Making it personal: Examining gendered personalization patterns of single-authored papers

Gita Ghiasi[1], Maxime Sainte-Marie[2], and Vincent Larivière[3]

[1] gita.ghiasi.hafezi@umontreal.ca
[2] msaintemarie@gmail.com
[3] vincent.lariviere@umontreal.ca
École de Bibliothéconomie et des Sciences de l'Information, Université de Montréal, CP6128, Station Centre-Ville, Montreal, Quebec,H3C3J7 (Canada).

Abstract

Authorial self-mentions as one of the primary forms of scientific communication in sole authorship could convey not only the salient characteristics of disciplinary epistemologies but a representation of the author. The latter is often reflected through the use of first-person pronouns. The objective of this study is to assess disciplinary, national, and gender differences in the use of first-person subject pronouns in sole-authored scientific articles. Following the analysis of all sole-authored articles published between 2008 and 2017 and indexed in the Web of Science, this study reveals that women refer to themselves in the singular form more often than men, which could be indicative of cultural associations of masculinity and authority in the system of science. The results of this study are of great importance to academic evaluative processes where sole-authorship is rewarded as an independent work and author's credibility.

Introduction

Despite the constant decline in numbers (Abt 2007; Barlow et al. 2018; Kuld and O'Hagan 2018), the reward system of science has traditionally considered sole-authorship as a measure of a researcher's ability to work independently (Gasparyan et al. 2013; Moore and Griffin 2006). Due to the dubious tendency of conferring greater credit to individual contributions, sole-authoring has been given special importance in several research review and evaluation systems, impacting promotion, tenure, and funding allocation management (Moore and Griffin 2006; Vafeas 2010). Moreover, it has also been argued that sole-authorship could offer a distinct level of recognition for the author than co-authorship, which could serve a researcher beyond the tenure and promotion (Moore and Griffin 2006) in the scientific system.

As regards to gender, sole-authorship could play a vital role in women's promotion and tenure reception: women are not only less likely to receive tenure when they co-author, but they also receive less credit for their contributions in co-authored papers (Sarsons 2017). Despite this, women are highly underrepresented as sole-authors (West et al. 2013), and female sole-authored papers are subject to lower citation rates than male-authored ones (Bendels et al. 2018; Larivière et al. 2013).

Along these lines, linguistic choices on authorial self-mention could also impose a considerable impact on scholarly communication, because it sheds light on the author's epistemological and social self-reflection by drawing on assumptions authors hold about their role in the research process (Hyland 2003). Gender differences in language use are of special interest here, as they can provide valuable insight into the social aspects of communication (Newman et al. 2008). Particularly, differences in authorial self-mention provide a rhetorical strategy to mirror an author's self-reflection of his/her contribution to a piece of research, presentation, and promotion of his/her knowledge claims, and research credibility (Hyland 2003).

In this regard, the main objective of this paper is to assess gender differences in the use of pronouns in single-authored articles. More specifically this study analyzes the referential meaning and pragmatic function of both singular ('I') and plural ('We') forms of first-person subject pronouns (termed as 'S' and 'P' hereafter, respectively) that researchers of each gender use to promote their scientific claims and themselves. For this purpose, a cross-country and

cross-disciplinary analysis of the use of first-person subject pronouns in abstracts and acknowledgments of single-authored articles is first provided, and gender is further factored into these differences.

Methods

Abstracts and acknowledgments of all 1,184,186 single-authored scientific articles (hereafter simply referred to as 'articles') published between 2008 and 2017 inclusively are extracted from the Web of Science (WoS). For each relevant article entry, the following attributes are extracted: article ID, author full given name, acknowledgment, abstract, publication year, and journal name. Discipline assignation is based on the National Science Foundation (NSF) field classification of journals used in the Science and Engineering Indicators (SEI) reports. Contrary to the WoS disciplinary classification, the NSF classification scheme assigns only one discipline to each journal, which prevents multiple counts of articles published in multidisciplinary journals. The relative citation of a paper is measured as the average yearly number of citations received by a paper divided by the average yearly number of citations to all the papers from the same year, in the same discipline and of the same document type. The normalized journal impact factor is defined similarly, considering the IF of the journal in which a paper is published. In this study, top-cited papers and top-impact journals refer to the top 5% cited papers and the top 5% high impact journals.

Following the procedure described in Larivière et al. (2013), gender is then assigned to article authors using universal and country-specific name lists in sequence. In the first step, gender is attributed to all authors based on 1990 US census data, which provides lists associating each given name to the percentage and gender of the population bearing that name. In cases where a name is used for both genders, a specific gender is assigned only if the corresponding gender is assigned ten times more frequently than the other. In a second step, all unassigned author names are then matched with corresponding entries in country-specific name lists, based on the geographical information given by the institutional affiliations of authors. These assignment procedures resulted in the author gender identification of more than 985,721 articles, which corresponds to 96% of total articles in which the full first name of the author is given.

Finally, in order to extract all first-person personal pronoun information contained in both abstracts and acknowledgments, the textual content of both attributes is first grammatically disambiguated using the TreeTagger part-of-speech tagger (Schmid 1999, 2013), as trained on the British National Corpus tagset (Leech et al. 1994). Following this, words segmented into either 'I' or 'We' and tagged as 'PNP' (corresponding to the 'personal pronoun' part-of-speech tag) by the algorithm are counted for each of the two article attributes. Articles are further categorized under 'I' and 'We', when one attribute equals to zero, and the other is equal or more than 1.

Findings

The results show that authors tend to rely more on S in their acknowledgments and on P while characterizing and summarizing their research. This sheds light on the differences in scientific communication when authors describe their research and express a personal statement, in the sense that P is more popular in the scientific and more formal context. However, the use of S increases in highly cited papers, while P is utilized at a higher rate in papers published in high impact journals. This trend highlights differences in rhetorical stance between two different recognition criteria: publishing in high impact journals and attracting citations. S is mostly used in social sciences, humanities, and psychology while P is most popular in mathematics, physics, and engineering (Fig. 2: Left). This corresponds to the findings of (Hyland 2003), who associated these opposite trends in 'hard' and 'soft' sciences to cross-disciplinary differences in the ways that research is conducted and accepted by the scientific community. Among highly

prolific countries, S is highly employed among North American countries (Canada and the US), while P is more used in East Asian countries (China and Japan) (Fig. 2: Right). These differences are further scrutinized using the Hofstede's individualism dimension of national culture (Hofstede 1984) (Fig. 3), which reveals a strong correlation between the two.

Share of papers using first-person subject pronouns
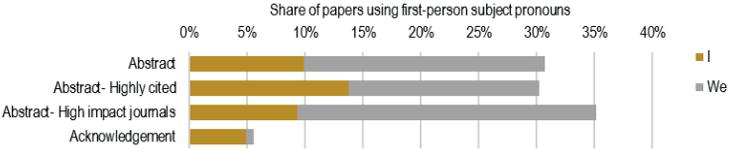
Figure 1- share of sole-authored papers using first-person subject pronouns in their abstracts and acknowledgments
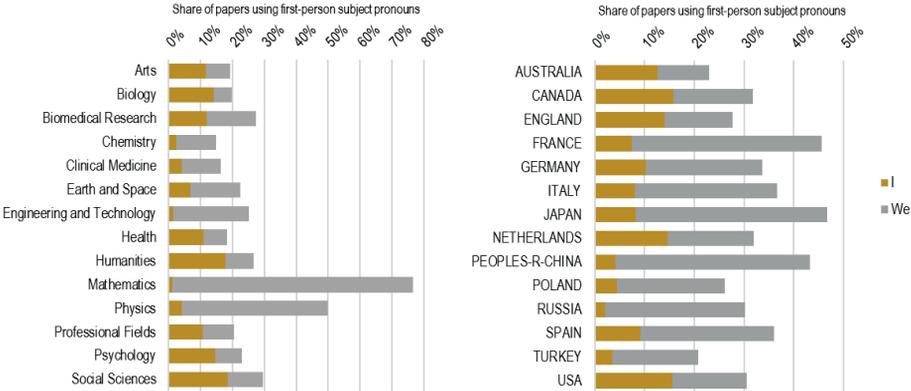
Figure 2- share of sole-authored papers using first-person subject pronouns (left) by filed and (right) by county

$y = -0.0035x + 0.3377$
$R^2 = 0.5332$
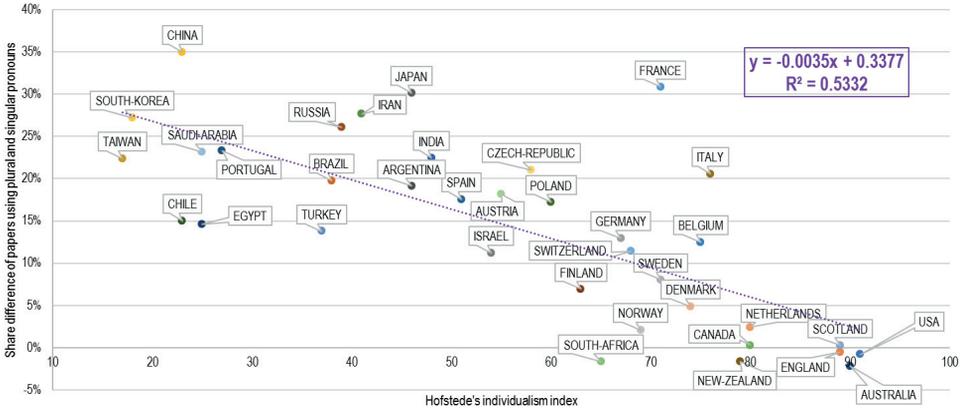
Figure 3- differences in shares of papers using singular and plural pronouns by Hofstede cultural individualism index

Gender Analysis

Considering only articles using S and P, results show that women use S more often than men in their papers and even more so in highly cited papers (Fig. 4). The largest gender differences

in the use of pronouns are in the field of social sciences, and professional fields, where women use S more often (Fig. 5), whereas in biomedical research women use P at a higher proportion than men than men.

While considering the national tendencies, gender differences in the use of pronouns in the scientific writing are more pronounced in countries with low power distance and are negligible in societies with high power distance such as China and Russia (Fig. 6 and 7). This could relate to the centralization of authority (Hofstede 1984) and its strong association with social uniformity.
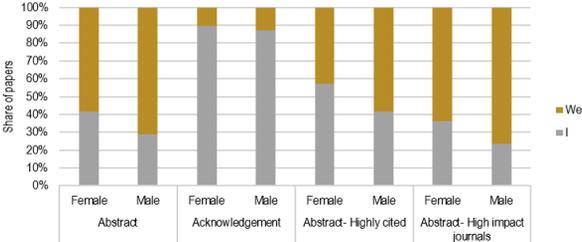

Figure 4- Share of sole-authored papers using singular and plural pronouns by gender
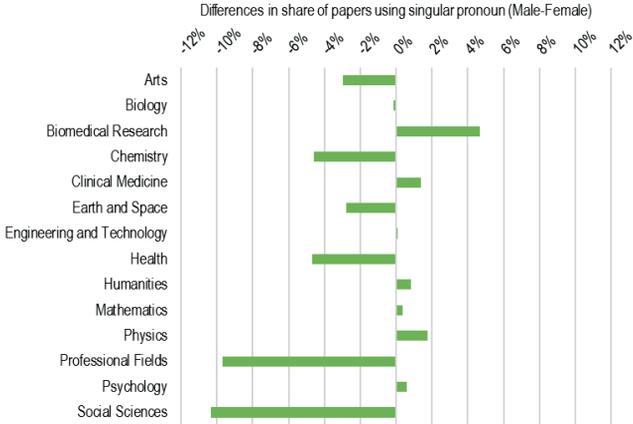

Figure 5- Gender differences in share of papers using singular pronoun by field (Male - Female)
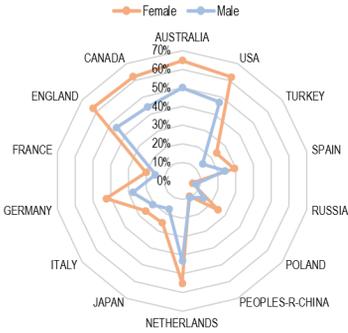

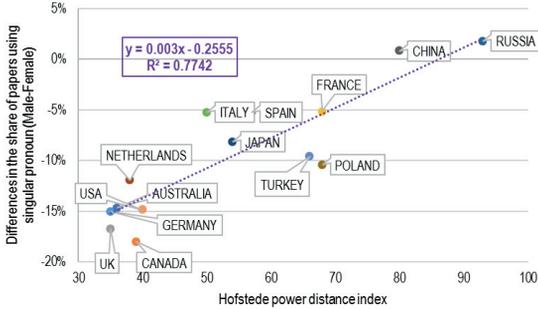**Figure 6- Share of papers using singular pronoun by gender and by country**

Figure 7- gender differences in the use of singular pronoun by Hofstede individualism index (Male - Female)

Discussion and Conclusion

Results of this study reveal that the use of 'We' generally prevails in the single-authoring context. While 'I' is practiced more often among top-cited papers, articles in top impact journals use 'We' more often. This shows that the use of the plural form might present stronger support for authors to present their research claims for the review process, while the singular person might present a stronger research creditability for a larger audience and the scientific community. Moreover, this study shows that 'I' is used more often in social sciences, humanities, and psychology, while 'We' is most used in mathematics, physics, and engineering disciplines. This is in line with the findings of Hyland (2003), who argued that since hard sciences are experimental, and results are replicable, therefore authors stay objective to their findings and use less intrusive and personal writing style. On the other hand, due to the use of interpretive approach, the level of personal engagement with the readers is important in soft sciences, in the sense that authors are required to present themselves as an informed researcher with a particular point of view to receive credibility for their perspective and research claims.

With regard to gender differences, this study shows that women refer to themselves as "I" more often than men and men use 'We' more often than women and that these differences are highly conspicuous in social sciences and professional fields. This could be indicative of cultural associations of masculinity and authority in the system of science. Because the use of plural pronouns may distance the author from the text, yet exhibit a temporary dominance by conferring the right to speak with authority on authors (Hyland 2001). The indication of less personal intrusion could also refer to the author and his/her possible colleagues and help acceptance of research arguments, based on which the reader assumes that scientific claims of the paper are supported by a research group or community (Zhou 2017). Therefore, one of the possible explanations for larger gender differences observed in social sciences is that women's contribution to papers is the result of an individual endeavor while for men is more of the team efforts.

Finally, this study confirms that the authors' use of singular and plural pronouns in the scientific papers are associated with how a national culture defines its self-image as "I" or "We" and that gender differences in communication style is less evident in countries with power distance culture, as plural pronouns are uniformly practiced to reflect authority.

The results of this study are of great importance to academic evaluative processes where sole-authorship is rewarded as an independent work and author's credibility. Authorial self-mentions as one of the primary forms of scientific communication in sole authorship could convey not only the salient characteristics of disciplinary epistemologies but a representation of the author, upon which research credibility and recognition within the scientific community could be attained. Therefore, gender differences in authorial self-mentions could also address under-recognition of the contribution of women to science.

References

Abt, H. A. (2007). The future of single-authored papers. *Scientometrics*, *73*(3), 353–358. doi:10.1007/s11192-007-1822-9

Barlow, J., Stephens, P. A., Bode, M., Cadotte, M. W., Lucas, K., Newton, E., et al. (2018). On the extinction of the single-authored paper: The causes and consequences of increasingly collaborative applied ecological research. *Journal of Applied Ecology*, *55*(1), 1–4. doi:10.1111/1365-2664.13040

Bendels, M. H. K., Müller, R., Brueggmann, D., & Groneberg, D. A. (2018). Gender disparities in high-quality research revealed by Nature Index journals. *PLOS ONE*, *13*(1), e0189136. doi:10.1371/journal.pone.0189136

Gasparyan, A. Y., Ayvazyan, L., & Kitas, G. D. (2013). Authorship problems in scholarly journals: considerations for authors, peer reviewers and editors. *Rheumatology international*, *33*(2), 277–284.

Hofstede, G. (1984). *Culture's Consequences: International Differences in Work-Related Values*. SAGE.

Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for specific purposes*, *20*(3), 207–226.

Hyland, K. (2003). Self-citation and self-reference: Credibility and promotion in academic publication. *Journal of the American Society for Information Science and technology*, *54*(3), 251–259.

Kuld, L., & O'Hagan, J. (2018). Rise of multi-authored papers in economics: Demise of the 'lone star' and why? *Scientometrics*, *114*(3), 1207–1225. doi:10.1007/s11192-017-2588-3

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature News*, *504*(7479), 211. doi:10.1038/504211a

Leech, G., Garside, R., & Bryant, M. (1994). CLAWS4: the tagging of the British National Corpus. In *Proceedings of the 15th conference on Computational linguistics-Volume 1* (pp. 622–628). Association for Computational Linguistics.

Moore, M. T., & Griffin, B. W. (2006). Identification of factors that influence authorship name placement and decisions to collaborate in peer-reviewed, education-related publications. *Studies in Educational Evaluation*, *32*(2), 125–135.

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes*, *45*(3), 211–236.

Sarsons, H. (2017). Recognition for Group Work: Gender Differences in Academia. *American Economic Review*, *107*(5), 141–145. doi:10.1257/aer.p20171126

Schmid, H. (1999). Improvements in part-of-speech tagging with an application to German. In *Natural language processing using very large corpora* (pp. 13–25). Springer.

Schmid, H. (2013). Probabilistic part-ofispeech tagging using decision trees. In *New methods in language processing* (p. 154).

Vafeas, N. (2010). Determinants of single authorship. *EuroMed Journal of Business*, *5*(3), 332–344. doi:10.1108/14502191011080845

West, J. D., Jacquet, J., King, M. M., Correll, S. J., & Bergstrom, C. T. (2013). The Role of Gender in Scholarly Authorship. *PLOS ONE*, *8*(7), e66212. doi:10.1371/journal.pone.0066212

Zhou, L. (2017). The Pragmatic Effect of the First Person Plural Pronouns in Single-authored Scientific Articles. In *3rd International Conference on Arts, Design and Contemporary Education (ICADCE 2017)*. Atlantis Press.