

The Naturalized Men and Over-represented Women: A Collocational Analysis of Gender Markers in Filipino Academic Discourse

Kamille D. Legaspi-Torres¹ 

Polytechnic University of the Philippines, Philippines

De La Salle University-Manila, Philippines

ARTICLE HISTORY

Received: January 15, 2026

Accepted: February 14, 2026

Published: February 28, 2026

ABSTRACT

Despite the alarming findings on gender inequality and the extensive literature on corpus analysis, the journal articles written by Filipino scholars remain unexamined. This gap is critical because journal articles serve as a foundation of knowledge from which research, classroom discussions, and other pedagogical practices are constructed. To address this gap, this research examined how gender is portrayed in a collection of published journal articles called CRAFT. Using Social Action Theory along with the collocational analysis of select feminine and masculine markers, the results show that gender equality is still a complex discussion in the country. While statistics seem to demonstrate women's visibility, this does not guarantee empowerment. Further analysis seems to suggest that the high frequency of feminine markers shows the constant need for gender identification among women. This interpretation is supported by the collocations of various feminine markers with indicators of classification, relation, and physical identification, along with negative appraisal. Conversely, the limited occurrence of masculine markers does not necessarily suggest new gender trends; rather, a potential perpetuation of stereotypical beliefs where men are considered the unmarked gender who are primarily functionalized by economic and professional agency. These findings emphasize the need for the continuous strengthening of gender sensitivity programs in the country and a systemic review of the academic editorial policies and practices to model and encourage gender equality. While the Filipinos have already achieved monumental success in bridging the gender gap, certain areas still need more attention.

Keywords: Gender parity, Corpus linguistics, Collocational analysis, Representations of Social Actor Theory

I. INTRODUCTION

Gender inequality continues to be a systemic concern experienced by the international community. At present, women still endure various gender-based constraints even though worldwide efforts are already in place. Studies show that they are still burdened by household labors that are uncompensated and unfairly distributed, restricted from accessing significant assets like land and technology, and victimized by

doi: 10.5281/zenodo.18754359

¹Corresponding Author: kdlegaspi@pup.edu.ph

forced marriages and genital mutilation (UN Women & United Nations Department of Economic and Social Affairs, 2025). These data emphasize that gender inequality is not just a theoretical matter but a serious universal problem.

Previous studies that investigated the intersection of gender and language highlight that the latter can be used to propagate inequality by encoding gender stereotypes and discrimination (Formanowicz & Hansen, 2021). Various domains illustrate these asymmetries in representation. For example, the stereotypes that describe women in leadership as primarily communal relegate their agentive capacity, resulting in underrepresentation (Tremmel & Wahl, 2023). Moreover, discussions on sexual issues (Dianitasari, 2021) and broader concepts of gender inequity, exclusion, and victimization (Amoniyani & Obasi, 2025) are mainly linked to the term woman. Likewise, diminutives that emphasize women's relational roles help normalize domestic violence (Angouri & Baxter, 2021).

Interestingly, despite numerous gender sensitivity initiatives, academia also shows indicators of disparities (Llorens et al., 2021). Empirical evidence reveals that certain schoolbooks of secondary students contain gender-based exclusions and portrayals (Saher et al., 2023). Additionally, academic outputs, like undergraduate essays, reveal a preference for masculine pronouns (Raza et al., 2025). Lastly, the research productions of doctoral students show that authors' age, particularly among men, significantly correlates with an increased susceptibility to using gender biased terminologies (Orgeira-Crespo et al., 2021). These findings confirm how language, even in a seemingly objective domain, can be an instrument in reflecting and reproducing social inequalities.

Corpus-based investigations have been particularly useful in investigating embedded biases in large collections of languages. There is an extensive body of international research that attempted to use this method to reveal gendered portrayals both in the academic and public settings (Biber & Burges, 2000; Brun-Mercer, 2021; Jevrić, 2017; Mosqueda & Sánchez, 2022; Pearce, 2008; Posch, 2023). Unfortunately, despite the global findings on gender inequality and the extensive literature on corpus analysis, the journal articles written by Filipino scholars remain underexplored. This gap is critical because journal articles serve as the foundation of knowledge from which research, classroom discussions, and other pedagogical practices are constructed. As an authority in academia, its potential to carry gender bias risks, leading its readers to unconsciously adopt such beliefs. Ultimately, if they remain unchecked, they have the potential to become tools for naturalizing harmful gender ideologies.

To address this gap and to ensure that the academic community and other research-based institutions will not inadvertently reinforce the very norms it hopes to subvert, this study advances the field in three significant ways. Primarily, it introduces a pioneering work on corpus linguistics called Corpus of Research Articles by Filipino Thinkers (CRAFT). This two-million-word corpus, built from published articles written exclusively by Filipino experts, offers itself as a productive site of investigation. I specifically used collocational analysis as a tool to scrutinize the ideologies hidden in the linguistic patterns of scholars from the Philippines. By positioning collocations as

meaningful representations of the mental shortcuts or schemas that a culture uses to categorize people (Stubbs, 2001), it acts as a window into the collective subconscious of a society. This makes its analysis in large-scale corpora capable of revealing lexical forms that potentially carry gender bias. Previous studies have been conducted to demonstrate its validity (Alkhamash & Al-Nofaie, 2020; Callaway, 2022; Mosqueda & Sánchez, 2022; Posch, 2023; Raphael, 2023; Stroble & Lebani, 2022; Zasina, 2019). Moreover, this study expands the previous investigations that focused solely on the markers –men and women (Biber & Burgess, 2000; Brun- Mercer, 2021; Jevrić, 2017; Mosqueda & Sánchez, 2022; Pearce, 2008; Posch, 2023). The strategic inclusion of other gender pairs, such as female and male, girl and boy, she and he, and her and him, helped to enhance its scope to capture more nuanced findings. Finally, this study went beyond relying on frequencies by anchoring the collocational analysis into Van Leeuwen’s (2003) Representation of Social Actor (RSA). Past investigations demonstrate the robust application of this theory in analyzing social actors, signifying its strong fit for this analysis (Almaghlouth, 2022; Chaemsaitong, 2021; Mustofa et al., 2025; Rashidi & Ghaedsharafi, 2015; Utama et al., 2020).

Essentially, the utilization of corpus linguistics, specifically collocational analysis guided by RSA, allowed this research to advance the study of gendered academic discourse in the Philippine context by contributing theoretically and empirically. Consequently, I attempted to answer the questions listed below:

1. How do the relative frequencies of the select masculine and feminine markers compare across the fields of specialization within CRAFT?
2. What are the significant collocates of gender markers within CRAFT?
3. Based on the collocates of the gender markers, how are women and men portrayed in the academic text according to the categorization of Van Leeuwen’s RSA Theory?

II. METHODS

While studies involving corpora of instructional materials (Rashidi & Ghaedsharafi, 2015; Briones, 2019; Zhao, 2022), literature, and learners’ outputs (Horton, 2008) have been investigated to assess their potential in encoding gender disparity, a distinct scarcity of the same investigation is found in journal articles written by Filipino scholars. Hence, a two-million-word (2,000,000) corpus called CRAFT was developed. As a body of work that contains highly selected scholarly works, it is positioned to be an effective reflection of gender parity in research.

To ensure the legitimacy of the investigation that involved CRAFT, three important considerations were carefully followed: academic rigor, temporal currency, and balance of representativeness. The academic rigor and temporal currency of CRAFT were ensured by sourcing only the articles published between 2015 and 2025 that satisfy the evaluation criteria of Web of Science (WOS) and Scopus indexing. Moreover, balance representativeness was secured by including ten (10) distinct fields of specialization which were coded using W1 through W10: Business, Tourism, and Hospitality (W1), Economics (W2), Agriculture (W3), Nursing and Allied Health Science (W4), Engineering

(W5), Psychology and Philosophy (W6), Political Science (W7), Applied Linguistics (W8), Education (W9) and Anthropology (W10).

The concordance feature of Sketch Engine was used to identify the collocations that met the set logDice threshold. This helped in ensuring that the retrieved collocations were not due to chance. I patterned this study after Frankenberg-Garcia et al. (2018), who set the logDice threshold to 5 or higher, with a minimum co-occurrence frequency of 5. Furthermore, I set the extraction parameters as follows: (a) a minimum corpus frequency of 5; (b) a minimum frequency in the given range of 3; and (c) a collocational span of -4 to +4. The resulting collocations were categorized according to Van Leeuwen's framework and verified by two inter-coders.

Although RSA provides a comprehensive framework for analyzing social representation, its category, Categorization, was strategically isolated because of the methodological alignments of its subtypes with collocational analysis. This strategic choice allowed me to extend the analysis from counting the frequency of occurrence to interpreting the collocational patterns as indicators of meaningful sociological insights.

Categorization has three main subtypes: functionalization, identification, and appraisal. According to Van Leeuwen (2003), writers are left with a choice to represent their social actors either through performance (Functionalization) or essence (Identification). When writers portray them through their roles and occupations, they are more functionalized. This functionalization eventually results in a dominance of markers that depict their agency and legitimacy. On the contrary, when social actors are characterized through their inherent attributes, the focus shifts from "what they do" to "who they are". This harmful shift downplays their agency and emphasizes the traits that can potentially normalize social inequality.

Identification categorizes social actors using three distinct features: classification, relational identification, and physical identification. First, when writers use Classification, they tend to use markers of gender, civil status, and ethnicity to characterize social actors. Meanwhile, terms related to personal, kinship, and work relations are often used to mark their relational identification. Lastly, writers may focus on their physical appearance to mark their physical identification. Van Leeuwen (2003) argues that these descriptors are rarely neutral. They are instead labels that may carry heavy cultural connotations and stereotypes, allowing evaluative bias.

The last part of the whole Categorization is Appraisal. Writers can use evaluative nouns or adjectives to portray their actors, revealing their evaluative stance that often signals power imbalances.

To ensure analytical rigor and minimize interpretive ambiguity, the following coding criteria were established for the categorization of significant collocates.

Table 1
Coding Criteria for Categorizing Collocates

Category	Coding Criteria	Illustrative Examples
Functionalization	verbs that transformed into nouns through suffixation (such as <i>-er, -ee, -ian</i>)	teacher, interviewee, technician
	nouns that transformed into a new noun through suffixation (such as <i>-ist, -eer</i>)	pianist, mountaineer
	nouns that transformed into new nouns through compounding	cameraman, chairperson
Identification	Classification	nouns or adjectives that depict age, gender, class, wealth, race, ethnicity, religion, and sexual orientation
	Relational Identification	nouns that depict familial or work relations
	Physical Identification	nouns, adjectives, or prepositions that denote physical characteristics
Appraisal	words and expressions that depict personal judgement	young, old, married daughter, parent tall, goth loved, pitied, survivors

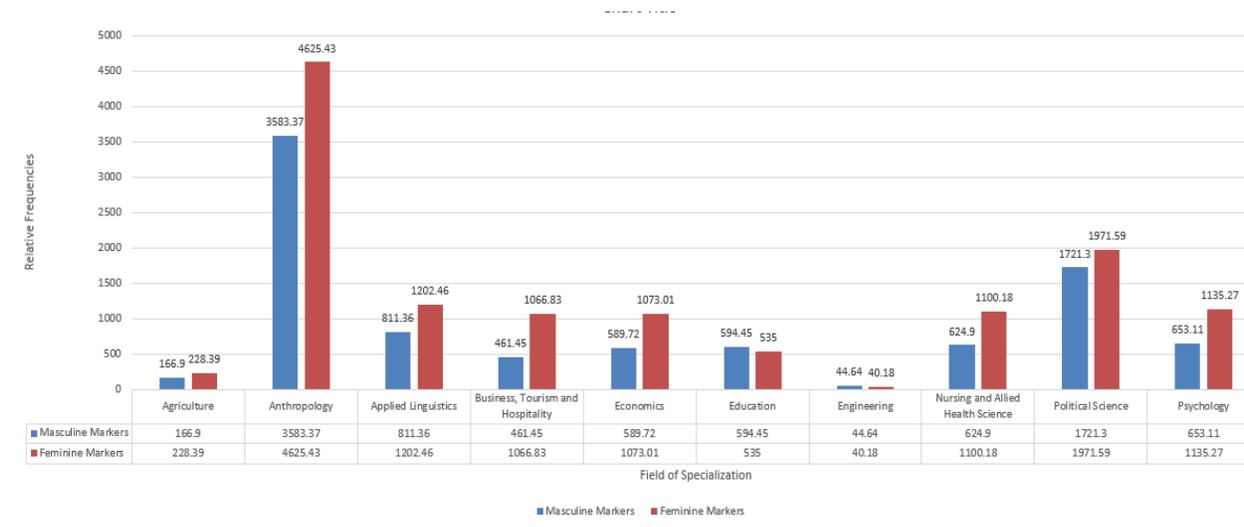
In instances where a collocate possessed polysemous meanings, I examined the specific concordance lines to verify its contextual function before final assignment. Since this research used published journal articles as a dataset, it does not require formal ethical clearance. However, to demonstrate the adherence of the corpus to the principle of Fair Use for transformative research, only two thousand (2,000) words were extracted from each article. This sampling method ensured that the articles were only taken for research purposes without any attempt to disrespect the original publishers and authors through unauthorized distribution. The identity of the Filipino researchers was also protected by using unique alphanumeric codes. This helped maintain their anonymity and shifted the focus of the analysis from individual authors to collective discursive patterns.

III. RESULTS

This part of the study presents the results of the investigation conducted using CRAFT. The frequency distribution and collocational patterns interpreted through the Categorization Framework of Van Leeuwen's RSA Theory demonstrate how Filipino scholars portray social actors.

Figure 1

Summation of the Relative Frequencies of the Gendered Markers across Fields of Specializations



On the surface level, Figure 1 shows that the frequency counts of the gendered markers are skewed toward the feminine markers. They outnumbered the masculine markers in eight out of the ten disciplines found in CRAFT. This initial data seems to suggest a high degree of female visibility, specifically within the social science domains. Interestingly, even previously male-dominated areas like Agriculture, Business, Tourism and Hospitality, Economics, and Political Science show a skew towards feminine markers. This pattern appears to support the globally acknowledged gender parity statistics of the Philippines and may corroborate research data that claims that the higher level of representation is a potential product of participation (Jaramillo et al., 2025). It is unexpected to see how this pattern exists in certain domains, although there are claims of underfunding in Economics, Culture, and practice-related research proposals where women are well represented (Larregue & Nielsen, 2023).

Conversely, technical fields like Engineering tend to yield low occurrences of gender markers, potentially due to their focus on iterative design rather than human interactions (Aydın-Günbatır and Röhrig, 2023). Finally, although both markers appeared less frequently in Education compared to other fields, it is observable that masculine markers outnumbered the feminine ones. This subtle masculine skew in a female-dominated discipline is reinforced by various investigations that utilized corpus analyses of academic-related corpora. These studies revealed specific ways the academe subtly demonstrates gender inequality, such as underrepresentation of women (Aguilar-Delavin, 2022; Çağlak, 2025), exclusion (Saher et al., 2023), and a preference for pronouns that depict masculinity over femininity (Raza et al., 2025).

Applying RSA for a more nuanced analysis, the study found a potential demonstration of subtle discrimination. Van Leeuwen (2003) argues that not all occurrences indicate social empowerment. In certain contexts, such as this, numerical visibility reflects a subtle manifestation of the society’s sustained patriarchal beliefs where men are considered the naturalized form; thus, the authority, expert, and legitimate professional, while women stand on the periphery as the notable and the “exception to the rule.” The constant need to identify women’s gender, like “female leaders,” “women entrepreneurs,” and “feminine farmers,” potentially validates that women are visible but not part of the presumed norm.

Table 2

Top Collocates of Feminine Gender Markers across Van Leeuwen’s Social Actor Categorization

Categorization	Collocations	LogDice Result	
Functionalization	initiators	8.42	
	students	8.41	
	workers	8.38	
	nurses	7.8	
	teachers	6.4	
	entrepreneurs	9.7	
	embroiders	8.05	
	leaders	7.78	
	farmers	7.74	
	archeologist	7.7	
	politicians	7.61	
OFW	7.05		
Classification	adolescents	9.79	
	young	9.28	
	married	8.9	
	unmarried	7.5	
	older	7.72	
	aged	7.45	
	old	6.71	
	married	8.9	
	Identification	children	8.44
		husband	9.12
Relational Identification		mother	8.58
		family	8.12
Physical Identification		parents	8.1
		daughter	7.77
Appraisement	pregnant	9.15	
	abused	7.71	
	survivors	7.7	
	battered	7.03	
	victims	7.23	

experienced	7.19
political	6.69

A deeper investigation of the collocational patterns within feminine markers shows two conflicting interpretations. Initially, Table 2 shows the diversity of roles that pattern with feminine markers. While there are stereotypical functions such as “embroiderers,” “nurses,” “teachers,” and “students,” statistically significant collocates were also observed to have been formed with functions that mark agency. Markers like “initiators,” “leaders,” “entrepreneurs,” “farmers,” “archeologists,” “politicians,” “OFWs,” and “workers” potentially echo the numerical data on gender equality among professional and skilled workers across various levels of educational background in the Philippines (Abad, 2025).

However, this diverse portrayal is seen as superficial when the markers of identification are integrated into the investigation. It can be observed that there are more stereotypical markers of classification, relation, and physical identification that pattern strongly with feminine markers. This emphasizes the potential complexity in the ways women are portrayed. While they are functionalized, they are also heavily identified using their inherent traits. In fact, the majority of these collocations depict biased ideologies that can restrict women instead of empowering them. For instance, a variety of collocations indicate age, such as “adolescent/s,” “young,” “older,” “aged,” and “old”. This detailed representation suggests that people still focus on women’s age. Such findings align with the previous research that revealed how the word “elderly” connects more with women, confirming that age is still a social taboo for them (Jevrić, 2017). Almost the same pattern is found in the investigation of the word “marriage” as a collocate of feminine markers, which potentially reveals that civil status continues to be one of the valuable criteria used to measure women’s worth (Solano, 2015).

Furthermore, “pregnant” as a physical identification, combined with a detailed list of collocations that relate to family life, uniquely patterns with different feminine markers. This potentially portrays women through their matrimonial and communal associations. The strong collocations of words that mark relational identification, such as “children,” “family,” “husband,” “mother,” “family,” “parents,” and “daughter,” can essentialize women, thereby risking them being limitedly defined by their reproductive capability and service to others. This resonates with the result of Dacon and Liu’s (2021) investigation, which highlighted the strong collocation of motherhood to females, and Almaghlouth (2022), who found that women are represented mostly through relational identification. Unfortunately, this might also mean that at present, a woman’s identity is portrayed in isolation, and, to a greater extent, they shoulder the multiple loads of work and domestic duties (Chen et al., 2017).

Lastly, appraisal-related data underscores the collocational bond between negative evaluation and female social actors. The significant associations of the terms “abused,” “survivors,” “battered,” and “victims” with the marker woman tend to show that women are portrayed through their vulnerability. While this can also be an attempt to highlight social issues, the persistence of these collocations may subtly reinforce

paternalistic views where women are essentialized as helpless, while the other gender is agentic. This result mirrors Xue’s (2025) observations of gendered roles in UN documents, where women are predominantly categorized as victims. Interestingly, masculine markers do not yield any collocation in this category. This reinforces a gendered hierarchy of representation. While women are subject to negative specification, men are shielded from the linguistic associations of victimization.

The overall data on top collocates of feminine markers points out the asymmetry in categorizing female social actors. On one hand, women are highly functionalized by portraying them with diverse roles and agency; on the other hand, they are still subjected to very detailed identification and negative appraisal. This suggests that while women are statistically visible in journal articles, their representation is conditional. They are intricately portrayed in the professional sphere, but they remain clearly defined by their social identifications. This indicates a traditional discursive construction that constrains rather than negates a valid form of gender parity.

Table 3

Top Collocates of Masculine Gender Markers across Van Leeuwen’s Social Actor Categorization

Categorization		Collocations	LogDice Result
Functionalization		migrant	8.45
		initiators	8.29
		participants	7.87
		students	7.79
		nurse	7.55
Identification	Classification	older	8.21
		young	9.88
		gay	9.16
		married	8.16
		Relational Identification	family
	Physical Identification	-	-
Appraisal		-	-

The analysis of markers found in CRAFT perpetuates the traditional portrayal of men—more functionalized than identified (Almaghlouth, 2022). Table 3 shows that masculine markers, although limited, strongly pattern with markers of roles and jobs rather than inherent traits. This potentially highlights their agency and the core tenet of Philippine hegemonic masculinity—the male’s role as the primary provider or breadwinner (Bagaporo & Papadelos, 2025). It also converges with international findings, such as those of Luo and Huang (2024), who observed the emphasis on men’s functions by being constantly addressed using their titles, while women are addressed by their full names.

Interestingly, amid the high occurrence of “migrant”, “Nurse” emerges as a collocation that potentially marks a progressive discursive shift. This pattern might be a subtle indication that the field is slowly diversifying its portrayal of nurses. Such observation is notable because men have been previously found to be inadequately interested in pursuing “Nursing” as a profession, given the stereotypes and other gender-related biases in the field (Paskaleva et al., 2020). Evidently, the “degendering” efforts using intrapersonal, interpersonal, and strategic adaptation have been working (Younas et al., 2021).

In contrast to feminine markers, which collocate with a variety of markers of identification, masculine markers strongly collocate only with a few indicators of age, civil status, gender, and relational identification. This restricted collocational range possibly supports the functionalization of men that emphasizes the stereotypical ways men are portrayed with agency rather than any of their physical or relational attributes. Combining the analysis of the two categories, the limited occurrence of masculine markers in CRAFT may not necessarily demonstrate a new trend in gender discourse, such as a neglect of men or a high level of feminization. In fact, this may suggest that Filipino scholars still portray men as the unmarked gender or the naturalized form; hence, the identification of their gender becomes irrelevant in the discourse. This is potentially supported by the minimal presence of functional and identification indicators that collocate with masculine markers. This shows that disparity in academic discourse might subtly lie in the portrayal of men as the naturalized form and the over-representation of women.

Overall, the results show that gender parity in the Philippines remains a complex linguistic negotiation. Women, on one hand, may have achieved statistical visibility, but it does not guarantee complete functionalization as they remain marked or exceptions to the norm, combined with over-representation using their inherent traits. Men, on the other hand, maintain a quiet dominance through their naturalization.

IV. DISCUSSION

This study emphasizes the paradox of gender parity in the Philippine academic discourse. The use of collocational analysis reveals that there is a surface-level dominance of feminine markers across eight out of ten fields of specialization. However, a deeper analysis guided by RSA shows that the statistical visibility acts as a discursive façade that obscures the reality of gendered portrayal in scholarly works. Instead of being a hallmark of gender parity or a new trend in gender discourse, the number shows the constant need for women to be identified as “women leaders,” “women entrepreneurs,” and “women archeologists.” This potentially renders the numerical data as a subtle indicator of discrimination where men maintain their dominance as the naturalized form, while women are continuously marked as the “exception to the rule” in many professional spheres. Furthermore, the overrepresentation of women presents a discursive asymmetry. While they are highly functionalized, they are also overly identified through their inherent characteristics. This potentially results in essentializing women in terms of their reproductive ability and natural traits while putting aside their agency.

In light of these findings, this investigation suggests a systemic review of the academic editorial policies and practices. Such an assessment is essential to encourage scholars to be more cautious and conscious in their gender representation. Eventually, they might turn out to be important agents who will overtly challenge the potential normalization of men through the adoption of more gender inclusive discourse. Furthermore, given the fact that journal articles serve as primary references for research production, academic instruction, and policy formulation, they must be positioned as a potential means of shaping the readers' ideologies. Hence, ensuring that these scholarly works are free from gender bias can make them a powerful tool for bridging the systemic gender gap. This study also recommends the intensification of gender sensitivity programs that can uproot even subtle indicators of linguistic bias. Although the Philippines is recognized for its efforts to address gender disparity, being a leading country in Eastern Asia and the Pacific region (Desiderio, 2024), a sustained awareness and reflexivity among scholars can help build an academe that fully contributes to a more inclusive and equitable production of knowledge.

While these implications highlight the potential of this study to inform institutional practice, they must be cautiously interpreted in light of certain limitations that also yield potential avenues of further investigation.

First, CRAFT contains only ten specialized fields. While it is not considered limited, it can be observed that it currently has a slight skew towards social sciences. This calls for incorporating more hard science-related disciplines to potentially elevate its representativeness. This addition is expected to enrich the current set of findings by making it not only more generalizable but also more nuanced. Second, the analysis used in this study relies on interpreting the presence or absence of linguistic markers. This interpretive subjectivity is inherent to these types of analyses, requiring the researcher to assign meaning to both linguistic voids and occurrences. A follow-up study using content analysis will help triangulate the findings and enrich the reliability of the results. Lastly, this research focused on limited pairs of gender binary markers. An extended discussion may be done by including more masculine and feminine markers such as gender-marked professional titles (e.g., waiter/waitress, actor/actress, landlady/landlord), gender-specific kinship terms (e.g., father/mother, daughter/son, uncle/aunt), and gender-coded adjectives. Furthermore, inclusivity may be promoted by examining the occurrence of non-binary markers like the singular use of "they" and "Filipinx". While the Filipinos have already achieved monumental success in bridging the gender gap, certain areas still need more attention.

V. DISCLOSURE STATEMENT

The author declares no competing financial or personal interests. While the author was a member of the research collective that developed the CRAFT corpus at De La Salle University-Manila, the current analysis was conducted independently and with a commitment to empirical objectivity.

VI. ACKNOWLEDGEMENT

This is to acknowledge the collaborative effort of the students pursuing their degree, Doctor of Philosophy in Applied Linguistics at the De La Salle University. Their collective effort under the guidance of Dr. Shirley N. Dita made CRAFT and this investigation possible.

VII. REFERENCES

- Abad, M. (2025, March 8). *IN NUMBERS: State of Filipino women in 2024*. Rappler. <https://www.rappler.com/philippines/numbers-state-filipino-women-2024/>
- Aguilar-Delavin, E. (2022). Philippine English modules during the COVID-19 pandemic: Their gender sensitivity and responsiveness. *International Journal of Multidisciplinary Applied Business and Education Research*, 3(9), 1798–1804. <https://doi.org/10.11594/ijmaber.03.09.19>
- Alkhamash, R., & Al-Nofaie, H. (2020). Do Saudi academic women use more feminized speech to describe their professional titles? Evidence from corpus. *Training, Language and Culture*, 4(3), 9–20. <https://doi.org/10.22363/2521-442X-2020-4-3-9-20>
- Almaghlouth, S. (2022). Mourning the lost: A social actor analysis of gender representation in the @FacesofCovid's tweets. *Frontiers in Psychology*, 13, Article 1042621. <https://doi.org/10.3389/fpsyg.2022.1042621>
- Amoniyani, O. M., & Obasi, J. C. (2025). A corpus study of gender identity representation in selected Nigerian newspapers. *Howard Journal of Communications*, 1–28. <https://doi.org/10.1080/10646175.2025.2476482>
- Angouri, J., & Baxter, J. (Eds.). (2021). *The Routledge handbook of language, gender, and sexuality*. Routledge.
- Aydın-Günbatır, S., & Roehrig, G. H. (2023). Elaborating nature of engineering through the family resemblance approach. *Science & Education*. <https://doi.org/10.1007/s11191-023-00452-5>
- Bagaporo, J. L. L., & Papadelos, P. (2025). Redefining breadwinning among internally displaced Filipino men in Southern Philippines. *The Journal of Men's Studies*. <https://doi.org/10.1177/10608265251349711>
- Biber, D., & Burges, J. (2000). Historical change in the language use of women and men. *Journal of English Linguistics*, 28(1), 21–37. <https://doi.org/10.1177/00754240022004857>
- Briones, A. C. D. (2019). Gender representation in Philippine junior high school English language textbooks. *Asia Pacific Journal of Multidisciplinary Research*, 7(4), 18–25.
- Brun-Mercer, N. (2021). Women and men in the United Nations: A corpus analysis of general debate addresses. *Discourse & Society*, 32(4), 443–462. <https://doi.org/10.1177/0957926521992145>
- Çağlak, I. (2025). Gender representation and sexism in international language coursebooks for tertiary education. *The Journal of Language Teaching and Learning*, 15(1), 20–45. <https://www.jltl.com.tr/index.php/jltl/article/view/727>

- Callaway, K. (2022). *Handsome women, gorgeous men: A corpus-based analysis of frequency, animacy and gender over time in adjectives of beauty*. The Linguistics Society at UGA. <https://openscholar.uga.edu/record/21641>
- Chaemsaitong, K. (2021). Representing people in execution news: Reference terms, identity, and ideology. *Journal of Pragmatics*, 183, 1–14. <https://doi.org/10.1016/j.pragma.2021.07.003>
- Chen, F., Bao, L., Lin, Z., Zimmer, Z., Gultiano, S., & Borja, J. B. (2017). Double burden for women in mid- and later life: Evidence from time-use profiles in Cebu, the Philippines. *Ageing & Society*, 38(11), 2325–2355. <https://doi.org/10.1017/S0144686X17000599>
- Dacon, J., & Liu, H. (2021). Does gender matter in the news? Detecting and examining gender bias in news articles. *Companion Proceedings of the Web Conference 2021*, 385–392. <https://doi.org/10.1145/3442442.3452325>
- Desiderio, L. (2024, June 14). *Philippines slips 9 notches in gender gap ranking*. Philippine Star. <https://www.philstar.com/headlines/2024/06/14/2362767/philippines-slips-9-notches-gender-gap-ranking>
- Dianitasari, F. (2021). Corpus-based analysis: “Women” for “leadership.” *Ethnolinguial*, 5(2), 126–141. <https://doi.org/10.20473/etno.v5i2.34155>
- Formanowicz, M., & Hansen, K. (2021). Subtle linguistic cues affecting gender in(equality). *Journal of Language and Social Psychology*, 41(2), 127–147. <https://doi.org/10.1177/0261927X211035170>
- Frankenberg-Garcia, A. (2018). Investigating the collocations available to EAP writers. *Journal of English for Academic Purposes*, 35, 93–104. <https://doi.org/10.1016/j.jeap.2018.07.003>
- Horton, R. H. (2018). A corpus analysis of *girl* and *boy* in spoken academic English and teaching activities to raise awareness about gendered discourse. *TESOL Working Paper Series*, 16, 1–18. <https://www.hpu.edu/research-publications/tesol-working-papers/2018/01horton.pdf>
- Jaramillo, A. M., Macedo, M., Oliveira, M., Karimi, F., & Menezes, R. (2025). Systematic comparison of gender inequality in scientific rankings across disciplines. *arXiv*. <https://doi.org/10.48550/arXiv.2501.13061>
- Jevrić, T. (2017). “Medical men” and “mad women”: A study into the frequency of words through collocations. [sic] – *A Journal of Literature, Culture and Literary Translation*, 8(1). <https://doi.org/10.15291/sic/1.8.lc.2>
- Larregue, J., & Nielsen, M. W. (2023). Knowledge hierarchies and gender disparities in social science funding. *Sociology*, 58(1), 45–65. <https://doi.org/10.1177/00380385231163071>
- Llorens, A., Tzovara, A., Bellier, L., Bhaya-Grossman, I., Bidet-Caulet, A., Chang, W. K., ... Dronkers, N. F. (2021). Gender bias in academia: A lifetime problem that needs solutions. *Neuron*, 109(13), 2047–2074. <https://doi.org/10.1016/j.neuron.2021.06.002>
- Luo, X., & Huang, C. (2024). Analyzing the gendered power dynamics in addressing practices: A corpus-based approach. *Proceedings of the 38th Pacific Asia Conference*

- on Language, Information and Computation, 1259–1267.
<https://aclanthology.org/2024.paclic-1.122/>
- Mosqueda, H. C., & Sánchez, I. R. (2022). Sexist stereotypes in the Spanish-speaking press: A collocational analysis of the lemmas *man* and *woman*. *Revista de Humanidades Digitales*, 7, 57–79. <https://doi.org/10.5944/rhd.vol.7.2022.34108>
- Mustofa, R., Zawawi, M., Jambak, M. R., Ghufroon, Z., & Wirmansyah, A. F. (2025). Social actors in the news of Iran’s attack on Israel in mass media: Theo van Leeuwen’s critical discourse analysis. *Lisanudhad: Jurnal Bahasa, Pembelajaran, dan Sastra Arab*, 12(1), 153–184. <https://doi.org/10.21111/lisanudhad.v12i1.13384>
- Orgeira-Crespo, P., Míguez-Álvarez, C., Cuevas-Alonso, M., & Rivo-López, E. (2021). An analysis of unconscious gender bias in academic texts by means of a decision algorithm. *PLoS ONE*, 16(9), Article e0257903.
<https://doi.org/10.1371/journal.pone.0257903>
- Paskaleva, T., Tornyova, B., & Shopov, D. (2020). Professional status of men in nursing. *Journal of IMAB – Annual Proceeding (Scientific Papers)*, 26(1), 3029–3033.
<https://doi.org/10.5272/jimab.2020261.3029>
- Pearce, M. (2008). Investigating the collocational behaviour of *man* and *woman* in the BNC using Sketch Engine. *Corpora*, 3(1), 1–29.
<https://doi.org/10.3366/E174950320800004X>
- Posch, C. (2023). Half-witted or hard-working-fun-loving women? A corpus-assisted study of gendered collocation in the New Zealand Alpine Club Journal Corpus. *Zeitschrift für Anglistik und Amerikanistik*, 71(3), 241–263.
<https://doi.org/10.1515/zaa-2023-2033>
- Raphael, E. B. (2023). Gendered representations in language: A corpus-based comparative study of adjective-noun collocations for marital relationships. *Theory and Practice in Language Studies*, 13(5), 1191–1196.
<https://doi.org/10.17507/tpls.1305.12>
- Rashidi, N., & Ghaedsharafi, S. (2015). An investigation into the culture and social actors representation in Summit series ELT textbooks within van Leeuwen’s framework. *SAGE Open*, 5(1). <https://doi.org/10.1177/2158244015576054>
- Raza, S., Niazi, M. M., & Saba, N. (2025). Exploring linguistic sexism in students’ written academic discourse of the Department of English at KUST. *Dialogue Social Science Review*, 3(1), 184–195.
<https://dialoguessr.com/index.php/2/article/view/161>
- Saher, A., Mehmood, R., & Farhad, A. (2023). Corpus-based investigation of gender representation in education: A content analysis of Pakistani and Indian high secondary school textbooks. *Journal of Development and Social Sciences*, 4(1).
[https://doi.org/10.47205/jdss.2023\(4-i\)03](https://doi.org/10.47205/jdss.2023(4-i)03)
- Solano, R. M. (2015). Man/woman versus hombre/mujer: A contrastive analysis of compound nouns, collocations and collocational frameworks. *Il, Elle : Entre Je(u)*.
<https://hal.science/hal-01645072v1>
- Stroble, M., & Lebani, G. (2022). *Use of gender-neutral language in respect of role nouns in Austria’s media: A corpus-based analysis*. Austrian Media Corpus.

https://amc.acdh.oeaw.ac.at/assets/content/pages/publikationen/Seminararbeit_Strobl_Gender_Neutral_Language_2023.pdf

- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. John Wiley & Sons.
- Tremmel, M., & Wahl, I. (2023). Gender stereotypes in leadership: Analyzing the content and evaluation of stereotypes about typical, male, and female leaders. *Frontiers in Psychology, 14*, Article 1034258. <https://doi.org/10.3389/fpsyg.2023.1034258>
- UN Women & United Nations Department of Economic and Social Affairs. (2025). *Progress on the Sustainable Development Goals: The gender snapshot 2025*. <https://unstats.un.org/sdgs/gender-snapshot/2025/GenderSnapshot2025.pdf>
- Utama, G. N., Amalia, R. M., & Yuliawati, S. (2020). Social actor representation of Indonesian migrant workers' executions in Saudi Arabia. *International Journal of Humanity Studies, 4*(1), 15–27. <http://dx.doi.org/10.24071/ijhs.v4i1.2170>
- Van Leeuwen, T. (2003). The representation of social actors. In C. R. Caldas-Coulthard & M. Coulthard (Eds.), *Texts and practices: Readings in critical discourse analysis* (pp. 32–70). Taylor & Francis. <https://doi.org/10.4324/9780203431382>
- Xue, J. (2025). Men as offenders while women as victims: A corpus-based study of men and women in the United Nations. *Frontiers in Communication, 10*. <https://doi.org/10.3389/fcomm.2025.1535312>
- Younas, A., Ali, N., Sundus, A., & Sommer, J. (2021). Approaches of male nurses for degendering nursing and becoming visible: A metasynthesis. *Journal of Clinical Nursing, 31*(5–6), 467–482. <https://doi.org/10.1111/jocn.15958>
- Zasina, A. J. (2019). Gender-specific adjectives in Czech newspapers and magazines. *Journal of Linguistics/Jazykovedný časopis, 70*(2), 299–312. <https://doi.org/10.2478/jazcas-2019-0060>
- Zhao, X. (2022). A corpus study of gender representation in Chinese secondary school EFL textbooks. *International Journal on Studies in English Language and Literature, 10*(11), 8–18. <https://doi.org/10.20431/2347-3134.1011002>