‘But this one was so . . . male.’ A Corpus-Based and LLM-Augmented Analysis of Language and Gender Bias in *Barbie*

Xin Luo, Wing-hei Lok, Yuyin-Hsu

Proceedings of the 39th Pacific Asia Conference on
Language, Information and Computation (PACLIC 39)

Emmanuele Chersoni, Jong-Bok Kim (eds.)

2025

‘But this one was so . . . male.’ A Corpus-Based and LLM-Augmented Analysis of Language and Gender Bias in *Barbie*

Xin Luo[1] and Wing-hei Lok[2] and Yuyin-Hsu[3]
Department of Language Science and Technology
The Hong Kong Polytechnic University
[1]tracyxin.luo@polyu.edu.hk, [2]winghei-lok@connect.polyu.hk,
[3]yu-yin.hsu@polyu.edu.hk

Abstract

The 2023 film *Barbie* has sparked discussions on women’s empowerment and patriarchal norms, weaving feminist themes with critiques of gender roles while subtly reflecting patriarchal undertones that marginalize women (Myisha et al., 2023). We conduct a corpus-based analysis to investigate gender bias and differences in utterances’ distribution of part-of-speech (POS), affective values, and gender-linked classifications across three distinct scenes: 1) Barbie Land, 2) Real World, and 3) Post-Barbie Land, each representing the matriarchal, conventional and patriarchal theme, respectively. Leveraging large language models (LLMs), we extend Bradley and Lang’s (1999) affective norms to assess the emotional properties of film scripts, marking a novel application of LLMs in film script analysis.

Findings reveal gender-based differences and reflect power dynamics in and across scenarios, supporting prior research regarding gendered lexical preferences (e.g., Argamon et al., 2003; Jasmani et al., 2011). The results of affective analysis indicate that females consistently demonstrate higher valence and arousal scores across scenes, aligning with feminine communication styles emphasizing emotional rapport and expressivity, whereas males prioritize dominance (Tannen, 1990; Holmes and Stubbe, 2003). These variations underscore how affective language use is co-constructed by gender and socio-narrative context. Notably, gender-linked classifications (Barbie/Ken tones) align more closely with theoretically grounded gendered linguistic features than the broad stereotypes. This study highlights the linguistic construction of gender roles and power dynamics in *Barbie*, offering insights into the interplay between subversive feminism and traditional patriarchal norms and demonstrating how language reflects and challenges gender stereotypes in the film.

1 Introduction

The 2023 film *Barbie* emerges as a cultural product that simultaneously promotes feminist narratives to mass audiences while critically engaging with persistent gender biases and patriarchal stereotypes - embodying what Byrnes (2025) terms ‘feminist ambivalence’. Within this framework, the language employed in the movie is designed to resonate with post-feminist ideas that emphasize women’s individuality, autonomy, and inherent uniqueness as distinct self-determining beings (Gill, 2007).

Previous studies have identified the distribution of POS as a key indicator of gendered language, with specific POS patterns reflecting distinct gendered linguistic features. Key’s (1972) foundational work on gender differences in linguistic behavior first observed variations in pronominal and nominal referents between male and female speakers. Lakoff (1975) later contextualized these differences within broader social hierarchies, arguing that they stem from women’s historically marginalized status – a dynamic that incentivizes women to prioritize linguistic markers of prestige, including more standardized forms than men (Trudgill, 1972; Labov, 1990).

Subsequent studies have elaborated on gendered POS tendencies. The feminine style, for instance, is characterized by a greater use of emphatic adverbs, such as *so, really* (Schofield and Mehr, 2016) and modal auxiliaries that soften declarative statements (McMillan et al., 1977; Biber et al., 1998; Mulac et al., 2001; Mehl and Pennebaker, 2003). These patterns link to Biber et al.’s (1998) framework of ‘involved’ (subjective, interpersonal) and ‘informative’ (objective, content-focused) speech, where feminine styles tend towards the former. Pronouns, one of the widely studied POS categories, reveal robust gender differences particularly in generic usage (e.g., *he, they*) and occupational referents (e.g., *nurse, doctor, president*) (Luo and Huang,

2024). Nouns, however, show more variable patterns. Binnenpoorte et al. (2005) found no significant gender-based differences in noun usage within telephone dialogues, while Argamon et al. (2003) documented notable disparities in fictional texts. This inconsistency underscores the critical role of contextual factors in shaping gendered language patterns and highlights the need for context-sensitive analysis in POS-based gender research.

Beyond POS distribution, this study examines gendered language through two dimensions: the Affective Norms of English Words (ANEW) and the Barbie/Ken tone (hereafter 'BT/KT'), categorized using the LLMs. Affective connotation of words exert a pervasive influence on cognitive processing, and a growing body of research documents gender-based disparities in how such words are employed. A critical challenge in this domain lies in selecting appropriate terms – a choice that underscores the importance of measuring words' affective meanings for scholars investigating gendered language and its role in shaping power dynamics. While existing literature primarily focuses on generating affective ratings for words across languages, including exploring how gender influences participants' evaluations (Warriner et al., 2013; Montefinese et al., 2014), far fewer studies investigate how ANEW-related words are employed differently by males and females. Notably, gender has been shown to correlate significantly with arousal and dominance ratings, aligning with the stereotype that women are perceived as more emotionally expressive than men (e.g., Fischer, 2000). One relevant investigation is Marrville's (2017) thesis, which found that verbs associated with high emotional dominance were strongly association with male characters, whereas those conveying low emotional control were more closely associated with female characters. However, this study is limited in scope focusing exclusively on verbs and examines only the dominance dimension with a self-paced reading task.

Gender stereotypes serve as a foundational lens in gender research, as they encapsulate culturally normalized perception of 'appropriate' behaviors for different genders. These stereotypes often manifest as dichotomous constructs, framing gendered features as distinct and mutually exclusive (Holmes and Stubbe, 2003). For instance, women communicate indirectly, while males adopt a direct communication style. However, scholars have increasingly challenged the rigidity of such dichotomy, noting that gendered linguistic features rarely align perfectly with stereotypes across all contexts. Cameron (1998) attributed this inconsistency to the collaborative nature of conversation, emphasizing that interactional dynamics, rather than fixed gender traits, shape communicative behavior. Similarly, Bing and Bergvall (1998) argued that male and female language use and behavior exist on an overlapping continuum, resisting strict categorization. This study employs a corpus-based approach, augmented by LLMs, to investigate gendered linguistic patterns and affective language in the 2023 film *Barbie*. Focusing on the tension between feminist and patriarchal ideologies, we analyze three key linguistic features: distribution of POS, affective values of word choices, and gender-linked classifications across different thematic scenarios.

2 Data and Methodology

The primary dataset comprises the complete transcript of the 2023 film *Barbie*, retrieved directly from official screenplay records. This source was selected for its authoritative representation of gendered dialogue, which explicitly navigates tensions between feminist and patriarchal ideologies. Text processing was processed using the Natural Language Toolkit (NLTK; Bird et al., 2009) to ensure systematic and replicable data preparation. The details of the data are summarized in Table 1. To analyze gendered language across contrasting power structures, the transcript was strategically subdivided into three sub-scenarios for targeted thematic and linguistic analysis: 1) Barbie Land (PreB) which was characterized by a matriarchal social order where female characters hold positions of power and agency, prior to narrative disruptions; 2) Real World (RW), capturing the gendered dynamics in a societal context that reflects conventional patriarchal norms; and 3) Post-Barbie Land (PostB), representing the altered Barbie Land under patriarchal influence. This subdivision captures the film's deliberate thematic juxtaposition—allowing direct comparison of linguistic patterns under matriarchal, patriarchal, and transitional power dynamics.

Drawing on established research documenting gendered differences in lexical and POS distributions (Argamon et al., 2003; Litvinova et al., 2017), we hypothesized that POS usage would exhibit significant variation across the film's sub-scenarios

	Females		Males	
Scenarios	Tokens	Word	Tokens	Word
PreB	2308	612	483	185
RW	2265	612	1626	524
PostB	3766	794	1874	598
Total	8339	2018	3983	1307

Table 1: Description statistics of dataset

(PreB, RW, PostB), reflecting their distinct gendered linguistic contexts. This hypothesis aligns with Key's (1972) proposal that grammatical categories serve as linguistic markers of gendered performance, which may be amplified or transformed in narratives explicitly addressing gender norms. Our POS analysis targeted two levels of granularity to capture both specific and broader functional patterns, including individual POS categories and functionally relevant POS groups (e.g., adjectives + adverbs (JJs+RBs); all verb forms (VBs)). We adopted the Affective Norms of English Words (ANEW) framework proposed by Bradley and Lang (1999), which operationalizes emotions along three dimensions: valence (ranging from 1 = unpleasant to 9 = pleasant), arousal (1=calm to 9 = excited), and dominance (1 = out of control to 9 = under control). To situate POS patterns within emotional and evaluative contexts, we applied this framework with established psycholinguistic norms (valence, arousal, and dominance) to evaluate 843 NNs and 361 adjectives and adverbs. We further augmented the analysis with LLM-based methods to extend ANEW's coverage and contextualize the affective scores within the narrative of *Barbie*.

To extend our study beyond traditional linguistic metrics, we employed GPT-3.5 Turbo to categorize the same 1204 target words (843 NNs and 361 adjectives and adverbs) used in our POS and affective analyses into two distinct tonal registers (BT or KT). This classification aimed to capture LLM's underlying words stereotypically associated with feminine (BT) and masculine (KT) linguistic style, respectively. To systematically investigate how different contextual frameworks influence the classification of gendered language, we designed three distinct prompts for a zero-shot classification task. All prompts were presented without examples to elicit the LLM's inherent reasoning and each prompt directly instructs the model to adopt a specific perspective for the 'BT/KT tone' classification task:

1. Film-contextualized (FC): Classify words as 'Barbie tone' or 'Ken tone' based on their usage in the film;

2. Stereotype-driven (SD): Classify words based on widespread gender stereotypes, disregarding movie context;

3. Theory-guided (TG): Classify word using Holmes and Stubbe's (2003) classification of masculine and feminine linguistic features, ignoring both movie context and stereotypes.

3 Results and Discussions

3.1 Distribution of POSs

Significant differences in POS usage emerged between male and female characters across the dataset. As illustrated in Figure 1, male characters demonstrated a significantly higher proportion of proper nouns (NNP) (χ^2=4.8647, p=0.0274), reflecting their tendency to reference specific entities or roles. Female characters employed past tense verbs (VBD) significantly more than male characters (χ^2=4.434, p=0.0352), suggesting a focus on narrating events or experiences. These patterns coorborate established gendered linguistic tendencies. The over-representation of nouns in male language use replicates the finding by Argamon et al. (2003), who linked noun dominance to men's stereotypically referential speech in formal contexts. The higher frequency of verbs in female speech corresponds to observations in Jasmani et al. (2011) regarding females' greater use of verbal constructs often tied to relational storytelling.

Analysis of the three sub-scenarios revealed significant context-dependent variations in gendered POS patterns, demonstrating how linguistic behavior adapts to sociocultural power structures. In PreB, female characters used significantly more common nouns (NN) (e.g., *president, doctor*) than males (χ^2=6.4358, p=0.0112) (Figure 2). This aligns with the scenario's thematic focus on female-centric agency, where common nouns often referenced communal roles central to Barbie Land's social structure. No statistically significant gender differences emerged in the distribution of POSs within RW (ps > 0.2435). This absence of divergence may reflect the scenario's portrayal of conventional societal norms, where gendered linguistic markers are less exaggerated than in the film's more ideologically charged settings. A striking

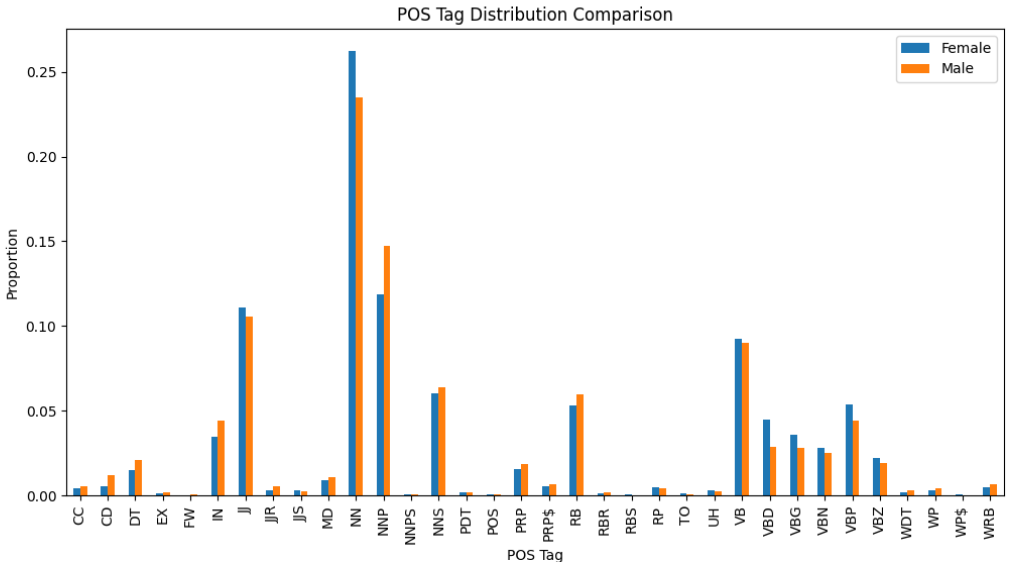

Figure 1: Distribution of POS in overall dataset

disparity was observed in proper noun (NNP) usage, with male characters displaying a significantly higher proportion of NNP in PostB (χ^2=8.0958, p=0.0044) (Figure 3). This pattern intensifies the overall trend of male-centric proper noun use, mirroring the scenario's emphasis on male dominance.

A close analysis of word frequencies within gender groups uncovered nuanced variations in lexical choice. Male characters predominantly referenced male-specific names (e.g., *Ken, Mojo, Aaron*) while female characters frequently referenced *Barbie* in their utterances. This gendered specificity in proper noun usage aligns with Bing and Bergvall's (1998) assertion that language reflects and reproduces power imbalances – centering the names of dominant groups within each scenario–females in PreB, and males in PostB. Female characters showed a marked preference for past tense auxiliary verbs (e.g., *was, did*), whereas male characters rarely used such forms. This aligns with classic hypothesis about gendered language, particularly Lakoff's (1975) observation that women's speech often employs hedging devices, including auxiliaries and Newman et al.'s (2008) finding that female language tends to emphasize relational or retrospective contexts, a pattern reflected in the preferential use of past tense constructions. Furthermore, gendered patterns extended to interpersonal language, with males frequently using terms such as *sir* and *man*, reinforcing hierarchical or peer-based masculinity, while females more commonly utilized the term *mom*, emphasizing relational or familial roles. Taken together, these findings suggest that POS distribution in *Barbie* is not only gendered but also contextually contingent, with power dynamics in each scenario (matriarchal, conventional, patriarchal contexts) shaping the linguistic strategies employed by male and female characters.

3.2 Gendered affective preferences

To ensure a robust and multi-faceted evaluation of affective norms, we leveraged three distinct LLMs, including GPT4o[1], deepseek-v3-fw[2], and Llama[3] to rate the valence, arousal, and dominance of target words. This selection provides architectural and training diversity, enabling cross-model validation and reducing the potential bias inherent in any single model's output. The results of LLM-based ratings showed no statistically significant discrepancies ($p > .05$), confirming the inter-model reliability for evaluating the affective norms of English words.

ANOVA revealed no statistical significant differences in overall affective norm scores between males and female tones ($p > .05$), suggesting a balanced distribution of emotional word usage between genders in the aggregate. However, nuanced gender differences emerged in interaction within scene context (PreB, RW, PostB) and POS categories (adjectives and adverbs (JJs+RBs) and nouns (NNs)).

[1] http://openai.com
[2] https://www.deepseek.com/
[3] https://www.llama.com/

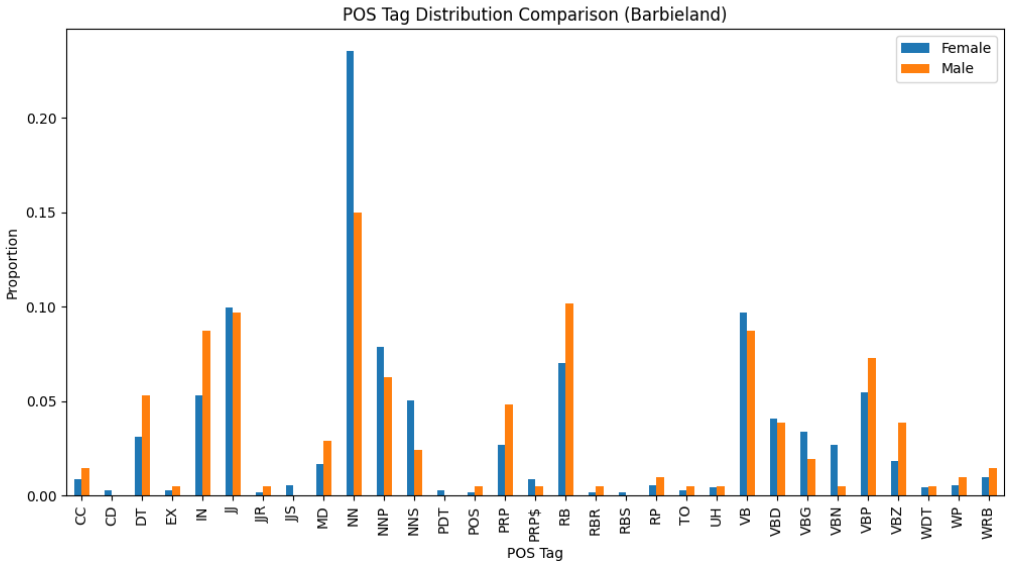

Figure 2: Distribution of POS in PreB

Gender differences in the affective norms of JJs+RBs were most prominent in scenario-specific analyses with significant effects limited to the PreB scene. In PreB as shown in Figure 4, females exhibited significantly higher valence scores for adjectives and adverbs compared to males. Their top 5 words *amazing, beautiful, fun, happy, perfect* – aligned with PreB's joyful, carefree narrative context, where female characters occupied central roles without external constraints. Females also showed elevated dominance scores in JJs+RBs usage, driven by terms such as *powerful, brave, free, perfect, and wonderful*. These lexical choices reflect PreB's thematic focus on female leadership and societal dominance. No statistically significant gender differences in JJs+RBs affective norms were observed in RW and PostB scenes ($p > .05$).

Gender differences in noun usage (NNs) across affective norms were more consistent across scenes, with males generally showing higher dominance scores (Figure 5 and 6). Their top 5 terms contributing to this pattern - *impeccable, best, professional, liberated, rich* – conveyed authority and control. No statistically significant differences were observed in noun valence, arousal, or dominance in PreB ($p > .05$), which aligns with the emphasis of the scene on collective female empowerment over individual hierarchy. Females' noun usage showed higher arousal scores in RW, driven by terms with negative emotional connotations, such as *destroy, anxiety, crisis, death, fascist*. This lexical pattern reflects Barbie's self-reflexive turmoil, blending postfeminist empowerment with underlying existential dread. In contrast, males' nouns in RW, such as *imagination, power, dreams, friends, understanding*, highlighted self-perceptions as leaders and significant figures, resulting in higher dominance scores. Females' nouns had higher valence and arousal scores, while males maintained higher dominance scores in noun usage in PostB.

Females consistently demonstrated higher valence and arousal scores across scenes, reflecting a linguistic emphasis on emotional rapport and expressive intensity – traits associated with feminine communication styles (Holmes and Stubbe, 2003). Males, in contrast, maintained higher dominance scores in noun usage across contexts, except in PreB, where females' elevated dominance in JJs+RBs mirrored Barbie Land's narrative of female societal leadership. These findings underscore how affective language use is co-constructed by gender and socio-narrative context.

3.3 Classification of 'Barbie Tone (BT)' and 'Ken tone (KT)'

GPT-3.5 Turbo was employed to classify all 361 JJs+ RBs into either 'BT' or 'KT' across the three experimental prompts. In Table 2, a significant majority of JJs+RBs were classified as BT, far exceeding KT classification in the film-contextualized classification (237 vs. 124). In contrast, stereotypes-driven classification favored KT over BT (150 vs. 211). Theoretically grounded classification leaned toward BT over KT.

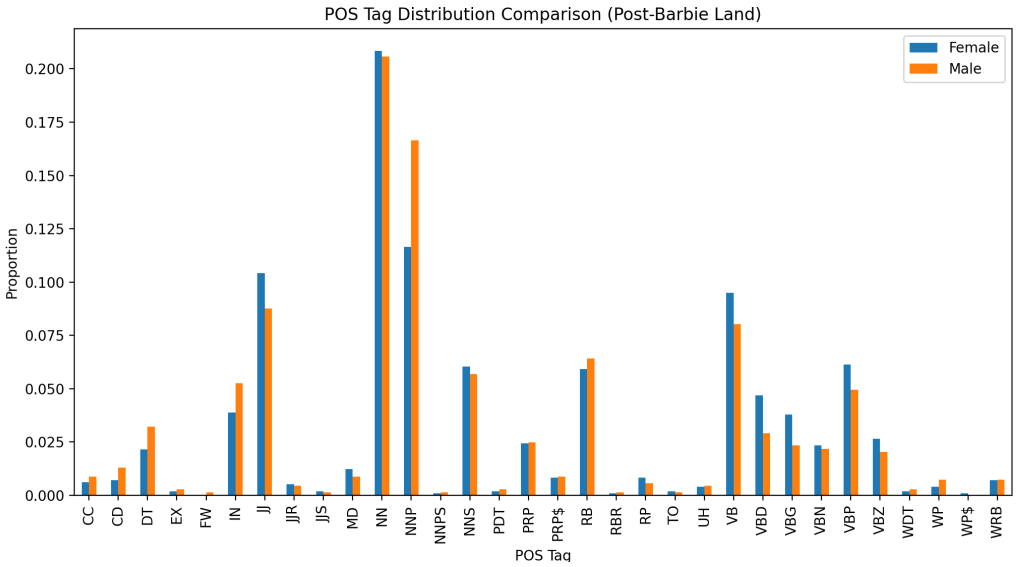

Figure 3: Distribution of POS in PostB

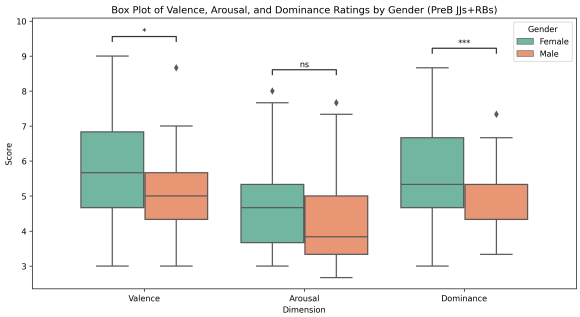

Figure 4: Affective norms value of JJs+RBs in PreB

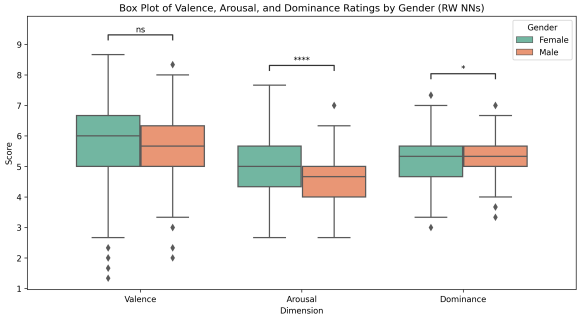

Figure 5: Affective norms value of NN in RW

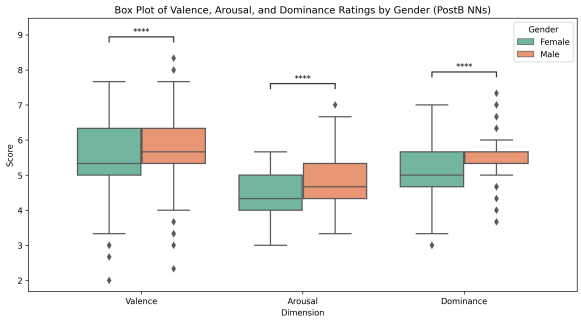

Figure 6: Affective norms value of NN in PostB

A highly significant difference between the film-contextualized (FC) and stereotypes-driven (SD) classification (χ^2=41.1887, $p < .001$), reflecting divergent classification patterns when contextualized film language versus general gender stereotypes. Consistent with Byrnes's (2025) concept of feminist ambivalence in *Barbie*, our quantitative result reveals no statistically significant difference between the FC and theory-guided (TG) classifications after Bonferroni correction for multiple comparisons (χ^2=3.9705, $p = 0.0463 > Bonferroni$ α=0.00167). This statistical alignment suggests the film's linguistic patterns organically embody the theoretical tension between feminine and masculine features. A significant discrepancy between stereotype-driven and theory-guided classifications (χ^2=19.2855, $p < .001$), indicating that general gender stereotypes diverge from theoretically grounded gendered linguistic features. Notable lexical discrepancies emerged for terms such as *com-*

Type	IA	FC	SD	TG
BT	361	237	150	210
KT	0	124	211	151
Total	361	361	361	361

Table 2: Classification of JJs+RBs as "Barbie tone" or "Ken tone" under different prompts

Note: Intrinsic Association (IA) refers to the model's classification based solely on its pre-existing knowledge, without any contextual prompt pr instructional framing.

Type	IA	FC	SD	TG
BT	840	561	335	451
KT	3	282	508	392
Total	843	843	843	843

Table 3: Classification of NNs as "Barbie tone" or "Ken tone" under different prompts

Note: Intrinsic Association (IA) refers to the model's classification based solely on its pre-existing knowledge, without any contextual prompt pr instructional framing.

plicated, exactly, pregnant. These were frequently categorized as KT under stereotypes-driven classification but aligned with BT in both the movie context and theory-guided classification. This divergence highlights a disconnect between culturally dominant stereotypes and actual linguistic features in the film, as well as theoretically defined gendered pragmatics (e.g., *expressive vs. assertive language*, Holmes and Stubbe, 2003).

GPT-3.5 Turbo demonstrated a strong overall bias toward classifying NNs as 'Barbie tone', labeling 840 out of 843 target nouns with this category and assigning only 3 as 'Ken tone' (Table 3). However, prompt-specific analyses revealed substantial variation in classification patterns. When comparing results between FC and SD prompts, a clear majority of nouns were classified as 'Barbie tone' in FC, the SD prompt triggered a sharp shift towards 'Ken tone' classifications. This difference was statistically significant (χ^2=120.5834, $p < .001$). Notably, stereotype-driven 'Ken tone' nouns include *CEO, airplanes, assistant, and autonomy* – terms linked to traditional masculine domains. A similar contrast emerged when comparing FC and TG prompt results, significantly more NNs were classified as 'Barbie tone' in the movie context (561 vs. 451), while the TG condition produced a higher proportion of 'Ken tone' classifications (282 vs. 392, χ^2=29.3677, $p < .001$). The comparison between SD and TG prompt results revealed that stereotype-driven classifications yielded far more 'Ken tone' labels (508 vs. 392), whereas the theory-guided classification resulted in more 'Barbie tone' (335 vs. 451, χ^2=31.5201, $p < .001$).

The analysis revealed no significant difference between the movie context and theory-guided classification for JJs and RBs (χ^2=41.1887, $p < .001$), indicating strong consistency between the film's linguistic portrayal and theoretically defined gendered features. This consistency indicates that the *Barbie* movie's use of descriptive language closely mirrors established linguistic framework (e.g., (Holmes and Stubbe, 2003)). Across both word categories, stereotype-based classifications differed significantly from both the movie context and theory-guided classifications, reflecting a clear mismatch between broad societal stereotypes and either the actual language of the film or theoretically grounded gendered linguistic traits. In particular, for NNs, even the movie context and Holmes and Stubbe's (2003) classification differed significantly, indicating nuanced distinctions in how gendered tones are operationalized in film versus linguistic theory. The tendency is consistent with Byrnes's (2025) concept of feminist ambivalence in *Barbie* that the film's linguistic patterns organically embody the theoretical tension between feminine and masculine features.

4 Conclusion

This study contributes to gendered language research by leveraging the 2023 film *Barbie* as a unique analytical lens to examine the interplay between feminist discourse and linguistic gendered performance. Consistent with Byrnes's (2025) interpretation of feminist ambivalence in this film, we demonstrate how the film's linguistic patterns organically embody contradictions between material culture and patriarchal stereotypes. Through POS distribution, affective norms and gendered-linked style classifications, we uncovered nuanced patterns that illuminate the contextual contingency of gendered language. Our findings revealed that gendered linguistic expressions are profoundly shaped by narrative context. Most notably, male characters employed more proper nouns (NNP), while female characters favored past tense verbs (VBD), consistent with prior research regarding gendered lexical preferences (e.g., Argamon et al., 2003; Jasmani et al., 2011). The results not only confirm the presence of gendered language in *Barbie* but also underscore that linguistic variations between gen-

ders are not static but dynamically co-constructed with socio-narrative environments.

Affective norms analyses further reinforced links between language and gendered communication styles. Females consistently exhibited higher arousal and valence scores across scenarios, aligning with Tannen's (1990) conceptualization of 'rapport talk', a speech style stereotypically associated with femininity. In contrast, males' elevated dominance scores resonated with 'report talk', reflecting a focus on authority and control. In particular, these patterns persist even as the film overtly promoted feminist ideals, indicating that *Barbie* retains and reinforces stereotypical gender norms through these linguistic patterns. This observation was further corroborated by the analysis of gendered tone (Barbie tone and Ken tone), which showed stronger alignment with theoretically grounded gendered features in Holmes and Stubbe (2003) than with broad social stereotypes, suggesting the films' portrayal of gendered language is nuanced, if not entirely subversive. These findings carry implications for understanding gendered communication in media and society. By demonstrating how narrative context shapes linguistic expressions of gender, our study highlights the limitations of decontextualized analyses of gender differences. The film's dual role in promoting feminism while retaining stereotypical linguistic patterns also underscores the complexity of media's role in reproducing or transforming gender forms. This study, while offering novel insights into gendered linguistic patterns and affective language in *Barbie*, is subject to several limitations that warrant consideration. First, regarding the scope of our data, the analysis is restricted to a single film and its scripted dialogue. This design, may limit the generalizability of the findings to broader media representation of gender and cannot capture the spontaneous nature of real-world gendered discourse. Second, concerning our computational methodology, our approach is primarily applicable to the English language only. Our experiments with LLMs aim to study how these models classify gendered words, as a general observation of the performance of LLMs. While our approach offers a robust method to augment and enrich the affective norms of words relevant to our research interests - particularly for cases where human responses are not yet available as a gold standard - it comes with limitations. We believe this methodology can be considered only for studying the English language. Applying it to other languages should be approached with caution, as most LLMs have been predominantly and, at times, biasedly trained on English-language data. Despite these limitations, our methodology offers a viable framework for investigating gendered language in narrative settings. Future research will extend this work by analyzing a broader corpus of media texts (e.g., *films, talk shows*) and comparing cross-cultural representations of gendered speech. By bridging computational linguistics and gender studies, this approach advances our ability to unpack the subtle, context-dependent ways gender is constructed and communicated through language.

References

Shlomo Argamon, Moshe Koppel, Jonathan Fine, and Anat Rachel Shimoni. 2003. Gender, genre, and writing style in formal written texts. *Text & Talk*, 23(3):321–346.

Douglas Biber, Susan Conrad, and Randi Reppen. 1998. *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

Janet M. Bing and Victoria L. Bergvall. 1998. The question of questions: beyond binary thinking. In Jennifer Coates, editor, *Language and Gender - A Reader*, pages 495–510. Blackwell Publishing Ltd.

Diana Binnenpoorte, Christophe Van Bael, Els den Os, and Lou Boves. 2005. Gender in everyday speech and language: a corpus-based study. In *INTERSPEECH*, pages 2213–2216.

Steven Bird, Ewan Klein, and Edward Loper. 2009. *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.".

Margaret M. Bradley and Peter J. Lang. 1999. Affective norms for english words (anew): Instruction manual and affective ratings. Technical report, Technical report C-1, the center for research in psychophysiology.

Alicia Byrnes. 2025. Surface and depth: ambivalence as postfeminist ideal in barbie. *Feminist Media Studies*, 25(3):767–773.

Deborah Cameron. 1998. Performing gender identity: young men's talk and the construction of heterosexual masculinity. In Jennifer Coates, editor, *Language and Gender - A Reader*, pages 270–284. Blackwell Publishing Ltd.

Agneta Fischer. 2000. *Gender and emotion: Social psychological perspectives*. Cambridge University Press.

Rosalind Gill. 2007. Postfeminist media culture: Elements of a sensibility. *European journal of cultural studies*, 10(2):147–166.

Janet Holmes and Maria Stubbe. 2003. "feminine" workplaces: stereotype and reality. In Janet Holmes and Miriam Meyerhoff, editors, *The Handbook of Language and Gender*, pages 573–595. Blackwell Publishing Ltd.

Mohd Faeiz Ikram Mohd Jasmani, Mohamad Subakir Mohd Yasin, Bahiyah Abdul Hamid, Yuen Chee Keong, Zarina Othman, and Azhar Jaludin. 2011. Verbs and gender: The hidden agenda of a multicultural society. *3L: The Southeast Asian Journal of English Language Studies*, 17(Special Issue):61–73.

Mary Ritchie Key. 1972. Linguistic behavior of male and female. *Linguistics*, 10(88):15–31.

William Labov. 1990. The intersection of sex and social class in the course of linguistic change. *Language variation and change*, 2(2):205–254.

Robin Lakoff. 1975. *Language and Woman's Place*. Harper & Row, New York.

Tatiana Litvinova, Pavel Seredin, Olga Litvinova, and Olga Zagorovskaya. 2017. Differences in type-token ratio and part-of-speech frequencies in male and female russian written texts. In *Proceedings of the Workshop on Stylistic Variation*, pages 69–73.

Xin Luo and Chu-Ren Huang. 2024. Analyzing the gendered power dynamics in addressing practices: A corpus-based approach. In *Proceedings of the 38th Pacific Asia Conference on Language, Information and Computation*, pages 1259–1267, Tokyo, Japan. Tokyo University of Foreign Studies.

Caelan Marrville. 2017. Gender and dominance in action: World view and emotional affect in language processing and use.

Julie R. McMillan, A. Kay Clifton, Diane McGrath, and Wanda S Gale. 1977. Women's language: Uncertainty or interpersonal sensitivity and emotionality? *Sex roles*, 3(6):545–559.

Matthias R. Mehl and James W. Pennebaker. 2003. The sounds of social life: a psychometric analysis of students' daily social environments and natural conversations. *Journal of personality and social psychology*, 84(4):857–870.

Maria Montefinese, Ettore Ambrosini, Beth Fairfield, and Nicola Mammarella. 2014. The adaptation of the affective norms for english words (anew) for italian. *Behavior research methods*, 46(3):887–903.

Anthony Mulac, James J. Bradac, and Pamela Gibbons. 2001. Empirical support for the gender-as-culture hypothesis: An intercultural analysis of male/female language differences. *Human Communication Research*, 27(1):121–152.

Nabila Myisha, Dinda Sabila, Angelia Brigita Maharani, Akira Hilal Ramadhan, and Mirza Fathima Jauhar Kamalia. 2023. Decoding the perpetuation of patriarchal culture in the barbie movie. *Cultural Narratives*, 1(2):71–82.

Matthew L. Newman, Carla J. Groom, Lori D. Handelman, and James W. Pennebaker. 2008. Gender differences in language use: An analysis of 14,000 text samples. *Discourse processes*, 45(3):211–236.

Alexandra Schofield and Leo Mehr. 2016. Gender-distinguishing features in film dialogue. In *Proceedings of the Fifth Workshop on Computational Linguistics for Literature*, pages 32–39.

Deborah Tannen. 1990. *You Just Don't Understand: Women and Men in Conversation*. William Morrow, New York.

Peter Trudgill. 1972. Sex, covert prestige and linguistic change in the urban british english of norwich. *Language in society*, 1(2):179–195.

Amy Beth Warriner, Victor Kuperman, and Marc Brysbaert. 2013. Norms of valence, arousal, and dominance for 13,915 english lemmas. *Behavior research methods*, 45(4):1191–1207.