Typical Writing Styles Among Genders: A Corpus Study in Asian L2 Learners' Academic Writing

Sulistyaningrum, The University of Queensland, Australia

The Asian Conference on Arts & Humanities 2018
Official Conference Proceedings

Abstract
Scholars have been attempting to distinguish the writing characteristics between male and female since the last several decades. Surprisingly, although the objects analysed by the earlier studies were diverse, all of them were written in the participants' first languages. Considering the insufficient study on the different linguistics aspects among genders in second-language discourse, this study hence explores the learner corpus of The International Corpus Network of Asian Learners of English (ICNALE) (Ishikawa, 2013), which provides the electronic collection of written essays produced by 2,800 EFL and ESL learners from ten different countries in Asia including Hong Kong, Pakistan, Philippines, Singapore, China, Indonesia, Japan, Korea, Thailand, and Taiwan. Using the 3.3g version of compilation and annotation software UAMCorpustool (O'Donnell, 2008), the data is annotated based on the stylistic features on Rubin and Greene (1992), Koppel et al. (2002), and Mulac and Lundell (1994). This corpus study aimed to investigate the comparative gender-based writing styles in argumentative essays written by ESL learners with B2 CEFR proficiency level to those written by EFL learners with the same English proficiency level. The findings show that (1) academic texts in general and argumentative essays of ICNALE, in particular, are characterized by gender-based writing styles; (2) the writings of L2 learners indicate the use of gender-based linguistic features, (3) both ESL and EFL learners use gender-based writing styles with an identical distribution.

Keywords: corpus, writing style, gender, ICNALE, UAMCorpustool

Introduction

Scholars have been attempting to differentiate the characteristics between males and females regarding language use (e.g. Lakoff, 1973; Bem, 1981). Several studies (Brouwer et al., 1979; Bradley, 1981; Berryman-Fink and Wilcox, 1983), however, claimed that sex role did not reflect the differences in the ways humans use language as means of communications. Females and males are considered to have the same nature in producing discourse. Furthermore, Berryman-Fink and Wilcox (1983) contended that the insufficient amount of information, as well as very few identifiable linguistic features, could not make the categorisation of language styles based on the gender possible. In other words, those studies believed that the preconceptions on the prototypical pattern were still depended on the limited amount of a recorded data.

Meanwhile, other studies believed that the different linguistics aspects between males and females only exist in literary works (Holmes, 1998; Koppel et al., 2002) and oral produced discourse (Mulac and Lundell, 1994) rather than that in academic writing. In terms of gender attribution, studies in spoken language were broadly developed in specific domains such as speech (Schirmer et al., 2005; Leaper and Ayres, 2007), verbal ability (Hyde and Linn, 1988), virtual communication (Furumo and Pearson, 2007), and conversation (Singh, 2001). As in spoken language, the author attribution in literature was broadly explored through the use of consistent choice of lexical, syntactic, and discourse features (Daelemans, 2013). Rubin and Greene (1992, p. 16) insisted that language style is more likely to be found in "reflexive and expressive" writing rather than that in "extensive and instrumental" writing. Thus, numerous studies can be found in this area exploring author's style in literary works (e.g. Culpeper, 2002; Stubbs, 2005; Starcke, 2006; Segundo, 2016). It has commonly been assumed that the language paradigmatic among genders is reduced in instructive texts because students learn and apply the same standard in academic writing (Mulac and Lundell, 1994).

Nevertheless, this sceptical view on the absence of male and female language variation in non-literary texts was challenged by a number of research studies investigating various research objects such as scientific articles (Argamon, et al., 2003; Sarawgi et al., 2011; Koppel et al., 2002), books (Argamon et al., 2003; Koppel et al., 2002), essays (Engelhard et al., 1992; Jones and Myhill, 2007; Mulac and Lundell, 1994; Rubin and Greene, 1992), business letters (Sterkel, 1988), emails (Colley and Todd, 2002), web blogs (Sarawgi et al., 2011), and online messages (Baron 2004; Zheng et al, 2006). Surprisingly, although the objects analysed by earlier studies were diverse, all of them were written in participants' first language. Yet, the investigation on gender-based characteristics in second-language discourse is insufficiently studied.

A statistical method to identify language style called stylometry has been widely used to assist the exploration of male and female unique language features. As a developing interdisciplinary, stylometric method combined the utility of "statistics and computer science" (Ramyaa and Rasheed, 2004) to identify a particular style applied by the author. The stylometric research considers certain linguistics variation that can distinguish one document from another. There are specific parameters or variables used to measure the characteristics of the writing. The categorisation in the

stylometry can be classified based on the content, genre, topic, author, and gender (Holmes, 1998; Ramyaa and Rasheed, 2004; Daelemans, 2013).

As a tool of pattern recognition, stylometry has been vigorously used in various disciplines such as literature, linguistics, forensics, sociolinguistics, psycholinguistics, and even medical diagnosis (Daelemans, 2013). However, before the emergence of computational stylometry, distinguishing the predefined language features was demanding and time-consuming since it was manually coded by humans (Zheng et al., 2005). It was not until the development of corpus linguistics that the limitless exploration in stylometry became possible. As massive huge source of language documentation completed with an automatic and computer-aided technique, a corpus empowers the analysis of the stylistic features of the language.

Corpus linguistics is a rapidly-growing discipline associated with the exploration of a corpus, an electronically searchable collection of spoken and written language, which has revealed many linguistics phenomena in this area (Granger, 2013; Gries, 2009; Hunston and Laviosa, 2006). This combination of specific software instruments and virtually stored documents covering diverse genres has enabled the advanced method of finding quantitative data to many research studies.

In reference to its typology, the corpus can consist of written or spoken texts produced by native speakers of the language as well as non-native speakers or language learners. The latter type is referred to learner corpus, a specific collection that records the discourse produced by language learners with the aim to "improve the learning and teaching of foreign/ second language" (Granger, 2013).

Granger (2013) stated that recent technological development has enabled the academics to compile the "learner data in large quantities, store it on the computer and analyse it automatically or semi-automatically using currently available linguistic software." Although computer corpus methodology has been conducted for the last several decades, the Computer Learner Corpora (CLC) of non-native English was started to develop in the late 1980s.

Having a potential of limitless exploration, digital corpus becomes a prospective source in many study areas, including the study of gender. Nonetheless, it is still insufficiently explored as a limited amount of research has been conducted using this electronic reference. Only few numbers of earlier studies (Koppel et al., 2002; Argamon et al., 2003) utilized digital corpus like the British National Corpus, while others built their own digital corpus, for example, from the data collected from AOL Instant Messenger (Baron, 2004) and LiveJournal blogs (Rosenthal and McKeown, 2011).

Linguists use a set of language features to evaluate the language variation among genders. Argamon et al. (2003), Zheng et al. (2006), and Koppel et al. (2002) suggested that gender-linked language differences can be traced from lexical, syntactic, structural, and content-specific features. Although the features vary, each gender is usually attributed to certain markers.

The dominant contrast of language differences between males and females is shown on the "involvement-informational dimensions", in which female's language indicates

a frequent use of elements classified as "involved", while male's language indicates a frequent use of elements classified as "informational" (Argamon et al., 2003). Compared to male's, female's writing shows the extensive usage of pronouns (Argamon et al., 2003; Colley and Todd, 2002; Koppel et al., 2002) and tag questions (Baron, 2004; Sterkel, 1988) as an intention to get involved in the situation they are discussing or to make an interaction with their readers. The way females use 'expressive language' (Rubin and Greene, 1992) as marked by the frequent use of intensifiers, e.g. 'strongly', 'really', 'very' (Mulac and Lundell, 1994; Sterkel, 1988; Rubin and Greene, 1992), affective markers, e.g. 'excited', 'anxious' (Baron, 2004; Colley and Todd, 2002; Mulac and Lundell, 1994), diminutives, e.g. 'kitty' for a cat, 'veggie' for vegetables (Baron, 2004) also reflects "an impression of heightened arousal, intimacy, and desire to engage the recipient's interest." (Colley and Todd, 2002). Hence, it is revealed that females use subjective approach to maintain the social connection and relationship, which is commonly referred as 'the involvedness' (Argamon et al., 2003; Rubin and Greene, 1992).

On the other side, males tend to use a set of gender-linked attributes such as quantifiers, e.g. 'one', 'some', 'more' (Koppel et al., 2002; Mulac and Lundell, 1994; Sterkel, 1998) and locatives, e.g. 'above', 'inside', 'left' (Mulac and Lundell, 1994) to directly present information or fact in their writing. Rubin and Greene (1992) defined this objective approach as "denotative" indication. Although male writing is more likely to exclude expressive or emotional expression, judgmental adjectives, e.g. 'distracting', 'moody', 'bad-tempered' (Mulac and Lundell, 1994) and profanity, e.g. 'damn' (Baron, 2004) are frequently used as a substitution. Argamon et al. (2003) believed that those distinctive elements in "involvement and informational dimensions" may occur based on how "people, objects, collectives and institutions are presented" by each gender in their writing.

Another noticeable pattern was shown by the preference of both genders to demonstrate directness, in which male's writing exhibits more illative connectives, e.g. 'therefore', 'thus' (Rubin and Greene, 1992), while the opposite gender exhibits the tendency to use hedges ('somewhat', 'probably'), perceptual verbs ('seems', 'looks'), adversative connectives ('but' 'otherwise'), auxiliaries of possibility ('could', 'may'), qualifiers ('nearly', 'kind of'), and conjunctions ('and', 'but', 'if') (Baron, 2004; Koppel et al., 2002; Lakoff, 1973; Mulac and Lundell, 1994; Rubin and Greene, 1992). Such stylistic elements in female's writing boldly underline the feeling of uncertainty and hesitancy as once again they tend to involve the readers' perception in their writing. Reflecting on those different features, this directness tendency is evidently shown although in many cases, female's writing shows more markers than male's.

Considering the way females and males present their writing, certain specific attributes were also drawn by previous studies, in which determiners, e.g. 'a', 'the', 'that' and sentence-initial conjunctions, e.g. 'and another reason is…' (Argamon et al., 2003; Koppel et al., 2002; Mulac and Lundell, 1994) are listed in male's characteristics, while dependent clauses ('which is supported…'), sentence-initial adverbials ('Before the lecture begins, she…'), active voice verbs ('eat', 'write', 'go'), negation ('no', 'nothing', 'none'), and prepositions ('on', 'at', 'by') (Mulac and Lundell, 1994; Koppel et al., 2002) are listed in female apparent characteristics. Besides, some investigation also collected a number of stylistic attributes that are still

questionable whether they belong to female or male characteristics. Thus, those attributes are listed for further to identify in the analysis. They are included additive connectives, e.g. 'and', 'also', adversative connectives, e.g. 'however', 'otherwise', 'yet', causal connectives, e.g. 'since', 'because', de-intensifiers, e.g. 'just', 'not really', egocentric sequences 'I believe', 'I think', progressive verbs, e.g. 'loving', 'reading', justifiers, e.g. 'It is hot because…', illustrators, e.g. 'for example', 'for instance', proximals, e.g. 'about', 'around', refusals, e.g. 'I do not know', 'I am not sure', and temporal connectives, e.g. 'next', 'first' (Mulac and Lundell, 1994, Rubin and Greene, 1992).

To examine the further possibility, other research papers also calculate the number of elements that constructed the text. Jones and Myhill (2007) stated that examination of the variety of paragraphing can show the aspects of "text-level linguistics". Thus, the stylometric analysis also can be drawn from the number of parts of speech, characters, words, sentences, paragraphs, abbreviations, acronyms, and even slang words (Baron, 2004; Jones and Myhill, 2007; Sterkel, 1988). Further, the use of punctuation such as full stops, commas, brackets, and dashes becomes one of the considerations to trace the gender-based writing styles (e.g. Calix et al., 2008; Engelhard et al., 1992; Jones and Myhill, 2007). Besides revealing the diverse dimensions among genders, this type of evaluation usually also measures the quality of the writing (e.g. Engelhard et al., 1992; Jones and Myhill, 2007; Francis et al., 2001).

As an extension of the earlier studies, this present study attempts to address the existing gap by focusing on the distinctive style of females and males in L2 writing. This study seeks whether the gender differences also occur in the discourse produced by L2 learners as in the discourse produced by native speakers. Accordingly, four research questions proposed are (1) Are there any linguistic differences between male's and female's writing in L2 learners' argumentative essays?; (2) How is the frequency of using gender-based writing styles between the Asian ESL and EFL groups?; (3) What are the stylistic features found in female's L2 writing?; and (4) What are the stylistic features found in male's L2 writing?

The learner corpus of ICNALE (The International Corpus Network of Asian Learners of English) will be used as the research object of this study as it provides the electronic collection of written discourse produced by EFL (English as a Foreign Language) and ESL (English as a Second Language) learners.

Conclusion

In the present study, a downloadable discourse of the learner corpus of ICNALE (The International Corpus Network of Asian Learners of English) was used as the research data.

The contributors of the essays in the ICNALE were 2,600 ESL and EFL students from ten different countries in Asia including Hong Kong, Pakistan, Philippines, Singapore, China, Indonesia, Japan, Korea, Thailand, and Taiwan who produced 5,200 essays or two argumentative essays per student. The English proficiency level of the learners ranges from A2 to B2+ based on Common European Framework of Reference (CEFR) level. A total number of 1,127 male students produced 2,254 essays, while 1,473 female students produced 2,946 essays. This study used the most

advanced level, B2, as the research data which consisted of 296 essays produced by female and 168 essays produced by male. The topics of the argumentative essays are:

1. It is important for college students to have a part-time job.
2. Smoking should be completely banned at all the restaurants in the country.

The downloaded data of the written essays were sorted based on the genders. Using the 3.3g version of compilation and annotation software UAMCorpustool (O'Donnell, 2008), the data were annotated based on the stylistic features on Rubin and Greene (1992), Koppel et al. (2002), and Mulac and Lundell (1994). The full list of the features was attached in the appendix. The features found in the annotation were examined one by one to make sure the target words are suitable for its specific function. This process will be followed by the analysis process of the language differences between males and females.

The findings show that there are significant differences in the gender-based writing styles between the writings of males and females in the L2 learners' writings. The writings of males are significantly characterised by the using of quantifiers, followed by locatives and determiners.

Table 1. The use of gender-based writing styles among genders

	Female		**Male**			
Features	**N**	**Percent**	**N**	**Percent**	**ChiSqu**	**Signif.**
Females' Features	13418	49.64%	6015	47.56%	14.951	+++
Males' Features	13613	50.36%	6633	52.44%	14.951	+++
Quantifiers	2590	9.58%	1292	10.22%	3.917	++
Locatives	3104	11.48%	1527	12.07%	2.909	+
Determiners	7919	29.30%	3814	30.15%	3.053	+
Illative Connectives	1035	3.83%	453	3.58%	1.46	
Adversative Connectives	1690	6.25%	734	5.80%	3.026	+
Causal Connectives	404	1.49%	172	1.36%	1.093	
Illustrators	332	1.23%	153	1.21%	0.025	
Additive Connectives	2926	10.82%	1294	10.23%	3.196	+
Temporal Connectives	385	1.42%	169	1.34%	0.486	
Conditional Connectives	411	1.52%	199	1.57%	0.159	
Intensifiers	569	2.10%	214	1.69%	7.598	+++
De Intensifiers	547	2.02%	236	1.87%	1.108	
Proximals	36	0.13%	18	0.14%	0.053	
Modal Adjuncts	108	0.40%	53	0.42%	0.081	
Auxiliaries of Possibilities	1777	6.57%	825	6.52%	0.037	
Perceptual Verbs	20	0.07%	9	0.07%	0.009	

The statistical data shows that female's linguistic features are used more regularly by EFL learners. EFL learners use illative, causal, temporal, conditional connectives and intensifiers more frequently compared to ESL learners. However, ESL learners' use

additive connectives, modal adjuncts, and auxiliaries of possibilities more frequently compared to EFL learners.

Table 2. The use of female's writing styles among ESL and EFL learners

Features	Female EFL Learners		Female ESL Learners		ChiSqu	Signif.
	N	Percent	N	Percent		
Females' Features	2754	50.09%	3261	45.61%	25.04	+++
Illative Connectives	626	8.20%	277	7.26%	3.109	+
Adversative Connectives	987	12.92%	479	12.55%	0.326	
Causal Connectives	247	3.23%	98	2.57%	3.882	++
Illustrators	187	2.45%	103	2.70%	0.64	
Additive Connectives	1569	20.54%	945	24.75%	26.293	+++
Temporal Connectives	283	3.71%	51	1.34%	50.502	+++
Conditional Connectives	240	3.14%	98	2.57%	2.947	+
Intensifiers	383	5.02%	103	2.70%	33.645	+++
De Intensifiers	303	3.97%	166	4.35%	0.938	
Proximals	18	0.24%	6	0.16%	0.751	
Modal Adjuncts	36	0.47%	50	1.31%	24.001	+++
Auxiliaries of Possibilities	896	11.73%	594	15.56%	32.921	+++
Perceptual Verbs	11	0.14%	6	0.16%	0.03	

Among 13 linguistic features listed, the use of intensifiers, adversative connectives, and additives connectives are significantly often. On the other side, the writings of males use all the three male's linguistic features listed.

Table 3. The use of male's writing styles among ESL and EFL learners

Features	Male EFL Learners		Male ESL Learners		ChiSqu	Signif.
	N	Percent	N	Percent		
Males' Features	7259	48.73%	4274	52.82%	35.024	+++
Quantifiers	598	21.79%	694	17.85%	15.987	+++
Locatives	659	24.02%	868	22.32%	2.613	
Determiners	1487	54.19%	2327	59.84%	20.975	+++

To sum up, First, academic texts in general and argumentative essays of ICNALE, in particular, are characterized by gender-based writing styles. Second, the writings of L2 learners indicate the use of gender-based linguistic features as in the writings of L1 learners. Third, both ESL and EFL learners use gender-based writing styles with an identical distribution.

Acknowledgements

I would like to express my sincere gratitude to Indonesia Endowment Fund for Education (Lembaga Pengelola Dana Pendidikan/LPDP) which has sponsored this research. I thank Zein Saeed and Peter Crosthwaite for their assistance and comments that greatly improved the manuscript.

References

Argamon, S., Koppel, M., Fine, J., & Shimoni, A. R. (2003). Gender, genre, and writing style in formal written texts. *Interdisciplinary Journal for the Study of Discourse*, *23*(3), 321-346.

Baron, N. S. (2004). See you online: Gender issues in college student use of instant messaging. *Journal of language and social psychology*, *23*(4), 397-423.

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological review*, *88*(4), 354.

Berryman-Fink. C. L. & Wincox, J. R. (1983) A multivariate investigation of perceptual attributions concerning gender appropriateness in Language. *Sex Roles,* 9, 663-681.

Brower, O. (1982). The influence of the addressee's sex on politeness in language use. *Linguistics*, *20*, 697-711.

Calix, K., Connors, M., Levy, D., Manzar, H., McCabe, G., & Westcott, S. (2008). Stylometry for e-mail author identification and authentication. *Proceedings of CSIS Research Day, Pace University*, 1048-1054. http://csis.pace.edu/ctappert/srd2008/c2.pdf

Colley, A., & Todd, Z. (2002). Gender-linked differences in the style and content of e-mails to friends. *Journal of Language and Social Psychology*, *21*(4), 380-392.

Culpeper, J. (2002). Computers, language and characterisation: an analysis of six characters in Romeo and Juliet. *Uppsala Universitetstryckeriet, 15,* 11-30.

Daelemans, W. (2013, March). Explanation in computational stylometry. In *International Conference on Intelligent Text Processing and Computational Linguistics* (pp. 451-462). Springer, Berlin, Heidelberg.

Engelhard Jr, G., Gordon, B., & Gabrielson, S. (1992). The influences of mode of discourse, experiential demand, and gender on the quality of student writing. *Research in the Teaching of English*, *26(*3), 15-336.

Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: countering criticisms against corpus-based methodologies. *English for specific purposes*, *24*(3), 321-332.

Francis, B., Robson, J., & Read, B. (2001). An analysis of undergraduate writing styles in the context of gender and achievement. *Studies in Higher Education*, 26(3), 313-326.

Furumo, K., & Pearson, J. M. (2007). Gender-based communication styles, trust, and satisfaction in virtual teams. *Journal of Information, Information Technology & Organizations*, *2,* 49-60.

Granger, S. (2013). A bird's-eye view of learner corpus research. *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, *6*, 3-33.

Gries, S. T. (2009). What is corpus linguistics?. *Language and Linguistics Compass*, *3*(5), 1225-1241.

Holmes, D. I. (1998). The evolution of stylometry in humanities scholarship. *Literary and linguistic computing*, *13*(3), 111-117.

Hunston, S., & Laviosa, S. (2006). Corpus linguistics. *Linguistics*, *7*(2), 215-244.

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104*(1), 53-69.

Jones, S., & Myhill, D. (2007). Discourses of difference? Examining gender differences in linguistic characteristics of writing. *Canadian Journal of Education*, *30*(2), 456-482.

Koppel, M., Argamon, S., & Shimoni, A. R. (2002). Automatically categorizing written texts by author gender. *Literary and Linguistic Computing*, *17*(4), 401-412.

Lakoff, R. (1973). Language and woman's place. *Language in society*, *2*(1), 45-79.

Leaper, C., & Ayres, M. M. (2007). A meta-analytic review of gender variations in adults' language use: Talkativeness, affiliative speech, and assertive speech. *Personality and Social Psychology Review*, *11*(4), 328-363.

Mulac, A., & Lundell, T. L. (1994). Effects of gender-linked language differences in adults' written discourse: Multivariate tests of language effects. *Language & Communication*, *14*(3), 299-309.

O'Donnell, M. (2008). The UAM CorpusTool: Software for corpus annotation and exploration. In *Proceedings of the XXVI Congreso de AESLA, Almeria, Spain* (pp. 3-5).

Ramyaa, C. H., & Rasheed, K. (2004). Using machine learning techniques for stylometry. In *Proceedings of International Conference on Machine Learning*.

Rosenthal, S., & McKeown, K. (2011, June). Age prediction in blogs: A study of style, content, and online behavior in pre-and post-social media generations. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1 (pp. 763-772). Association for Computational Linguistics.

Rubin, D. L., & Greene, K. (1992). Gender-typical style in written language. *Research in the Teaching of English*, *26*(1), 7-40.

Sarawgi, R., Gajulapalli, K., & Choi, Y. (2011, June). Gender attribution: tracing stylometric evidence beyond topic and genre. In *Proceedings of the Fifteenth*

Conference on Computational Natural Language Learning (pp. 78-86). Association for Computational Linguistics.

Schirmer, A., Kotz, S. A., and Friederici, A. D. (2005). On the role of attention for the processing of emotions in speech: Sex differences revisited. *Cognitive Brain Research*, *24*(3), 442-452.

Segundo, P. (2016). A corpus-stylistic approach to Dickens' use of speech verbs: Beyond mere reporting. *Language and Literature*, *25*(2), 113-129.

Singh, S. (2001). A pilot study on gender differences in conversational speech on lexical richness measures. *Literary and Linguistic Computing*, *16*(3), 251-264.

Starcke, B. (2006). The phraseology of Jane Austen's *Persuasion*: Phraseological units as carriers of meaning. *ICAME Journal*, *30*, 87-104.

Sterkel, K. S. (1988). The relationship between gender and writing style in business communications. *The Journal of Business Communication*, *25*(4), 17-38.

Stubbs, M. (2005). Conrad in the computer: examples of quantitative stylistic methods. *Language and Literature, 14*(1), 5-24.

Zheng, R., Li, J., Chen, H., & Huang, Z. (2006). A framework for authorship identification of online messages: Writing‐style features and classification techniques. *Journal of the Association for Information Science and Technology*, *57*(3), 378-393.

Appendix – Coded Stylistic Features

Female's Linguistic Features

Features	Key Examples
Illative connectives	Therefore, so, consequently, as a result, as a consequence, hence, thus, accordingly, then
Adversative connectives	However, but, yet, otherwise, nevertheless, nonetheless, still, though, although, even so, despite that, in spite of that, anyway, anyhow, notwithstanding
Causal Connectives	Because, since, in order to
Illustrators	For example, for instance, as an illustration, such as, to illustrate, namely, like
Additive Connectives	And, also, with, together with, along with, as well as, in addition, including, too, besides, furthermore, moreover, plus
Temporal Connectives	Next, after, lastly, first, afterwards, subsequently, thereafter, thereupon, then
Conditional Connectives	If, as long as
Intensifiers	A lot, quite, really, very, extremely, at all, ever, too, so
De-Intensifiers	Just, only, not really, rather, approximately, roughly
Proximals	About, around, nearly, roundabout, thereabouts, more or less, close to, almost
Modal Adjuncts	Maybe, hopefully, probably, possibly, perhaps, conceivably, feasibly, likely
Auxiliaries of possibility	Could, may, would, should
Perceptual verbs	Looks, seems, sounds, feels

(Rubin and Greene, 1992)

Male's Linguistic Features

Features	Key Examples
Quantifiers	Some, many, plenty, heaps, load, loads, tons, both, each, either, few, neither, several, couple, hundred, hundreds, thousand, thousands, million, millions, billion, billions, a bit, all, a lot of, a number of, a plethora of, enough, sufficient, no lack of, lots of, quantities of, a good deal of, a great deal of, adequate, as much as, ample, abundant, quantity of, numbers of, a little bit
Locatives	Above, inside, in, at, on, near, there, here, below, indoor, outdoor, within, centre, middle, corner, front, around, center
Determiners	A, the, that, an, any, other, another

(Koppel et al., 2002; Mulac and Lundell, 1994)