

# “However, we argue that ...”: The Construction of Authorial Identities in English Research Articles

Tofan Dwi Hardjanto

English Department, Universitas Gadjah Mada, Yogyakarta, Indonesia

**Abstracts**—This research investigates how professional authors construct their authorial identities in English research articles (RAs). Sixty research articles were selected from three disciplines, biology, linguistics, and medicine, published by native and non-native authors in reputable international journals. First-person references in the articles were analyzed qualitatively and quantitatively with the help of corpus linguistic methods to examine the identities they were used to express, their density, and their cross-cultural and cross-disciplinary variation. The findings showed that native and non-native authors frequently employed authorial references in their English RAs. The results showed striking similarities in the density of authorial references and the identities they express in the RAs published by both groups. However, there was a significant difference in authorial references and identities by these groups of authors across different disciplines, with the hard sciences employing significantly more frequent authorial references than the soft sciences. The findings suggest that while research publication at the highest level does not seem to affect the construction of authorial identities, disciplinary practices significantly affect authorial identity construction. It is argued that authorial references may be an essential feature of written academic interaction at the highest level. They allow authors to create identities serving as humble servants to seek cooperation and emphasize solidarity with their disciplinary communities and as knowledge originators to stress personal contribution to their respective disciplines.

**Index Terms**—academic writing, authorial identities, disciplinary variation, research article

## I. INTRODUCTION

There appear to be conflicting ideas about academic writing, especially in English. On the one hand, English academic writing has been commonly viewed as impersonal reporting, promoting directness, accuracy, and objectivity (Hacker, 2008; Lipson, 2005; Manser, 2006; Strunk Jr. & White, 2000; Taylor, 2005). This view emphasizes the impersonality of academic writing, commonly characterized by the prominent use of nominalization, agentless passive constructions, and impersonal constructions. Lachowicz (1981) argues that impersonality gives prominence to the collective responsibility of academic endeavor and objectivity and open-mindedness. Moreover, it highlights the “common share of knowledge with the community” (p. 111). This viewpoint is based on the positivist concept that academic research is simply empirical and objective. As a result, it is best presented as if a human agency is not involved (Hyland, 2001, p. 208). Thus, as an objective report, academic writing leaves no room for authorial presence, as evidenced by the common discouragement of the use of pronominal reference, particularly first-person pronouns in academic writing (see, e.g., Arnaudet & Barrett, 1984; Lester, 1993; Spencer & Arbon, 1996).

However, academic writing has recently been more than just an objective, distant and impersonal report. It has recently been seen as displaying interaction between the writer and the reader (e.g., Hyland, 2004, 2005, 2008). In this view, text production represents an external reality and involves the use of language to acknowledge, construct, and negotiate social relations (Hyland, 2005). Furthermore, as Hyland (1994) argues, rather than being factual and impersonal, effective academic writing relies on interactional components to strengthen propositional material throughout the text and make readers aware of the author’s point of view. In an attempt to demonstrate how such writer-reader interactions are realized in academic writing, researchers have examined such linguistic features as attitude markers (e.g., Mur-Dueñas, 2010), evaluative adjectives (e.g., Swales & Burke, 2003), evaluation (e.g., Hunston, 1993; Thompson & Ye, 1991), hedges (e.g., Hardjanto, 2016; Hyland, 1996, 1998; Salager-Meyer, 2011; Yang, 2013), and stance adverbials (e.g., Conrad & Biber, 2000). However, the study of authorial identities in academic writing has received little attention, especially in research articles. This lack of research is interesting as writers’ ability to create a credible portrayal of themselves and their work by aligning themselves with their communities’ socially shaped identities is critical to their pragmatic competence (Hyland, 2002a). Examples (1) and (2) below illustrate how authors use the first-person singular pronoun *I* in (1) and the inclusive first-person plural pronoun *us* in (2) to express their authorial identities in their RAs. In example (1), the single author Nikolaev uses the first-person singular pronoun *I* to indicate the purpose of his research. In (2), Gao and Skolnick use the accusative first-person plural pronoun to refer to both the writers and the reader. The authors attempt to engage the reader in the discourse in this case.

- (1) Using a sample of over 5,000 debt issues, *I* test whether firms with more extensive use of covenants in their public debt contracts exhibit timelier recognition of economic losses in accounting earnings (Nikolaev, 2010, p. 51).

(2) Let *us* consider all geometrically possible quaternary structures of dimers (Gao & Skolnick, 2010, p. 22520).

The present research examines authorial identities in English academic writing. It focuses mainly on first-person pronouns as “the most visible expression of a writer’s presence in a text” (Hyland, 2002b, p. 354). Moreover, they not only serve as “powerful metadiscursive resources in manifesting authorial identities” but also “most overtly mark attribution of stance to the writer.” More specifically, it addresses the following questions:

- a) What linguistic forms are used to refer to the authors in English research articles? Are they used differently by native and non-native authors in different disciplines?
- b) What authorial identities are these linguistic forms used to express? Are they expressed differently by native and non-native authors in different disciplines?

Thus, in answer to the questions above, this study attempts to explore authorial identities through the use of pronominal references, especially first-person pronouns, that authors use to indicate their presence in their research articles in terms of their forms and roles. An attempt was also made to examine whether there is significant variation in the use of first-person references by native and non-native authors across three different disciplines.

## II. AUTHORIAL IDENTITIES IN ENGLISH ACADEMIC DISCOURSE

The notion of authorial identities refers to the use of linguistic devices to indicate the presence of the author or writer in the text. Such authorial identities are most noticeably expressed through personal pronouns, especially first-person pronouns (Hyland, 2002b). This notion has been referred to differently in different studies. The term *authorial identity* (Flowerdew & Wang, 2015; Hyland, 2002a, 2002b; Işık-Taş, 2018; Ivanič, 1994, 1995, 1998; Matsuda, 2015; Sanderson, 2008) has also been called *authorial presence* (Chávez Muñoz, 2013; Dontcheva-Navrátilová, 2013; Hartwell & Jacques, 2014; Poudat & Loiseau, 2005), *authorial reference* (Basic & Veselica-Majhut, 2016), *authorial self* (Hryniuk, 2018), *authorial stance* (Pho, 2012), *authorial voice* (Fløttum et al., 2006; Matsuda & Tardy, 2007; Tardy, 2012; Tardy & Matsuda, 2009), *personal attribution* (Martín Martín, 2003), *personal reference* (Vladimirov, 2007), *self-mention* (Hyland, 2001) and *subjectivity* (Rentel, 2012). As the notion of authorial identity is very close to authorial presence and authorial voice, they will be used interchangeably in this research.

As expressions of authorial identities, first-person pronouns have been examined in various genres in such academic writing as research articles (e.g., Hyland, 2001; Işık-Taş, 2018; Sanderson, 2008), RA abstracts (Kim, 2015; Martín Martín, 2003), RA methods sections (Harwood, 2005b; Martínez, 2018), and theses (Hyland, 2002b; Isler, 2018; Karoly, 2009). Comparisons have also been attempted between the use of first-person pronouns in research articles written in English and those in Bulgarian, French, German and Russian (Vassileva, 1998), Croatian (Basic & Veselica-Majhut, 2016); French (Hartwell & Jacques, 2014), French, German and Italian (Rentel, 2012), German (Sanderson, 2008), Persian (Tayyebi, 2012), Polish (Hryniuk, 2018), Russian (Krapivkina, 2015), Spanish (Mur-Dueñas, 2011), and Turkish (Işık-Taş, 2018). These studies have shown that scientific disciplines (see, e.g., Harwood, 2005b; Hyland, 2001, 2002a) and writing cultures (see, e.g., Mur Dueñas, 2007; Mur-Dueñas, 2011; Vassileva, 1998) significantly affect the use of first-person pronouns as expressions of authorial identities.

It has also been shown that there are significant differences in the use of first-person singular pronouns and first-person plural pronouns. First-person singular pronouns tend to be used more commonly in soft sciences such as linguistics and economics to indicate a more visible and authoritative identity (Harwood, 2005b; Hyland, 2001, 2002a). On the other hand, first-person plural pronouns to express authorial identities are more dominant in hard sciences. Hyland (2001, 2002a) argues that such variation is attributable to disciplinary communities’ different epistemological beliefs and practices (Harwood, 2005b).

In addition to their linguistic forms and their frequency and distribution, first-person pronouns have also been investigated in terms of the functions and identities in academic discourse (e.g., Harwood, 2005c; Hyland, 2001, 2002a; Ivanič, 1998; Martínez, 2005; Sheldon, 2009; Starfield & Ravelli, 2006; Tang & John, 1999). Previous studies have shown that first-person pronouns can be used, for example, to (a) help to identify a community’s current research concern (Harwood, 2005c), (b) explain what was done (Kuo, 1999), (c) appear as a competent member of an academic community (Tayyebi, 2012), (d) highlight and express the contribution of an author (Vladimirov, 2007), and (e) create cohesion in writing (Zhang, 2012). Furthermore, Hyland (2002a) proposes a taxonomy of the discourse functions of first-person pronouns. He argues that first-person pronouns in academic discourse can be used to (a) elaborate an argument, (b) explain a procedure, (c) express self-benefits, (d) state a goal/purpose, and (e) state results or claim (see also Harwood, 2005a; Kuo, 1999).

Authorial references can also be used to express multiple identities. Building on Ivanič (1998), Tang and John (1999) propose a model of authorial identities. This model explores various authorial identities based on the degree of their presence in the text and the authority invested in their statements. The authorial identities proposed by Tang and John (1999) include the author as a) representative, b) guide, c) architect, d) recounter of the research process, e) opinion holder and f) originator. These identities range from the least powerful (the author as the representative) to the most powerful (the author as the originator). This model has since attracted numerous attempts to develop similar models of authorial identities or roles (see, e.g., Dahl, 2009; Dontcheva-Navrátilová, 2013; Hyland, 2002b, 2002a; Işık-Taş, 2018; Millán, 2010; Mur Dueñas, 2007; Sheldon, 2009; Xia, 2018). Despite the different labels, these models are essentially similar to the one proposed by Tang and John (1999). We adopt this framework in this study.

### III. METHODOLOGY

This research is corpus-based (Lee, 2008), as it is assisted with the help of a collection of digital text and computer techniques. This study rests on the principle that more valid conclusions could be reached and more generalizable results obtained with more data (Ädel, 2006). For this study, the data were collected from a corpus of English RAs from three disciplines (biology, linguistics, and medicine) written by authors affiliated with institutions in English-speaking and non-English-speaking countries and published in reputable international journals (see Appendix for the selected journals). Only empirical research articles were selected because of their vital role in producing scientific knowledge (Knorr-Cetina, 1981; see Swales, 1990, 2004). No theoretical or review papers were selected. On the basis of these criteria, 60 articles were selected randomly, 20 each from biology, linguistics, and medicine. Table 1 below presents the number of articles and words from the three fields under investigation. Out of the 20 articles from each discipline, ten were written by native authors (NAs), and another ten were written by non-native authors (NNAs). The identification of the identities of these groups, whether they are native or non-native authors, is based on their institutional affiliations (Işık-Taş, 2018; Sheldon, 2009; Xia, 2018). Because it was very difficult to find articles written by single authors in biology and medicine, it was decided to collect only articles written by multiple authors. Thus, single-authored articles were excluded “to avoid interference in the interpretation of results from other variables” (Lor és-Sanz, 2011, p. 177).

TABLE 1  
NUMBER OF ARTICLES AND WORDS IN THE CORPUS

No.	Discipline	No. of articles	No. of words	Mean
1.	Biology	20	172,688	8,634
2.	Linguistics	20	136,293	6,815
3.	Medicine	20	115,578	5,779
<b>Total</b>		<b>60</b>	<b>424,559</b>	<b>21,228</b>

First, the data were analyzed in terms of the linguistic realizations of authorial identities, i.e., first-person plural pronouns and such phrases as *the authors* and *the researchers* used as authorial references. Grammatically, first-person pronouns belong to the general category of personal pronouns categorized according to the deictic type of person (Huddleston & Pullum, 2002). Table 2 below shows English first-person pronouns grouped according to their formal grammatical categories. The data were also analyzed quantitatively using a concordance program called WordSmith Tools Version 5 (Scott, 2008) to examine the frequency and distribution of the linguistic forms used to express authorial identities. Moreover, the data were then analyzed quantitatively for possible cross-cultural and cross-disciplinary variation in the use of authorial references in the corpus. Finally, the data were analyzed in terms of the authorial identities that first-person references express in the corpus to address the second research question. The analysis of the authorial identities was based on the existing model proposed by Tang and John (1999): the author as an architect (ARC), guide (GUI), opinion holder (OPI), originator (ORI), recounter (REC) and representative (REP). Data normalization was necessary as this study examines the cross-cultural and cross-disciplinary use of authorial references in English RAs. The data were normalized per 100,000 words (p100kw). Finally, to test whether there is any significant difference in the use of authorial references cross-culturally and cross-disciplinarily, a chi-square ( $\chi^2$ ) test was run using Minitab 18 (Minitab, 2017). A significant level or *p*-value was set at 0.05, which is commonly adopted in applied linguistics (Gomez, 2002; Martínez, 2005; Mur-Dueñas, 2011; Sanjaya, 2013; Xia, 2018).

TABLE 2  
ENGLISH FIRST-PERSON PRONOUNS  
(MODIFIED FROM HUDDLESTON & PULLUM, 2002, P. 426)

Case	First-person Pronouns	
	Singular	Plural
Nominative	I	we
Accusative	me	us
Genitive	my, mine	our, ours
Reflexive	myself	ourselves

### IV. RESULTS AND DISCUSSION

This section presents and discusses the results of the data analyses in terms of authorial references that native and non-native authors use in three different disciplines, i.e., biology, linguistics, and medicine, and the identities that authorial references express.

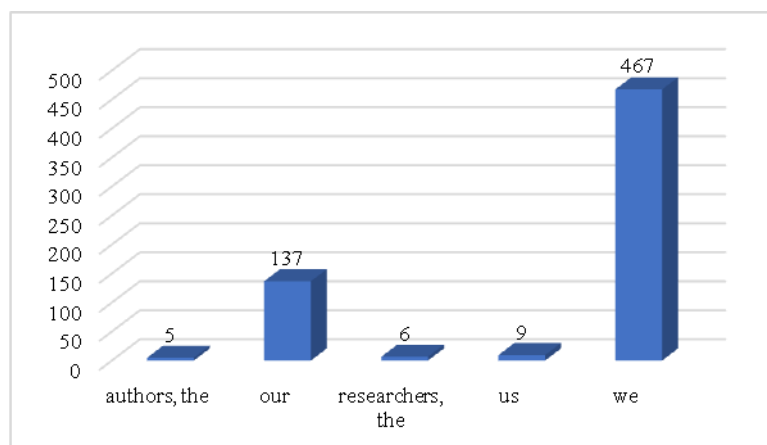


Figure 1 Overall Frequency of Authorial References in the Corpus (p100kw)

### A. Authorial References

In a corpus of 424,559 words, 2,650 tokens of authorial references were found, or 624 per 100,000 words (p100kw), or 44.17 per paper. As Figure 1 shows, the most frequently used authorial reference was the first-person pronoun *we* (74.87%), and the least was the noun phrase *the authors* (0.75%). No token was found for the noun phrase *the writers* used to refer to the RA authors. This noun phrase was not used at all in both the biology and medicine subcorpora, but in the linguistic subcorpus, it was used 19 times. However, all these tokens of *the writers* were not used to refer to the RA authors. A highly significant difference was observed in the use of authorial references in the corpus ( $\chi^2_{(4)} = 1275.04$ ,  $p < 0.001$ ). This finding suggests that, first of all, authorial identity markers are relatively frequently employed in English research articles published internationally by reputable journals. After all, academic writing is not faceless, not as formal as it is commonly believed to be (Hyland, 2001). Secondly, professional authors tend to employ the subjective first-person plural pronoun *we* when making themselves visible to their respective academic communities. This seems to indicate the application of politeness strategies in research articles (Myers, 1989). Authors commonly employ the first-person plural subjective pronoun, particularly inclusive *we*, to engage with their target readers, i.e., disciplinary community or peer researchers, in the knowledge-making process. The best means of gaining acceptance of one's claims is humbleness towards one's peers, reviewers, or the discipline in general (Hyland, 2001).

This finding is slightly different from that of Hyland (2001), who reported an average of 27.8 tokens per paper or 505 tokens of authorial references per 100,000 words (see also Işık-Taş, 2018). This slight difference may result from the search items in both studies. Hyland excluded inclusive first-person plural pronouns, whereas such pronouns were included in the search in this study. The use of the two most common authorial references *we* and *our* is exemplified below. In these and other remaining examples, BIO stands for biology, LIN for linguistics, MED for medicine, NA for native authors, and NNA for non-native authors, whereas the numbers indicate the article numbers.

- (3) Herein, *we* have revealed a particular effect of rG4s on 5'-UTR translation (BIO01NA).
- (4) *Our* results provide empirical evidence to explain the findings of Cumming et al. (2005) (LIN17NNA).
- (5) *We* defined sex-specific quarters of consumption for each type of sugary drink (MED20NNA).

As Figure 2 below shows, the use of authorial references by native and non-native authors seems to be evenly distributed (figures in parentheses indicate the number of words). As can be seen, both groups of authors used the first-person plural subjective and possessive pronouns *we* and *our* extensively with relatively the same degree of density. No statistically significant difference was observed ( $\chi^2_{(4)} = 6.734$ ,  $p = 0.151$ ). This finding suggests that socio-cultural factors and writing cultures do not seem to affect the use of authorial references in research articles published in reputable international journals.

This finding contradicts some previous studies (e.g., Abdi & Farrokhi, 2015; Dontcheva-Navrátilová 2013; Walková 2018), which reported more frequent authorial references by native authors than that by non-native authors. The primary reason for this discrepancy seems to be the readership size and reputation of the journal. In these studies, the non-native authors who published their RAs in nationally-oriented journals with a more limited readership size tend to use a significantly lower frequency of first-person pronominal references than the native speakers who published their RAs in internationally reputable journals. This is corroborated by Işık-Taş (2018), who reported that locally-oriented Turkish authors used a much lower frequency of authorial references than native and internationally-oriented Turkish authors. She found that native authors used approximately the same number of authorial references as the number of internationally-oriented Turkish authors in their RAs published internationally by reputable journals. The frequent use of authorial references in English RAs published by professional authors in international journals might be attributable to the fierce competitiveness of international publications. Moreover, their awareness of this competitiveness has led them to employ authorial references as a strategy to show their confidence in the presentation of their claims (Işık-Taş, 2018; Mur-Dueñas, 2011). This might also be caused by the strict gatekeeping in the review process in international publications.

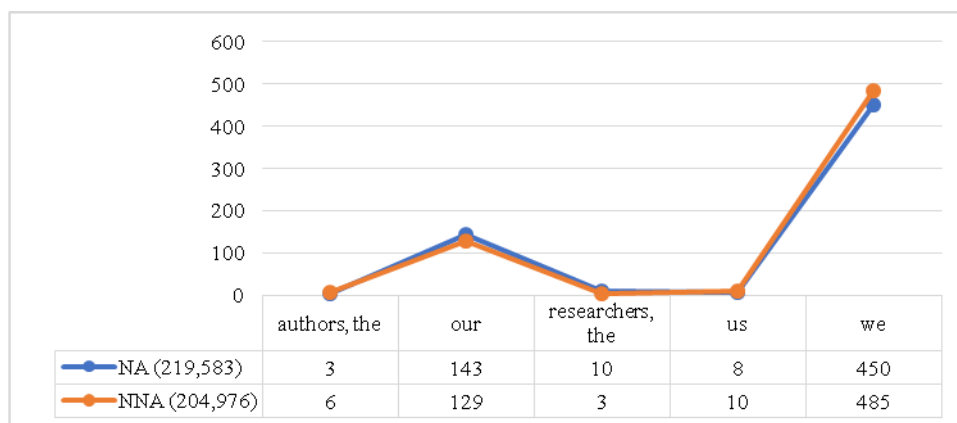


Figure 2 Frequency and Distribution of Authorial References by Author in the Corpus (p100kw)

Figure 3 below shows two crucial points about first-person references used in biology, linguistics, and medicine. First, most importantly, professional authors in the three disciplines used authorial references significantly differently ( $\chi^2_{(8)} = 167.983$ ,  $p < 0.001$ ), with scholars in biology employing the most (887 p100kw) and those in linguistics the least (357), while those in medicine in between (546). This significant disciplinary variation is hardly surprising as previous studies have reported similar findings (e.g., Harwood, 2005b; Hyland, 2001, 2002a; Hyland & Jiang, 2018). This variation may result primarily from disciplinary communities' different epistemological beliefs and practices (Hyland, 2001, 2002a). However, contrary to the traditional view of academic writing, this study has shown that academic writing in biology and medicine, which can be classified as belonging to the hard sciences (Becher, 1989, 1994), is not "author-evacuated" (Geertz, 1983) after all. Academic writing in these disciplines is traditionally regarded as impersonal, simply reflecting indisputable facts (Mulkay, 1979), and thus not inherently self-promotional (Harwood, 2005b). Yet, self-promotional use of first-person plural pronouns, especially in biology, is significantly much more frequent than that in linguistics. This appears to reflect changes in rhetorical disciplinary practices.

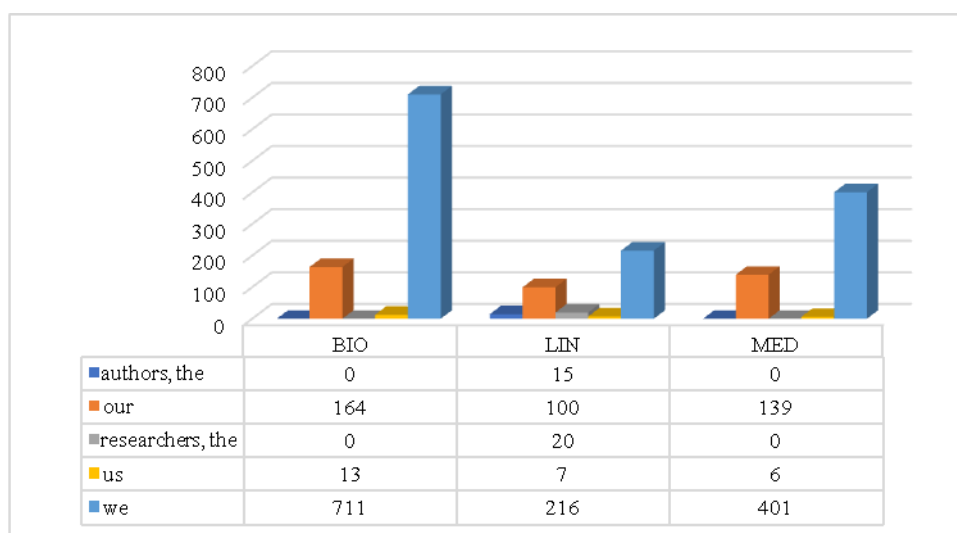


Figure 3 Frequency and Distribution of Authorial References by Discipline in the Corpus (p100kw)

Secondly, we can see that *we* and *our* are the two dominant pronominal forms employed in the three disciplines despite their variable frequencies: the plural pronoun *we* was used most frequently in the disciplines. This finding confirms Hyland (2001). Furthermore, we can also observe that only in linguistics are the noun phrases *the author* and *the researchers* used as authorial references. Despite their low frequencies, these third-person references that are used as authorial references might be characteristic of academic writing in social sciences in general and in linguistics in particular.

#### B. Authorial Identities

To answer the second research question, we analyzed the data regarding the overall frequency of first-person references used to express authorial identities, their cross-cultural variation, and their cross-disciplinary variation. Figure 4 displays the overall frequency of authorial identities in the corpus. First-person references most frequently express the author-as-recounting identity in the corpus (360 tokens p100kw or 52.67%). The identities shown in Figure 4 are presented in order of the face-threatening risk (see, e.g., Işık-Taş, 2018; Xia, 2018) or level of authorial presence (see, e.g., Sheldon, 2009; Tang & John, 1999), starting with the lowest risk author-as-representative identity (REP) to

the highest risk author-as-originator identity (ORI). The recounter identity projects the authors as storytellers, reporters, or conductors (Sheldon, 2009) of the research processes. It bears considerable face-threatening risks to the authors. This finding is consistent with Hyland (2002a), who reported that more than half of self-references, especially in biology, were used to explain a procedure. The author recounted the research processes they had gone through. However, it is significantly different from those of Sheldon (2009) and Tang and John (1999) primarily because of the disciplines in the case of Sheldon and text types in the case of Tang and John. In her linguistics and language teaching native author sub-corpus, Sheldon found that only 20% of first-person pronouns were used to express the recounter identity or conductor in her term. In this study, 37.99% of recounter identities were used in linguistics. In addition, in their research, Tang and John (1999) did not find any token of the recounter identity as the texts they examined were not research-based. Yet, in terms of the originator role (15.89%), this finding supports Sheldon's (14%).

Furthermore, we can see that many authorial references (72.38%) were used to express medium-risk identities, i.e., as architects and recounters and a sizeable proportion (25.25%) high-risk identities, i.e., as opinion holders and originators. To a large extent, this finding supports Işık-Taş (2018), who reported an average of 24.15% of first-person pronouns used to express high-risk functions in her native and internationally-oriented non-native subcorpora. In short, it seems that most authorial references are used to express sizable risk identities. This finding might suggest that RA publication in reputable international journals encourages professional authors to assume high-risk identities, such as presenting a knowledge claim or claiming responsibility and credit for the content.

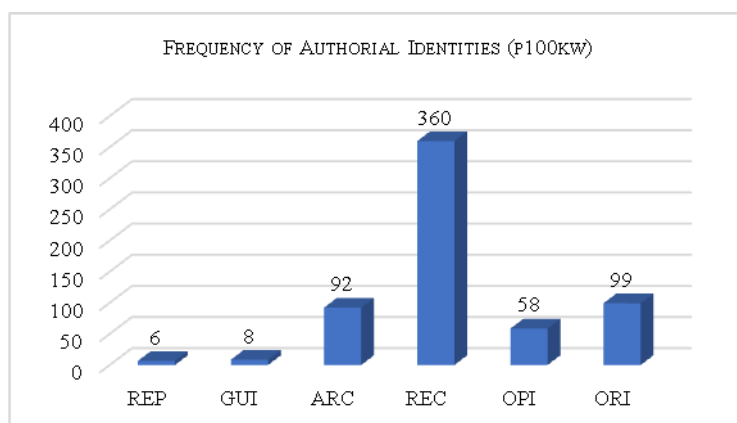


Figure 4 Overall Frequency of Authorial Identities in the Corpus (p100kw)

Figure 5 presents the use of first-person references by native and non-native professional authors to express authorial identities. As we can see, there seems to be no significant difference in the identities expressed by first-person references that both groups of authors used ( $\chi^2_{(5)} = 7.725$ ,  $p = 0.172$ ), with the author-as-recounter role being the most frequent identity expressed by first-person references in the two sub-corpora and the author-as-representative identity the least frequent. Again, this is in line with Işık-Taş (2018), who found no significant difference in the authorial identities expressed by first-person references in the native and non-native internationally-oriented Turkish subcorpora. This finding also suggests that native and non-native authors tend to assume considerable risk identities when publishing in internationally reputable journals.

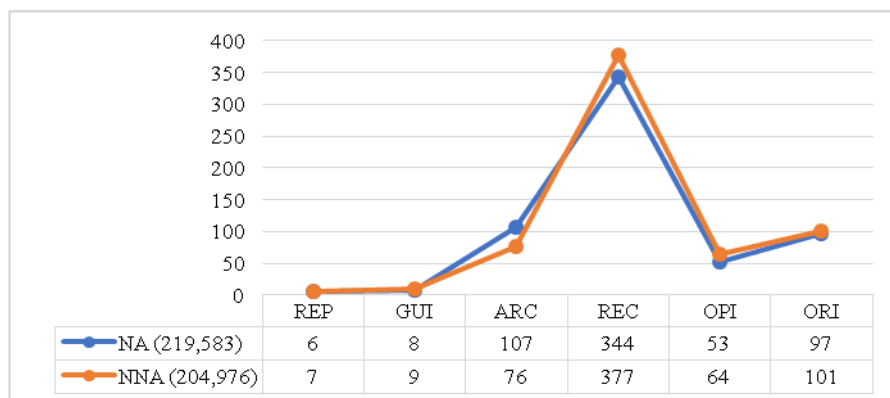


Figure 5 Frequency and Distribution of Authorial Identities by Author in the Corpus (p100kw)

Figure 6 shows a marked difference in the use of these authorial identities ( $\chi^2_{(10)} = 232.840$ ,  $p < 0.001$ ). As can be seen, most identities are dominantly used in biology, except for the guide and opinion holder identities. However, A chi-square test for these two identities does not show any significant difference ( $\chi^2_{(2)} = 2.88889$ ,  $p = 0.236$  for GUI and  $\chi^2_{(2)} = 2.35754$ ,  $p = 0.308$  for OPI). The same test for the other identities shows statistically significant difference ( $\chi^2_{(2)} =$

6.33333,  $p = 0.042$  for REP,  $\chi^2_{(2)} = 84.8110$ ,  $p < 0.001$  for ARC,  $\chi^2_{(2)} = 214.414$ ,  $p < 0.001$  for REC,  $\chi^2_{(2)} = 124.303$ ,  $p < 0.001$  for ORI). Another striking difference is that the presence of the authors in biology and medicine RAs far outnumbers that in linguistics RAs. Again, this seems to suggest that different epistemological beliefs and practices of disciplinary communities (Hyland, 2001, 2002a) play a crucial role in the extent to which professional authors “are willing to make a solid personal commitment to the most authorially powerful aspects of their texts, those which carried both the most risks and potentially gained them the most credit” (Hyland, 2002a). In what follows, each identity will be discussed in more detail.

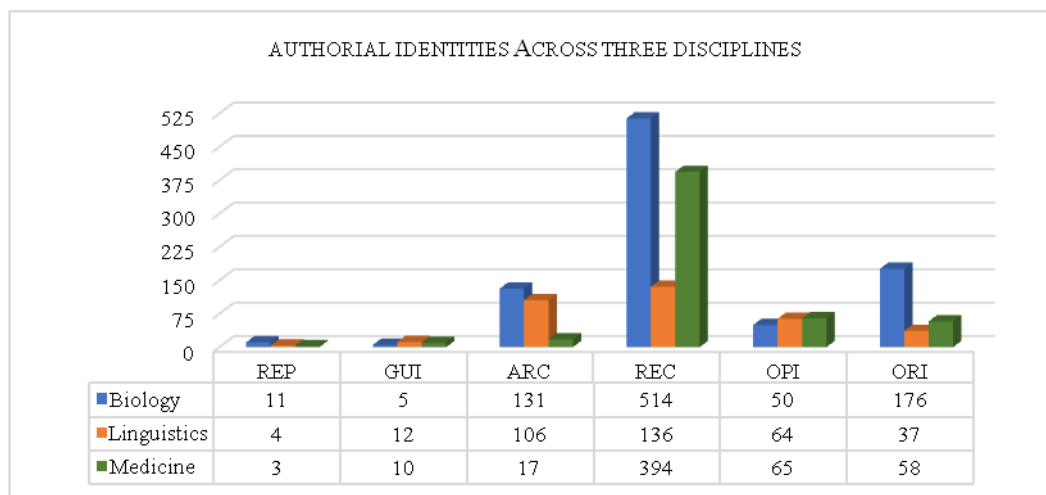


Figure 6 Frequency and Distribution of Authorial Identities by Discipline in the Corpus (p100kw)

#### (a). Authors as Representatives

The author-as-representative identity indicates a reference to people in general or discourse communities. It is the least authoritative power that authors show in their academic writing. Inclusive first-person references commonly express this identity. Some examples below illustrate how professional authors indicate their presence in their RAs as representatives of their discourse communities.

(6) Thus, it is challenging to identify the activities and properties of individual transposons. As a result, *we* only have a partial understanding of how transposons contribute to chromatin folding (...) (BIO01NA).

(7) This comparison of typical shell-noun uses in the N-be-that construction in popular and professional science articles contributes to *our* understanding of the preferred forms of evaluation in the two genres (LIN09NA).

In these examples, the first-person plural pronouns *we* and *our* are used to refer to the authors of the RAs and the readers, especially the research and disciplinary communities, i.e., biologists in the case of (6) and linguists in (7). While in (6), the RA authors indicate a research gap by pointing out the authors' as well as the readers' lack of understanding of the contribution of transposons to chromatin folding and their impact on gene regulation, in (7), the authors point to the contribution of their research to the authors' and the readers' understanding of the preferred forms of evaluation in the two genres. In short, professional authors sometimes employ *we* and *our* inclusively to show their identity as representatives, representing especially their communities of practice.

#### (b). Authors as Guides

The author-as-guide identity directly refers to the authors or readers of the RAs. It is commonly expressed using first-person pronouns. As the representative, this identity shows that authors do not wish to indicate a visible authorial presence in their RAs. However, this identity is reader-oriented. When assuming this identity, authors involve the readers in the research and draw their attention to it (Tang & John, 1999). Sometimes the inclusive first-person plural pronoun *we* is employed to indicate possibilities (8)-(9). Such pronouns in conditional constructions serve to draw the readers' attention and negotiate the authors' arguments as such constructions “bring the writer and reader closer to agreement” (Hyland, 2010, p. 173). The inclusive pronoun *we* is also sometimes used when authors present their interpretation and understanding of their data (10). Involvement of the readers in data interpretation may indicate a reader-writer mutual understanding. Thus such authorial presence can reflect positive politeness in academic writing (Harwood, 2005c; Myers, 1989). The authors view the readers as having the same expertise in data interpretation.

(8) If *we* assume that learners can recognize proper nouns and marginal words, *we* can add these figures to the coverage provided by word families at each level (...) (LIN13NA).

(9) Therefore, if *we* compare outcomes in the groups with blood pressure just above and just below the threshold, *we* can estimate the causal impact of the recommendations made after a screening diagnosis of hypertension (...) (MED19NNA).



- (10) Although the findings were based on self-reports, *we* can assume that the data are of high quality because the participants were highly motivated university graduates (...) (MED10NA)

(c). *Authors as Architects*

The author-as-architect identity indicates the use of first-person pronouns to organize and structure academic writing (Tang & John, 1999). Unlike the guide identity, this identity is realized by exclusive first-person references. These references can serve some functions, such as stating a purpose (11), limiting the research scope (12), and indicating what will be discussed next (13). The use of the first-person pronoun *we* to state the research purpose suggests a step in occupying the niche in Swales' CARS model (Swales, 1990). Overall, authors frequently use first-person pronouns to create coherent RAs by showing what they do to make their RAs easy to understand.

- (11) In this article, *we* provide the estimated numbers of new cancer cases and deaths in 2018 in the United States nationally and for each state (...) (MED11NA).  
 (12) In this paper, *we* focus closely on the writing development of a novice writer, Amal, and an experienced writer, Oscar (LIN04NA).  
 (13) However, the trend showed some dependence on molecular features such as mRNA expression levels, as *we* discuss below (BIO05NA).

(d). *Authors as Recounters*

Next, we turn to the author-as-recounters identity. RA authors very commonly use exclusive first-person pronouns to show themselves as recounters of the research process (Tang & John, 1999) or conductors of research (Sheldon, 2009). This role allows authors to explain their data collection and analysis procedures, as shown in examples (14)-(16) below.

- (14) To control for the quality of the probes and samples, *we* filtered out individuals with > 5% of probes associated with a detection P value > 10<sup>-3</sup> (...) (BIO09NNA).  
 (15) In this study, in order to ensure maximum credibility and validity of findings, *we* collected data through triangulation of instruments (LIN16NNA).  
 (16) *We* classified CVD cases using ICD-CM codes (international classification of diseases-clinical modification, 10th revision) (MED09NNA).

These examples show how authors collected and analyzed their data and how they presented themselves as agents responsible for any finding resulting from their procedural decisions. Such authorial presence indicates authors as competent members of their disciplinary communities. In short, authorial presence in data collection and analysis procedures contributes to the authors' positive self-representation.

(e). *Authors as Opinion Holders*

Professional authors frequently utilize the exclusive first-person pronoun *we* to express their attitudes, opinions, or views. This pronoun is commonly used together with epistemic verbs of cognition (Halliday & Matthiessen, 2004), such as *think*, *believe*, and *assume*, as illustrated in (17)-(19) below. When stating attitudes, opinions, or views, authors are committed to their statements and expose themselves to criticism and refutation. Therefore, this role can be viewed as highly face-threatening (cf. Işık-Taş, 2018). In other words, by explicitly showing their presence in their writing, authors indicate their responsibility for their personal opinions.

- (17) (...) *we thought* that the decrease in monoterpene biosynthesis could have caused an increase in sesquiterpene biosynthesis (BIO17NNA).  
 (18) *We believe* that the ACL (...) can serve as part of a lexical syllabus to raise learners' awareness of word co-occurrence and help them prioritize the learning of lexical items (LIN11NA).  
 (19) *We assumed* a common heterogeneity variance across all pairwise comparisons (...) (MED08NNA).

(f). *Authors as Originators*

The author-as-originator identity allows authors to claim authority and indicate ownership of the content of their writing. In Tang and John's (1999) model, this identity is the most powerful presence that authors can make in their writing and, therefore, the most face-threatening. This role allows authors to highlight their contribution to the field and promote their merits as researchers (Harwood, 2005b). In (20)-(22) below, the authors present their findings (20), knowledge claim (21), and contribution of their research to the field (21).

- (20) *We* found that both lincRNAs and mRNAs contain cis-regulatory potential in their promoters (...) (BIO03NA).  
 (21) *We* have claimed that (...) lexical diversity values will no longer be a function of text length (LIN17NNA).  
 (22) (...) *our* findings reinforce the existing evidence on the negative impact of ultra-processed foods on the overall incidence of chronic diseases (...) (MED10NNA).

## V. CONCLUSION

This study has explored how professional authors announce their presence in English RAs published in reputable international journals through first-person references. It has been shown that native and non-native authors frequently use authorial references in different disciplines. The results revealed that at the highest level of research publication, native and non-native authors show some notable similarity in the frequency of first-person references and their



identities, suggesting that socio-cultural factors do not seem to affect the employment of authorial references in reputable international publications. This seems to support the view that rhetorical choices are influenced by publication context (Işık-Taş, 2018), in which authors publishing in a national context tend to employ significantly fewer authorial references than those in an international context (see, e.g., Abdi & Farrokhi, 2015; Dontcheva-Navrátilová, 2013; Işık-Taş, 2018; Walková, 2018).

It has also been shown that authorial references are used differently in different disciplines in terms of frequency and the identities they express. They are prevalent in the hard sciences, but they are unexpectedly rare in the soft sciences like the humanities. This seems to suggest that different disciplines allow different degrees of authorial presence. Authors' rhetorical choices appear to be affected by "the social and epistemological practices of their disciplines and represent an important way of signaling membership and honoring what is accepted as professional engagement, appropriate intrusion, and persuasive conviction" (Hyland, 2001, p. 224).

It can be argued that authorial references may be an essential feature of written academic interaction at the highest level. They may play a crucial role in the promotion of a competent scholarly identity and the acceptance and ratification of knowledge claims. They allow authors to create identities serving as humble servants to seek cooperation and emphasize solidarity with their disciplinary communities and as knowledge originators to stress personal contribution to their respective disciplines.

This study has focused on authorial presence realized by first-person references in RAs published in reputable international journals by native and non-native authors in three disciplines, i.e., biology, linguistics, and medicine. Future research might address whether first-person references are employed differently in different RA sections to express authorial identities. Another interesting question is whether female and male authors from different cultural backgrounds and disciplines create different authorial identities when publishing in international and national journals. Answers to these questions and those related to the readers' processing of authorial presence will undoubtedly contribute to a better understanding of how scholars interact with one another through their scholarly writing and mainly how they create authorial identities in their text.

#### APPENDIX. LIST OF SELECTED JOURNALS

##### BIOLOGY

BIO01-10 Genome Biology (<https://genomebiology.biomedcentral.com/>)

BIO11-20 The Plant Cell (<https://academic.oup.com/plcell>)

##### LINGUISTICS

LIN01-10 English for Specific Purposes (<https://www.sciencedirect.com/journal/english-for-specific-purposes>)

LIN11-20 Journal of English for Academic Purposes (<https://www.sciencedirect.com/journal/journal-of-english-for-academic-purposes>)

##### MEDICINE

MED01-10 British Medical Journal (<https://www.bmj.com/>)

MED11-20 CA: A Cancer Journal for Clinicians (<https://acsjournals.onlinelibrary.wiley.com/journal/15424863>)

#### REFERENCES

- [1] Abdi, J., & Farrokhi, F. (2015). Investigating the projection of authorial identity through first-person pronouns in L1 and L2 English research articles. *International Journal of Language and Literature*, 3(1), 156–168. <https://doi.org/10.15640/ijll.v3n1a21>
- [2] Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. John Benjamins.
- [3] Arnaudet, M., & Barrett, M. (1984). *Approaches to academic reading and writing*. Prentice-Hall.
- [4] Basic, I., & Veselica-Majhut, S. (2016). Explicit Author Reference in Research Articles in Linguistics in English and Croatian. *Linguistics and Literature Studies*, 4(4), 233–242.
- [5] Becher, T. (1989). *Academic Tribes and Territories: Intellectual Enquiry and the Cultures of Disciplines*. The Society for Research into Higher Education and Open University Press.
- [6] Becher, T. (1994). The significance of disciplinary differences. *Studies in Higher Education*, 19(2), 151–161.
- [7] Chávez Muñoz, M. (2013). The "I" in interaction: Authorial presence in academic writing. *Revista de Lingüística y Lenguas Aplicadas*, 8, 49–58.
- [8] Conrad, S., & Biber, D. (2000). Adverbial marking of stance in speech and writing. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 57–73). Oxford University Press.
- [9] Dahl, T. (2009). Author identity in economics and linguistics abstracts. In E. Suomela-Salmi & F. Dervin (Eds.), *Cross-Linguistic and Cross-Cultural Perspectives on Academic Discourse* (pp. 123–134). John Benjamins.
- [10] Dontcheva-Navrátilová, O. (2013). Authorial presence in academic discourse: Functions of author-reference pronouns. *Linguistica Pragmática*, 3(1), 9–30.
- [11] Fløttum, K., Dahl, T., & Kinn, T. (2006). *Academic voices: Across languages and disciplines*. John Benjamin.
- [12] Flowerdew, J., & Wang, S. H. (2015). Identity in academic discourse. *Annual Review of Applied Linguistics*, 35, 81–99.
- [13] Gao, M., & Skolnick, J. (2010). Structural space of protein–protein interfaces is degenerate, close to complete, and highly connected. *Proceedings of the National Academy of Sciences*, 107(52), 22517–22522.
- [14] Geertz, C. (1983). *Local Knowledge: Further Essays in Interpretive Anthropology*. Basic Books.
- [15] Gomez, P. C. (2002). Do we need statistics when we have linguistics? *DELTA*, 18(2), 233–271.

- [16] Hacker, D. (2008). *Rules for Writers* (6th ed.). Bedford.
- [17] Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An Introduction to Functional Grammar* (3rd ed.). Hodder Arnold.
- [18] Hardjanto, T. D. (2016). Hedging through the use of modal auxiliaries in English academic discourse. *Humaniora*, 28(1), 37–50.
- [19] Hartwell, L. M., & Jacques, M. (2014). Authorial Presence in French and English: Pronoun + Verb. Patterns in Biology and Medicine Research Articles. *Discours*, 15, 3–21.
- [20] Harwood, N. (2005a). I hoped to counteract the memory problem, but I made no impact whatsoever”: Discussing methods in computing science using I. *English for Specific Purposes*, 24, 243–267.
- [21] Harwood, N. (2005b). ‘Nowhere has anyone attempted ... In this article I aim to do just that’ A corpus-based study of self-promotional I and we in academic writing across four disciplines. *Journal of Pragmatics*, 37, 1207–1231.
- [22] Harwood, N. (2005c). We do not seem to have a theory. The theory I present here attempts to fill this gap”: Inclusive and exclusive pronouns in academic writing. *Applied Linguistics*, 26(3), 343–375.
- [23] Hryniuk, K. (2018). Expressing authorial self in research articles written by Polish and English native-speaker writers: A corpus-based study. *Studies in Second Language Learning and Teaching*, 8(3), 621–642.
- [24] Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge University Press.
- [25] Hunston, S. (1993). Professional conflict—Disagreement in academic discourse. In M. Baker, G. Francis, & E. Tognini-Bonelli, *Text and Technology: In Honour of John Sinclair* (pp. 115–134). John Benjamins.
- [26] Hyland, K. (1994). Hedging in academic writing and EAP textbooks. *English for Specific Purposes*, 13(3), 239–256.
- [27] Hyland, K. (1996). Writing without conviction? Hedging in science research articles. *Applied Linguistics*, 17(4), 433–454.
- [28] Hyland, K. (1998). *Hedging in scientific research articles*. John Benjamin.
- [29] Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes*, 20, 207–226.
- [30] Hyland, K. (2002a). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, 34(8), 1091–1112.
- [31] Hyland, K. (2002b). Options of identity in academic writing. *ELT Journal*, 56(4), 351–358.
- [32] Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing*, 13, 133–151.
- [33] Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173–192.
- [34] Hyland, K. (2008). Disciplinary voices: Interactions in research writing. *English Text Construction*, 1(1), 5–22.
- [35] Hyland, K. (2010). Community and Individuality: Performing Identity in Applied Linguistics. *Written Communication*, 27(2), 159–188. <https://doi.org/10.1177/0741088309357846>
- [36] Hyland, K., & Jiang, F. (Kevin). (2018). “In this paper we suggest”: Changing patterns of disciplinary metadiscourse. *English for Specific Purposes*, 51, 18–30. <https://doi.org/10.1016/j.esp.2018.02.001>
- [37] Işık-Taş, E. E. (2018). Authorial identity in Turkish language and English language research articles in Sociology: The role of publication context in academic writers’ discourse choices. *English for Specific Purposes*, 49, 26–38. <https://doi.org/10.1016/j.esp.2017.10.003>
- [38] Isler, C. (2018). The use of first-person pronouns in master’s theses written in English by Turkish authors. *Journal of English Language and Education*, 4(1), 12–20.
- [39] Ivanič, R. (1994). I Is for interpersonal: Discoursal construction of writer identities and the teaching of writing. *Linguistics and Education*, 6(1), 3–15.
- [40] Ivanič, R. (1995). Writer Identity. *Prospect*, 10(1), 8–31.
- [41] Ivanič, R. (1998). *Writing and identity: The discoursal construction of identity in academic writing*. Benjamins.
- [42] Karoly, K. (2009). Author identity in English academic discourse: A comparison of expert and Hungarian EFL student writing. *Acta Linguistica Hungarica*, 56(1), 1–22.
- [43] Kim, E. (2015). Quantitative evidence on the uses of the first-person pronoun (I and We) in journal paper abstracts. *Journal of the Korean Society for Information Management*, 32(1), 227–243.
- [44] Knorr-Cetina, K. D. (1981). *The manufacture of knowledge*. Pergamon.
- [45] Krapivkina, O. A. (2015). Verbalization of the writer in academic prose. *Journal of Siberian Federal University. Humanities & Social Sciences*, 8, 1625–1634.
- [46] Kuo, C. (1999). The Use of Personal Pronouns: Role Relationships in Scientific Journal Articles. *English for Specific Purposes*, 18(2), 121–138.
- [47] Lachowicz, D. (1981). On the use of the passive voice for objectivity, author responsibility and hedging in EST. *Science of Science*, 2(2), 105–114.
- [48] Lee, D. Y. W. (2008). Corpora and discourse analysis. In V. K. Bhatia, J. Flowerdew, & R. H. Jones (Eds.), *Advances in Discourse Studies* (pp. 86–99). Routledge.
- [49] Lester, J. D. (1993). *Writing research papers* (7th ed.). Harper Collins.
- [50] Lipson, C. (2005). *How to Write a BA Thesis: A Practical Guide from Your First Ideas to Your Finished Paper*. University of Chicago Press.
- [51] Lorés-Sanz, R. (2011). The construction of the author’s voice in academic writing: The interplay of cultural and disciplinary factors. *Text & Talk*, 31(2), 173–193. <https://doi.org/10.1515/text.2011.008>
- [52] Manser, M. (2006). *The Facts on File: Guide to Good Writing*. Infobase Publishing.
- [53] Martí i Martí, P. (2003). Personal attribution in English and Spanish scientific texts. *Barcelona English Language and Literature Studies*, 12. Retrieved February 7, 2019 from <http://www.publicacions.ub.es/revistes/bells12/PDF/art09.pdf>.
- [54] Martí i Martí, I. A. (2005). Native and non-native writers’ use of first-person pronouns in the different sections of biology research articles in English. *Journal of Second Language Learning*, 15, 174–190.
- [55] Martínez, R. (2018). “Specially in the last years...”: Evidence of ELF and non-native English forms in international journals. *Journal of English for Academic Purposes*, 33, 40–52. <https://doi.org/10.1016/j.jeap.2018.01.007>

- [56] Matsuda, P. K. (2015). Identity in written discourse. *Annual Review of Applied Linguistics*, 35, 140–159.
- [57] 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, 235–249.
- [58] Millán, E. L. (2010). Extending this claim, we propose... The writer's presence in research articles from different disciplines. *Ibérica*, 20, 35–56.
- [59] Minitab. (2017). *Minitab 18 Statistical Software*. Minitab, Inc.
- [60] Mulkay, M. (1979). *Science and the Sociology of Knowledge*. George Allen & Unwin.
- [61] Mur Dueñas, P. (2007). 'I/we focus on...': A cross-cultural analysis of self-mentions in business management research articles. *Journal of English for Academic Purposes*, 6(2), 143–162. <https://doi.org/10.1016/j.jeap.2007.05.002>
- [62] Mur-Dueñas, P. (2010). Attitude markers in business management articles: A cross-cultural corpus-driven approach. *International Journal of Applied Linguistics*, 19(1), 50–72.
- [63] Mur-Dueñas, P. (2011). An intercultural analysis of metadiscourse features in research articles written in English and in Spanish. *Journal of Pragmatics*, 43(12), 3068–3079.
- [64] Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics*, 10, 1–35.
- [65] Nikolaev, V. V. (2010). Debt covenants and accounting conservatism. *Journal of Accounting Research*, 48(1), 51–89.
- [66] Pho, P. D. (2012). Authorial stance in research article abstracts and introductions from two disciplines. In N. Baumgarten, I. D. Bois, & J. House (Eds.), *Subjectivity in language and in discourse* (pp. 97–114). Emerald.
- [67] Poudat, C., & Loiseau, S. (2005). Authorial presence in academic genres. In E. Tognini-Bonelli & G. D. L. Camiciotti (Eds.), *Strategies in Academic Discourse* (pp. 51–68). John Benjamins.
- [68] Rentel, N. (2012). Subjectivity in academic discourse: A cross-linguistic comparison of the author's presence in French, Italian and German research articles in linguistics. In N. Baumgarten, I. D. Bois, & J. House (Eds.), *Subjectivity in language and in discourse* (pp. 339–354). Emerald.
- [69] Salager-Meyer, F. (2011). Scientific discourse and contrastive linguistics: Hedging. *European Science Editing*, 37(2), 35–37.
- [70] Sanderson, T. (2008). Interaction, identity and culture in academic writing: The case of German, British and American academics in the humanities. In A. Ädel & R. Reppen (Eds.), *Corpora and discourse: The challenges of different settings*. John Benjamin.
- [71] Sanjaya, I. N. S. (2013). *Hedging and Boosting in English and Indonesian Research Articles*. (Unpublished doctoral dissertation). Pennsylvania State University.
- [72] Scott, M. (2008). *WordSmith Tools Version 5*. Lexical Analysis Software.
- [73] Sheldon, E. (2009). From one I to another: Discursive construction of self-representation in English and Castilian Spanish research articles. *English for Specific Purposes*, 28(4), 251–265. <https://doi.org/10.1016/j.esp.2009.05.001>
- [74] Spencer, C., & Arbon, B. (1996). *Foundations of writing: Developing research and academic writing skills*. National Textbook Co.
- [75] Starfield, S., & Ravelli, L. J. (2006). "The writing of this thesis was a process that I could not explore with the positivistic detachment of the classical sociologist": Self and structure in New Humanities research theses. *Journal of English for Academic Purposes*, 5(3), 222–243. <https://doi.org/10.1016/j.jeap.2006.07.004>
- [76] Strunk Jr., W., & White, E. B. (2000). *The Elements of Style* (4th ed.). Pearson.
- [77] Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- [78] Swales, J. (2004). *Research Genres*. Cambridge University Press. <https://doi.org/10.1017/cbo9781139524827>
- [79] Swales, J., & Burke, A. (2003). It's really fascinating work": Differences in evaluative adjectives across academic registers. In P. Leistyna & C. F. Meyer (Eds.), *Corpus analysis: Language structure and language use* (pp. 1–18). Rodopi.
- [80] Tang, R., & John, S. (1999). The "I" in identity: Exploring writer identity in student academic writing through first-person pronoun. *English for Specific Purposes*, 18(Suppl.), 23–39.
- [81] Tardy, C. M. (2012). Current conceptions of voice. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres*. Palgrave.
- [82] Tardy, C. M., & Matsuda, P. K. (2009). The construction of author voice by editorial board members. *Written Communication*, 26(1), 32–52.
- [83] Taylor, R. B. (2005). *The Clinician's Guide to Medical Writing*. Springer.
- [84] Tayyebi, M. (2012). Personal pronouns in English and Persian medical research articles. *English for Specific Purposes World*, 36(12), 1–12.
- [85] Thompson, G., & Ye, Y. (1991). Evaluation of the reporting verbs used in academic papers. *Applied Linguistics*, 12(4), 365–382.
- [86] Vassileva, I. (1998). Who am I/who are we in academic writing? A contrastive analysis of authorial presence in English, German, French, Russian and Bulgarian. *International Journal of Applied Linguistics*, 8(2), 163–190.
- [87] Vladimirou, D. (2007). I suggest that we need more research: Personal reference in linguistics journal articles. *Lancaster University Postgraduate Conference in Linguistics and Language Teaching*, 1, 139–157.
- [88] Walková, M. (2018). Author's self-representation in research articles by Anglophone and Slovak linguists. *Discourse and Interaction*, 11(1), 86–105. <https://doi.org/10.5817/DI2018-1-86>
- [89] Xia, G. (2018). A cross-disciplinary corpus-based study on English and Chinese native speakers' use of first-person pronouns in academic English writing. *Text & Talk*, 38(1), 93–113. <https://doi.org/10.1515/text-2017-0032>
- [90] Yang, Y. (2013). Exploring linguistic and cultural variations in the use of hedges in English and Chinese scientific discourse. *Journal of Pragmatics*, 50(1), 23–36.
- [91] Zhang, Y. A. (2012). Generic analysis on 1st personal pronouns in the international conference presentation. *International Conference on Education Technology and Management Engineering, Lecture Notes in Information Technology*, 16–17, 345–349.

**Tofan Dwi Hardjanto** is a senior lecturer in linguistics in the English Department, Faculty of Cultural Sciences, Universitas Gadjah Mada, Yogyakarta, Indonesia. He is interested in a broad area of research, including academic writing, discourse analysis, and pragmatics.