Academic writing in a second language: Essays on research and pedagogy (pp. 231–274). Ablex.
Bourdieu, P. (1991). Language and symbolic power (G. Raymond & M. Adamson, Trans.). Polity Press.
Charles, M. (2006). Phraseological patterns in reporting clauses used in citation: A corpus-based study of theses in two disciplines. English for Specific Purposes, 25, 310–331.
Cheng, C.-Y. (2002). An onto-hermeneutic interpretation of twentieth-century Chinese philosophy: Identity and vision. In C.-Y. Cheng & N. Bunnin (Eds.), Contemporary Chinese philosophy (pp. 366–404). Blackwell.
Coffin, C. (2009). Incorporating and evaluating voices in a film studies thesis. Writing & Pedagogy, 1, 163–193.
Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum.
Connor, U., & Moreno, A. I. (2005). Tertium comparationis: A vital component in contrastive research methodology. In D. Bruthiaux, W. G. Atkinson, W. Eggington, V. R. Grabe, & R. B. Kaplan (Eds.), Directions in applied linguistics: Essays in honor of Robert B. Kaplan (pp. 153–164). Multilingual Matters.
Davide, K., & Vandelanotte, L. (2011). Tense use in direct and indirect speech in English. Journal of Pragmatics, 43, 236–250.
Fairclough, N. (1992). Discourse and social change. Polity Press.
Fløttum, K., Dahl, T., & Kinn, T. (2006). Academic voices: Across languages and disciplines. John Benjamins.
Gilbert, G. N. (1977). Referencing as persuasion. Social Studies of Science, 7, 113–122.
Glänzel, W., & Schoepflin, U. (1995). A bibliometric study on ageing and reception processes of scientific literature. Journal of Information Science, 21, 37–53.
Halliday, M. A. K. (2006). Studies in Chinese language. Continuum.
Harwood, N., & Petrić, B. (2012). Performance in the citing behavior of two student writers. Written Communication, 29, 55–103.
Hawes, T., & Thomas, S. (1997). Tense choices in citations. Research in the Teaching of English, 31, 393–414.
Hedges, L. V. (1987). How hard is hard science, how soft is soft science? The empirical cumulativeness of research. American Psychologist, 42, 443–455.
Hood, S. (2004). Appraising research: Taking a stance in academic writing [Unpublished doctoral dissertation]. University of Technology.
Hu, G., & Wang, G. (2014). Disciplinary and ethnonlinguistic influences on citation in research articles. Journal of English for Academic Purposes, 14, 14–28.
Hu, G. (2018). Disciplinary knowledge making and academic discourse. In Y. Leung, J. Katchen, S. Hwang, & Y. Chen (Eds.), Reconceptualizing English language teaching and learning in the 21st century (pp. 553–573). Crane Publishing.
Hyland, K. (1999). Academic attribution: Citation and the construction of disciplinary knowledge. Applied Linguistics, 20, 341–367.
Hyland, K. (2002). Activity and evaluation: Reporting practices in academic writing. In J. Flowerdew (Ed.), Academic discourse (pp. 115–130). Longman.
Institute of Computing Technology of the Chinese Academy of Sciences. ICTCLAS (Version 2011). Retrieved March 3, 2012, from http://ictclas.nlpir.org.
Ji, X. (2006). Chan yu wenhua [Zen and culture]. China Yan Shi Press.
Kristeva, J. (1986). Word, dialogue and novel. In T. Moi (Ed.), The Kristeva reader (pp. 34–61). Columbia University Press.
Kwon, M. H., Staples, S., & Partridge, R. S. (2018). Source work in the first-year L2 writing classroom: Undergraduate L2 writers’ use of reporting verbs. *Journal of English for Academic Purposes, 34*, 86–96.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press.

Latour, B. (1993). *We have never been modern* (C. Porter, Trans.). Harvard University Press.

Latour, B., & Woolgar, S. (1986). *Laboratory life: The social construction of scientific facts*. Princeton University Press.

Loghmani, Z., Ghonsooly, B., & Ghazanfari, M. (2020). Engagement in doctoral dissertation discussion sections written by English native speakers. *Journal of English for Academic Purposes, 45*, 1–13.

Li, X. (2010). The influence of Chinese topic-prominent features on the acquisition of four types of English subjects. *Chinese Journal of Applied Linguistics, 33*(5), 80–94.

Lillis, T., Hewings, A., Vladimirou, D., & Curry, M. J. (2010). The geolinguistics of English as an academic lingua franca: Citation practices across English-medium national and English-medium international journals. *International Journal of Applied Linguistics, 20*(1), 111–135.

Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology, 46*, 806–834.

Moed, H. F., & Garfield, E. (2004). In basic science the percentage of “authoritative” references decreases as bibliographies become shorter. *Scientometrics, 60*, 295–303.

Pecorari, D. (2006). Visible and occluded citation features in postgraduate second-language writing. *English for Specific Purposes, 25*, 4–29.

Peng, J.-E. (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.

Petric, B., & Harwood, N. (2013). Task requirements, task representation, and self-reported citation functions: An exploratory study of a successful L2 student’s writing. *Journal of English for Academic Purposes, 12*, 110–124.

Salager-Meyer, F. (1999). Referential behavior in scientific writing: A diachronic study (1810–1995). *English for Specific Purposes, 18*, 279–305.

Schembri, N. (2009). Citation practices: Insights from interviews with six undergraduate students at the University of Malta. In L. J. O’Brien & D. S. Giannoni (Eds.), *University of reading: Language studies working papers* (Vol. 1, pp. 16–24). University of Reading.

Shi, L. (2011). Common knowledge, learning, and citation practices in university writing. *Research in the Teaching of English, 45*, 308–334.

Swales, J. (1987). Utilizing the literatures in teaching the research paper. *TESOL Quarterly, 21*, 41–68.

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

Swales, J. (2014). Citational practice in a corpus of student biology papers: From parenthetical plonking to intertextual storytelling. *Written Communication, 31*, 118–141.

Thompson, P., & Tribble, C. (2001). Looking at citations: Using corpora in English for academic purposes. *Language Learning & Technology, 5*, 91–105.

Wang, L. (1985). *Zhongguo xiandai yufa [Modern Chinese grammar]*. The Commercial Press.

Wette, R. (2017). Source text use by undergraduate post-novice L2 writers in disciplinary assignments: Progress and ongoing challenges. *Journal of Second Language Writing, 37*, 46–58.

White, H. D. (2004). Citation analysis and discourse analysis revisited. *Applied Linguistics, 25*, 89–116.

White, P. R. R. (2003). Beyond modality and hedging: A dialogic view of the language of intersubjective stance. *Text, 23*, 259–284.

Wu, L.-L., Huang, M.-H., & Chen, C.-Y. (2012). Citation patterns of the pre-web and web-prevalent environments: The moderating effects of domain knowledge. *Journal of the American Society for Information Science and Technology, 63*, 2182–2194.

Zhang, W. (2007). A morphological study of “word” and “zi” (character) in English and Chinese from a contrastive perspective. *Journal of Hubei University (Philosophy and Social Science), 34*(5), 83–89.