Gender-preferential Linguistic Elements in Applied Linguistics Research Papers: Partial Evaluation of a Model of Gendered Language

*Reza Ghafar Samar**
Mohsen Shirazizadeh
Tarbiat Modares University

This article intended to investigate whether the gender-preferential linguistic elements found by Argomon, Koppel, Fine and Shimoni (2003) show the same gender-linked frequencies in applied linguistics research papers written by non-native speakers of English. In so doing, a sample of 32 articles from different journals was collected and the proportion of the targeted features to the whole number of words was calculated. The results indicated that, although the gender-linked patterns of use for many of those features were also observed in our sample, the difference between men and women in the frequency of using those elements was not statistically significant. This non-significant difference shows that either the confinements of genre or those of using a second language or both are keeping L2 writers from expressing their gender to its fullest capacity in the texts they produce.
Keywords: gendered language, genre, second language, applied linguistics

Investigation of the possible differences between males and females in their use of language has been a topic of interest in sociolinguistics for a long time. Many researchers have thus far studied these possible gender-based differences though the justification has been quite elusive (e.g., Trudgill 1972; Lakoff 1975; Labov 1990; Coates 1998). Within this line of inquiry, many have focused their attention on phonological or pragmatic aspects

---

* Corresponding author. E-mail: rgsamar@modares.ac.ir

of speech among first language speakers (Key 1975; Holmes 1990; Labov 1990; Eckert 1997). Recently, however, some attempts have been made to come up with a pattern or algorithm of male/female-preferential linguistic features in written texts. Such a pattern is intended to help one guess the gender of anonymous texts" authors. Argomon, Koppel, Fine and Shimoni(2003), for instance, offered a pattern of male/female linguistic preferences according to which the gender of a text"s author can be recognized with 80 percent of accuracy. The aim of this study is, nevertheless, to focus on the role of gender in the academic texts produced by non-native speakers of English, partially evaluating the validity of Argomon et al."s model of gendered texts in an L2 academic context.

Gender and language variation

The link between sex and language use variation has attracted so much interest in the past decades. The large number of studies conducted on this topic may convince one that all possible sources of language use variation are, to a lesser or greater extent, influenced by gender (Crawford, 1995). One of the oldest areas of gender-linked variation is that of phonology. Numerous differences have been recognized in different languages between men and women with respect to pronunciation. All these studies appear to directly or indirectly refer to the issue that women, irrespective of their class and age, use more standard pronunciation than their male counterparts (Labov, 1966; Trudgill, 1972; Milroy, 1980).

There are also differences between the two sexes in their use of words. Lakoff (1973), a most distinguished figure in this area, maintains that women use more color descriptors (e.g. lavender), more empty adjectives (e.g. lovely), and weaker insult words (e.g. oh fudge). At the grammatical level, Lakoff (1973) claimed that women do use more hedges and tag questions than men do. Tag questions were regarded by lakoff as sign of uncertainty; a claim which was later seriously disputed by many other scholars (Holmes, 1984; Weatherall, 2002). Women were also found to have a more rising intonation which was considered another

evidence for their uncertain use of language. They also used more indirect requests and standard grammars than men.

Since the Lakoff's (1973; 1975) groundbreaking publications, an avalanche of empirical studies have focused on the issue (Weatherall, 2002). As Crawford (1995) maintained, from 1976 to 1992, more than twenty four references on average have been made to her work per year in scholarly journals. The focus of the studies since then has been much varied, including a large gamut of issues. Although the inception of interest in gender-linked language variation was rather limited to oral language, a noticeable number of studies have recently shown more interest in written language as well. Rubin and Greene (1992), for instance, found that politeness markers are more typical of males than females. Females have also been found to use more linguistic markers of excitement (Colley & Todd, 2002; Rubin &Greene, 1992), hedges (Mulac, Studley, Blau , 1990), and shorter sentences (Mulac et al., 1990; Rubin & Greene, 1992) than male writers.

Janssen and Murachver (2004a) in an investigation of the role of gender in New Zealand literature also found that female authors use more compliments, adjectives, questions, third person pronouns, adverbs, adjectival metaphor and simile, elliptical sentences, and intensifiers than their male counterparts. Male authors, on the other hand, were found to use more first person pronouns, geographical references, and focus adverbs than female authors. Janssen and Murachver (2004b) also found that the use of gender-preferential elements may be related to the topic of writing. For instance, they found that texts with socio-emotional description topics contain more female-preferential elements while political topics contained more male-preferential devices.

Computer mediated communication and especially its commonest form, email communication, has also attracted so much attention by researchers studying language and gender. Electronic mails, though written, share many features of the oral communication. As Barron (1998) maintains, the underlying social dynamics of emails are those of writing while their lexical and stylistic properties are more similar to speech. Such a dual nature

will thus make emails much interesting subjects for inquiry. Thomson and Murachver (2001), investigating the role of gender in emails, realized that it is possible for both the researchers and even the receiver of the email to guess the gender of the sender with a high level of accuracy. They found that emails written by female writers included more modals, intensifiers, questions, compliments, apologies and also more personal information than those written by males. Males, however, gave more personal opinion and used more insults.

The language of weblogs has also been investigated in relation to the influence of the authors' gender on it. Huffaker and Calvert (2005), for example, found that the language of boy bloggers were more active, inflexible, and resolute than that of girls. Kennedy, Robinson and Trammell (2005) also analyzed the comments posted by males and females. They found that the comments posted by women were more expressive and inclusive while those posted by men were more instrumental, assertive and competitive.

Argomon et al. (2003) in their inclusive machine learning research set out to explore differences between female and male writing in a large sample taken from the British National Corpus. They found many different lexical and functional words which were used with substantially different frequencies according to the gender of the author. For instance, in nonfiction texts, these researchers found first person singular pronouns, second person pronouns and third person female pronouns to be used with significantly higher frequency by females. Third person plural pronouns and the possessive "its" were also examples of male-preferential words. In fiction texts, first person plural pronouns, third person male pronouns, possessive "its" and third person plurals were examples of male-preferential features while second person pronouns and third person female pronouns acted as instances of female-preferential linguistic features. This study therefore resulted in the construction of two statistical models for gender categorization in fiction and nonfiction. The models were purported to predict the gender of the author with around 80 percent of accuracy. It was then used as the underlying structure

for „Gender Genie", an online gender recognition program for written texts.

Employing Argomon et al."s model, Herring and Paolillo (2006) set out to examine if part of the features of their model can hold true in weblogs considering two different weblog subgenres (i.e., diary and filter). The results indicated that the filter entries included more male features while diaries included more female features irrespective of the gender of the author. These findings highlight the fact that it might be the influence of (sub)genre and not necessarily the effect of the author"s biological gender which determines the frequency of some so-called gender linked linguistic devices in a certain text. This complexity led Herring and Paolillo (2006) to call for "a more fine-grained genre analysis of apparently gendered language use in other communicative contexts [which] might reveal genre to be a conditioning factor" (p. 455).

A much important point to note here is that many studies have thus far found several discoursal and lexico-grammatical features of texts to be *genre-preferential* (e.g. Samarj, 2005; Hyland, 2008). In fact, the genre in which the author writes may act as a determining factor in choosing some linguistic features over the others (Swales, 1990). The native language of the authors has also been an important factor in choosing several linguistic elements of the text (Scott and Tribble, 2006; Hyland, 2008). In this study, we have, therefore, included texts from one single genre and written only by non-native speakers of English. This may help us come up with a clearer picture of what the effects of the authors" gender would be. In other words, such a research design would enable us to partly examine if the effect of the writers" gender is strong enough to overpower those of genre and native language of the author. In particular, this paper intends to answer with some details the following broad research question:

Is there any difference between the language of male and female non-native authors in a corpus of applied linguistics articles?

Method

Corpus

In order to investigate the validity of Argomon et al. (2003) model of gendered texts in a different and more specific communicative context, a sample of articles from different language related journals were collated. The journals included "the Asian ESP Journal", "The Asian EFL Journal", "TESL-EJ", "Language Learning and Technology", "Second Language Research", "Discourse and Society", "RELC" and "First Language". Half of the journals were open access while the rest were not. The authors were emailed and asked about their native langue. Four articles from each of the journals (2 by men and two by women) were finally selected making a total of 32 single-authored journals written by non-native speakers of English.

Gender-preferential features

The gender-preferential features studied in this article are a subset of those linguistic elements found by Argomon et al. (2003) as gender linked. These researchers, in their computational study, identified several classes of syntactic and lexical linguistic elements which vary in texts according to the author's gender. In particular, significant differences were found between male and female writers in their use of certain pronouns and noun modifiers. The specific male-preferential features focused on in this study are "that", "these", "one", "two", "more", "some", "its", "he" and "him". The female-preferential features selected and studied here are "I", "you", "She", "her", "their", "myself", "yourself" and "herself". These linguistic features stood as the dependent variables in this study with gender being the independent variable.

Procedure and data analysis

The number of the targeted gendered features and the total number of words were counted in each of the articles. The proportion of each of the gendered features to the total number of words was then considered as the dependent variables of this

study. The descriptive statistics of the dependent variables were then calculated to come up with a rough picture of how the supposedly gendered linguistic features varied in applied linguistics research papers produced by non-native writers. A number of independent t-tests were then run to see if there is a significant difference between male and female writers in their use of the targeted linguistic features.

Results

Table 1 shows the descriptive statistics for the use of linguistic features by both male and female writers. As shown in the table, there were some differences between the means of the two groups in using any of the features. However, a quick comparison of the means in the use of Argomon et al."s male-preferential features (i.e., A, the, that, these, one, two, more, some, its, he & him) shows that all these features except for the word "more" were either used equally or more frequently in the applied linguistics research papers produced by non-native male writers. The supposedly female-preferential features (i.e., I, you, she, her, their, myself & yourself), excluding the pronoun "I", were also used equally or more frequently by female writers in this study (see Figure 1).

In order to examine the data further and see if the observed differences between males and females in the use of these linguistic forms are statistically significant, a number of independent samples T-tests were used. The main purpose of this part of the data analysis was, in other words, to see if the selected male/female-preferential linguistic elements can also stand as male/female-preferential linguistic forms in applied linguistics research papers written by non-native writers.

The result of the T-tests indicated that there were no significant differences between males and females in their frequency of the use of "A" (t= .45; p> .05), "the" (t= .46; p> .05), "that" (t=.129; p> .05), "these" (t=.241; p> .05), "one" (t=.31; p> .05), "two" (t= 1.234; p> .05), "more" (t= -.77; p> .05), "some" (t=.77; p>.05), "its" (t=1.236; p>.05), "him" (t= 1.633; p>.05), "I" (t=1.026; p>.05), "you" (t=-.512; p>.05), "she" (t= -1.496; p>.05),

Table 1.

Descriptive Statistics of Gender-preferential Elements

Linguistic features	Gender	N	Mean	Std. Deviation
A	Male	16	.0173	.00371
	Female	16	.0167	.00359
The	Male	16	.0580	.01564
	Female	16	.0556	.01176
That	Male	16	.0099	.00324
	Female	16	.0097	.00364
These	Male	16	.0015	.00062
	Female	16	.0015	.00079
One	Male	16	.0020	.00066
	Female	16	.0019	.00069
Two	Male	16	.0020	.00115
	Female	16	.0016	.00083
More	Male	16	.0024	.00103
	Female	16	.0028	.00170
Some	Male	16	.0014	.00081
	Female	16	.0014	.00058
Its	Male	16	.0008	.00045
	Female	16	.0006	.00072
He	Male	16	.0012	.00119
	Female	16	.0003	.00028
Him	Male	16	.0001	.00018
	Female	16	.0000	.00006
I	Male	16	.0041	.00502
	Female	16	.0024	.00366
You	Male	16	.0006	.00101
	Female	16	.0008	.00094
She	Male	16	.0003	.00062
	Female	16	.0008	.00094
Her	Male	16	.0002	.00022
	Female	16	.0008	.00178
Their	Male	16	.0041	.00197
	Female	16	.0065	.00366
Myself	Male	16	.0000	.00005
	Female	16	.0000	.00010
Yourself	Male	16	.0000	.00000
	Female	16	.0000	.00000

"her" (t=-1.236; p>.05), and "myself" (t=-1.279; p>.05). The reflexive pronoun "yourself" was in fact nonexistent in our sample and thus excluded from the analysis.

Of all the selected linguistic features to be studied in this paper, only "he" and "their" turned out to be used significantly differently by the authors of the two sexes. The result of the t-tests indicated that the pronoun "he" was used more frequently by male writers in our study (t= 2.833; p<.01) whereas the possessive "their" was used more frequently be female writers (t=-2.24; p<.05).

In sum, the result of this study indicated that a considerable number of the linguistic variables recognized by Argomon et al. (2003) as being gender- linked in a large sample from the British National Corpus do not show the same quality in a limited sample of applied linguistic research papers written by non-native writers. It should however be emphasized that the patterns for the use of these linguistic elements in our study (as shown in the descriptive statistics) were to a large extent similar to the ones found by Argomon et al. Most of these patterns of gender-linked linguistic differences were however not *statistically* significant in our study.
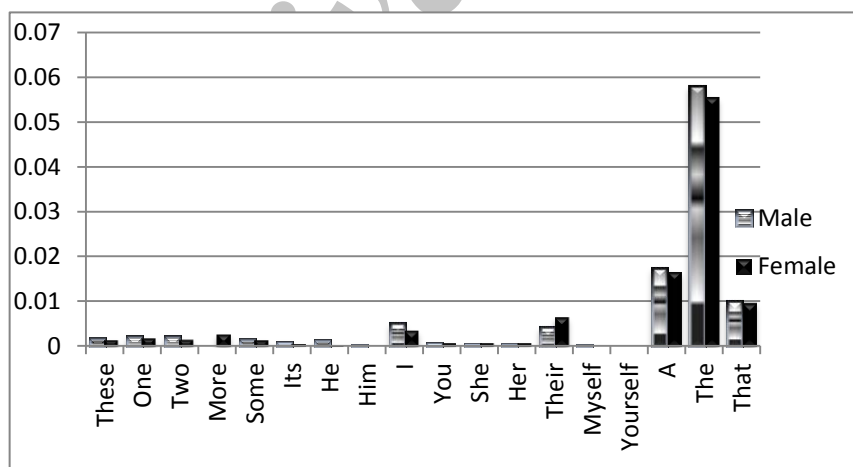

Figure1. Mean of the proportion of linguistic features for males and females

Discussion

The results of this study indicated that many of the linguistic features which were previously found by Argomon et al. (2003) to be gender- linked in a large sample of the British National Corpus were not significantly gender-preferential in a sample of applied linguistics research papers written by non-native speakers. In fact, out of the 18 linguistic features studied here, only two (i.e., he & their) turned out to be significantly gender- linked. In other words, we can claim that the results of this study partially invalidated Argomon et al. (2003) model of gendered language.

One evident explanation for this discrepancy between our results and those of Argomon et al. may be due to the genre specificity of our sample. While the Argomon et al."s sample was a large sample of texts from the BNC, ours was limited to some research papers in applied linguistics and written by non-native speakers of English. This issue of the influence of genre on the style of writing and the frequency of the use of some special linguistic forms has previously been attested to by a number of researchers (Swales, 1990; Hyland, 2003).

Not surprisingly, Argamon et al. (2003) acknowledged that there is "a strong correlation between the characteristics of male (female) writing and those of non-fiction (fiction)". This very claim made by the original proposers of the gender-linked linguistic features makes it much clear that the role of gender may be more or less mediated by some other variables such as the genre of the texts. Quite similarly, Herring and Paolillo (2006) found that the genre of diary weblogs includes more female linguistic features while filter type of weblogs included more male stylistic features independent of the author"s gender. The further insight obtained from the results of the present study might be that not only are the macro-genre (e.g., non-fiction) of texts important elements in determining how frequently a set of features are used in them, but also the micro-genres (e.g., applied linguistics research papers) may exert significant influence thereupon. In fact, one can conclude from the findings of this study that the effect of text genre can, at least in our corpus, overpower the effect of the author"s gender.

Another explanation for the obtained result can be found in the fact that all the papers studied here were written by second language writers. It can be hypothesized from the results that the effect of gender is possibly neutralized in using a second language. This neutralization of the effect of social factors in using a second language was also observed by Katal and Evazzadeh (2010) who investigated the role of social class in the use of compliments by second language learners. They maintained that both working class and middle class learners used the same formulas in expressing their compliments.

One point which should not, however, be neglected in discussing the results is that Argamon et al.'s male and female-preferential linguistic elements stayed, to a large extent, male and female-preferential in our study as well. The noticeable point was that this gendered preference was not statistically significant. Part of this non-significant preference could be due to the limited number of articles we selected as our sample. Another probable explanation may be that the gender of language users are more evident or more strongly expressed when using their first language rather than their second language. As a result, we may observe a descent in the power of gender from first language use to second language use. The same explanation may also be applied to the role genre may play in modifying the influence of gender on language. In fact, it might be claimed that the genre or sub-genre in which the author is wring in act as a filter for the writers in expressing their gender. Some genres, thus, will provide more room for the expression of femaleness while others may act vice versa.

As found, the influence of gender on the texts produced in applied linguistics journals are still there. Yet, this influence is much greater on the non-fiction texts produced by native speakers as reported by Argamon et al. (2003). It is therefore concluded that either genre or being an L2 speaker or even both act as filters for authors in expressing their gendered identity. In other words, in our sample the effect of gender seems to be overshadowed by the effect of genre or the native language of the writer.

One final point which should be made regarding the results of the study is about the two linguistic elements which were found to be significantly different in their frequency of use by females and males. The first feature, the pronoun "he", a supposedly male-preferential element, turned out be used significantly more by males in our study as well. This preference on the part of men can be explained by the fact that they may use "he" more frequently when referring to categorical third person singular, a tendency which is most probably not high among female writers. This more frequent use of this pronoun may also be due to men"s higher tendency to refer to the literature written by male researchers. The possessive "their" was the only feature used significantly more by female writers. This may also be due to women"s tendency to use "their" as a categorical pronoun when referring to third person singular. In other word, while males may prefer to use "he", women may prefer to use "she" or "their" when referring to categorical third person singular. It is these preferences which may have shaped the texts, hence our data.

This partial invalidation of Argamon et al."s model of gendered language, which may also be called partial validation from a different viewpoint, can raise many questions over the issue of gendered linguistic variation in view of genre and second language use patterns.

Conclusion

In this study, we investigated the validity of Argamon et al. (2003) proposed pattern for gendered language. The findings indicated that the role of gender is diminished in applied linguistics research articles written by non-native speakers of English. This diminishing of influence may be due to the genre of the texts or to the non-nativeness of the writers. All in all, it can be hypothesized from the results of this study that applied linguistics research papers are to some extent gender neutral. Put simply, we may claim that the genre of applied linguistic research does not provide enough room for writers to express and manifest their gender. This may be the case with many other genres. Yet, many more studies are required to shed light over this hypothetic claim.

Acknowledgement

The Authors

Reza Ghafar Samar, assistant professor in TEFL, is the head of the ELT department at Tarbiat Modares University, Tehran, Iran. He is the director of TELL journal and has a number of publications in national and international journals. He is currently teaching "sociolinguistics", "second language acquisition" and "discourse analysis" at MA and PhD level.

Mohsen Shirazizadeh is a PhD candidate in English Language Teaching at Tarbiat Modares University, Tehran, Iran. He has published a number of articles in international and national journals such as "Journal of English as an International Language", "The Modern Journal of Applied Linguistics", and "The Journal of Comparative Language and Literature Research". His fields of interests include "second language writing", "language assessment" and "English as a global Lingua franca".

References

Argamon, S., Koppel, M., Fine, J., & Shimoni, A. (2003). Gender, genre, and writing style in formal written texts. *Text, 23*, 321–346.

Baron, N. S. (1998). Letters by phone or speech by other means: the linguistics of email. *Language and Communication*, *18*, 133-170.

Coates, J. (1998). *Language and gender: A reader* . Oxford: Blackwell.

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

Crawford, M. (1995). *Talking difference*. London: Sage.

Eckert, P. (1997). Gender and sociolinguistic variation. In J. Coates (Ed.), *Readings in Language and Gender (pp*. 64–75). Oxford: Blackwell.

Herring, S. C., & Paolillo, J. C. (2006). Gender and genre variation in weblogs. *Journal of Sociolinguistics, 10*, 439-459

Holmes, J. (1984). Hedging your bets and sitting on the fence: some evidence for hedges as support structures. *Te Reo, 27*, 47-62

Holmes, J. (1990). Hedges and boosters in women̈s and men̈s speech. *Language and Communication, 10,* 185-205

Huffaker, D. A., & Calvert, S. L. (2005). Gender, identity, and language use in teenage blogs. *Journal of Computer-Mediated Communication,10*, article 1.

Hyland, K. (2003). Dissertation acknowledgments: The anatomy of a Cinderella genre. *Written Communication*, *20*, 242-268.

Hyland, K. (2008). Academic clusters: text patterning in published and postgraduate writing. *International Journal of Applied Linguistics, 18*, 41-62.

Janssen, A., & Murachver, T. (2004a). The role of gender in New Zealand literature: Comparisons across periods and styles of writing. *Journal of Language and Social Psychology*, *23*, 180-203.

Janssen, A., & Murachver, T. (2004b). The relationship between gender and topic in gender-preferential language use. *Written Communication* , *21*, 344-367.

Katal, A., & Evazzadeh, S. (2010, October). *We Are All Upper Class in L2: Complimenting in L1 vs. L2*. Paper presented at 8th international TELLSI conference, Tehran, Iran.

Kennedy, A., Tracy L. M., Robinson J. S. & Trammell. K. (2005, October). *Does gender matter? Examining conversations in the blogosphere*. Paper presented at Internet Research 6.0: Internet Generations, Chicago, Illinois.

Key, M. R. (1975). *Male/Female Language*. Metuchen, NJ: Scarecrow Press.

Labov, W. (1966). Hypercorrection by the lower middle class as a factor in sound change . In W. Bright (Ed.), *Sociolinguistics* , (pp. 88- 101). The Hague: Mouton.

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

Lakoff, R. T. (1975). *Language and Women's Place*. New York: Harper Colophon Books.

Milroy, L. (1980). *Language and social networks*. Oxford: Blackwell.

Mulac, A., Studley, L. B., & Blau, S. (1990).The gender-linked language effect in primary and secondary students" impromptu essays. *Sex Roles*, *23*, 439-469.

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

Scott, M. & Tribble, C. (2006). *Textual Patterns*. Amsterdam: Benjamins.

Swales, J. (1990). *Genre Analysis*. Cambridge: Cambridge University Press.

Trudgill, P. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich, *Language in Society* ,*1*, 179–195.

Thomson, R., & Murachver, T. (2001). Predicting gender from electronic discourse. *British Journal of Social Psychology*, *40*, 193-208.

Weatherall, A. (2002). *Gender, language and discourse*. London: Routledge.