A LINGUISTIC AND NON-LINGUISTIC ANALYSIS OF
GENDER DIFFERENCE IN WRITING STYLE IN
preadolescent BLOGS

ANGELINE HII HUI WEN

This project is submitted in partial fulfilment of the requirements for a
Bachelor of Education with Honours
(Teaching English as a Second Language)

Faculty of Cognitive Sciences and Human Development
UNIVERSITI MALAYSIA SARAWAK
2009
ACKNOWLEDGEMENTS

First of all, I would like to express my heartfelt appreciation to my supervisor, Madam Rosnah Mustafa for guiding me through this challenging research. Your guidance, advice and support were invaluable in my work and had made this seemingly impossible project possible.

Thanks to my family members for all their love and support: my beloved grandmother who offered unconditional love and encouragement to pursue my studies; my dear parents, for providing untiring support; my little sister, for putting up with me and adding cheers when times were tough.

Last but not least, I would like to thank all my friends and coursemates who have been helpful and supportive throughout my four years of study in UNIMAS. Thanks for all the beautiful memories, prayers, and finally, not forgetting my Father in heaven, who is always there to guide me through the ups and downs of life.
# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS**  
**LIST OF TABLES**  
**LIST OF FIGURES**  
**LIST OF ABBREVIATIONS**  
**ABSTRACT**  
**ABSTRAK**

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.0</td>
<td>Chapter overview</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>Introduction</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Background of study</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Research problem</td>
<td>5</td>
</tr>
<tr>
<td>1.4</td>
<td>Aims and objectives of study</td>
<td>6</td>
</tr>
<tr>
<td>1.5</td>
<td>Significance of study</td>
<td>8</td>
</tr>
<tr>
<td>1.6</td>
<td>Operational definition of terms</td>
<td>9</td>
</tr>
<tr>
<td>1.6.1</td>
<td>Blogs</td>
<td>9</td>
</tr>
<tr>
<td>1.6.2</td>
<td>Writing style</td>
<td>9</td>
</tr>
<tr>
<td>1.6.3</td>
<td>Gender</td>
<td>10</td>
</tr>
<tr>
<td>1.6.4</td>
<td>Adolescent bloggers</td>
<td>10</td>
</tr>
<tr>
<td>1.6.5</td>
<td>Linguistic analysis</td>
<td>10</td>
</tr>
<tr>
<td>1.6.6</td>
<td>Non-linguistic analysis</td>
<td>11</td>
</tr>
<tr>
<td>1.7</td>
<td>Scope of study</td>
<td>12</td>
</tr>
<tr>
<td>1.8</td>
<td>Chapter review</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>LITERATURE REVIEW</td>
<td>14</td>
</tr>
<tr>
<td>2.0</td>
<td>Preview</td>
<td>14</td>
</tr>
<tr>
<td>2.1</td>
<td>Gender and language use</td>
<td>15</td>
</tr>
<tr>
<td>2.1.1</td>
<td>Gender difference in spoken texts</td>
<td>16</td>
</tr>
<tr>
<td>2.1.2</td>
<td>Gender differences in written texts</td>
<td>16</td>
</tr>
<tr>
<td>2.2</td>
<td>Gender and online language</td>
<td>19</td>
</tr>
<tr>
<td>2.3</td>
<td>Computer-mediated discourse</td>
<td>21</td>
</tr>
<tr>
<td>2.4</td>
<td>The nature of <em>Netspeak</em> as a new linguistic medium</td>
<td>22</td>
</tr>
<tr>
<td>2.5</td>
<td>Blogs</td>
<td>23</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Personal blogs</td>
<td>24</td>
</tr>
<tr>
<td>2.6</td>
<td>The Gender Genie as a predictor of author gender</td>
<td>25</td>
</tr>
<tr>
<td>2.7</td>
<td>Gender differences in e-mail communication</td>
<td>28</td>
</tr>
<tr>
<td>2.8</td>
<td>Gender and writing style in fiction and non-fiction texts</td>
<td>29</td>
</tr>
<tr>
<td>2.9</td>
<td>Gender and emoticons</td>
<td>31</td>
</tr>
<tr>
<td>2.10</td>
<td>Summary</td>
<td>33</td>
</tr>
</tbody>
</table>
3 METHODOLOGY
3.0 Preview
3.1 Research design
3.2 Population and sampling
3.3 Instruments for data collection
3.3.1 Linguistic features
3.3.2 Non-linguistic features
3.4 Data collection procedures
3.5 Data analysis
3.6 Limitations of the study
3.7 Chapter review

4 FINDINGS AND DISCUSSION
4.0 Preview
4.1 Analysis of blogs
4.2 Findings
  4.2.1 What are the linguistic features found in adolescent male- and female authored blogs and to what extent are these linguistic features similar to those in the checklist adapted?
  4.2.2 What are the non-linguistic / paralinguistic features found in adolescent male- and female-authored blogs and to what extent are these non-linguistic / paralinguistic features similar to those in the checklist adapted?
  4.2.3 How does writing vary linguistically between adolescent male- and female-authored blogs?
  4.2.4 How does writing vary non-linguistically / paralinguistically between adolescent male- and female-authored blogs?
4.3 Discussion
4.4 Summary

5 SUMMARY, RECOMMENDATIONS AND CONCLUSION
5.0 Chapter Overview
5.1 Summary of study
5.2 Implications of the Findings
5.3 Recommendations for Future Research
5.4 Conclusions
REFERENCES  

APPENDICES  

A Checklist to analyse the occurrence of linguistic and non-linguistic features in blog writing 113 
B Definition and examples of linguistic and non-linguistic features 114 
C Samples writings of male blog writers 115 
D Samples writings of female blog writers 127
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Linguistic features found in blog writing</td>
<td>43</td>
</tr>
<tr>
<td>2</td>
<td>Non-linguistic features found in blog writing</td>
<td>48</td>
</tr>
<tr>
<td>3</td>
<td>Average length per post by male and female blog writers</td>
<td>59</td>
</tr>
<tr>
<td>4(a)</td>
<td>Occurrences of linguistic features in adolescent male blog writers</td>
<td>64</td>
</tr>
<tr>
<td>(b)</td>
<td>Occurrences of linguistic features in adolescent female blog writers</td>
<td>65</td>
</tr>
<tr>
<td>5</td>
<td>Linguistic features employed in male and female blog writing</td>
<td>70</td>
</tr>
<tr>
<td>6</td>
<td>Non-linguistic features as found prevalent in the blogs</td>
<td>73</td>
</tr>
<tr>
<td>7</td>
<td>The linguistic pattern employed by adolescent male and female blog writers in writing</td>
<td>79</td>
</tr>
<tr>
<td>8(a)</td>
<td>Occurrences of non-linguistic features in male blog writers</td>
<td>91</td>
</tr>
<tr>
<td>(b)</td>
<td>Occurrences of non-linguistic features in female blog writers</td>
<td>91</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>Criteria for analysis by The Gender Genie</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>Random Number Generator</td>
<td>39</td>
</tr>
</tbody>
</table>
ABSTRACT

A LINGUISTIC AND NON-LINGUISTIC ANALYSIS OF GENDER DIFFERENCE IN WRITING STYLE IN ADOLESCENT BLOGS

Angeline Hii Hui Wen

This study examines gender variation on writing style employed by adolescent male and female blog writers (13 – 22 years of age), who create and maintain blogs made publicly accessible on the Internet. The objectives of the study were to find out the linguistic and non-linguistic features employed when writing blogs so as to determine the extent to which these writing styles were similar linguistically and non-linguistically in line with the checklist adapted from various findings and claims. 40 personal blogs, with five up-to-date entries were selected using the stratified random sampling method with equal number of male- and female- authored blogs. The online identify was gathered from the explicit display of demographic information. Data were analysed and checked against the above checklist to discover possible gender variation in writing. The blog analysis was done at word and sentence level. Overall, presentation of the results and findings showed that the features employed reflected the criteria in the adapted checklist in great similarities with notable difference in the low occurrence of taq questions across all gender. It was also noted that female blog writers showed greater consistency in the use of blogwords in their writing. Besides, the non-linguistic features also showed a high degree of resemblance to the adapted checklist with the significant difference that the spacing element was not applicable at any point in the blog samples studied. The study also discovered that textual emoticons were used more frequently as compared to graphical icons. All these findings confirmed with the literature reviewed that female writing were more ‘interactional’ with personalization of texts and male writing were more ‘informational’ with specification of things and concepts (Biber, 1995 cited in Argamon et al., 2003). As a whole, a close examination of blog analysis categorized the patterns into male preferential features and female preferential features, suggesting possible gender markers of the language. In particular, personal pronouns, intensifier, taq question, blogwords, emoticons and paralanguage were related to female writing whereas preposition, determiner, quantifier and hyperlinks were closely associated with male writing. Contrary to prediction, the results indicated that male blog writers tended to use more modal auxiliary words, which was fairly justifiable according to Saal (2005) to establish a close writer-reader relationship. All in all, the linguistic and non-linguistic features as in the present study demonstrated gender-specific capabilities and might be further replicated to determine the relationship of gender and language use in other contexts of study.
ABSTRAK

LINGUISTIK DAN BUKAN-LINGUISTIK ANALISIS BAGI PERBEZAAN JANTINA DALAM GAYA PENULISAN DALAM BLOG REMAJA

Angeline Hii Hui Wen

CHAPTER 1
INTRODUCTION

1.0 Chapter Overview

The first chapter includes an introduction that provides a general picture on blogs as the social medium of interaction and language use that differs between men and women. The background of study will later presents a description about existing studies and literature regarding the effects of gender on writings. Next, the research problem will express the need for further study. Following that, the aims and objectives that answer the research questions are stated to indicate the purpose of the study. Significance of the study and the definition of terms will provide an understanding on the worth of study and important concepts used throughout this research. Lastly, the boundaries of this proposed study will be covered by the scope of the study.
1.1 Introduction

Gender difference in language use has long been heatedly discussed and the rapid efflorescence of the blogosphere begins to draw concern to the use of language in the computer medium as well. In general, blog is often known as a reversed-chronological online diary that allows commentary about oneself and similarly, blogging, which refers to the practice of writing blogs, continues to emerge as a new form of Internet communication that gained its popularity in the recent year. The ease publishing of content to a website, simple-to-use interface, and an engaged connection with other bloggers has provided a convenient way for users to share their thoughts, life experiences or discuss anything with the world. In addition, the construction of blogs is no longer viewed as requiring expert knowledge since users can customize the appearance of their blogs that best suits their needs with the HyperText Markup Language (HTML) editor which helps description of texts by use of links, headings, paragraphs, embedded images and etc. With all these features, people are offered an unprecedented opportunity to express their thoughts and add new dimension to self-discovery through interacting with the global audience, for little or no cost. In the current study, the influence of gender in the writing of blogs is yet to be discovered. As noted by Hoar (1992), gender differences are profound to form 'genderlects' which contains features that mark it as stereotypically masculine or feminine (cited in Ivy & Backlund, 2004, p.183). The significance of these discriminating features in relation to each gender could therefore distinguish a male writing from a female writing and vice versa.
1.2 Background of Study

As reported by Argamon et al. (2003), most previous works has investigated the influence of gender on spoken discourse (Trudgill 1972; Holmes 1990; Labov 1990; Eckert 1997), informal writing such as student essays (Mulac et al 1990; Mulac & Lundell 1994) and electronic messaging (Herring, 1996). The claim that men and women employ different linguistic styles in writing has always been a great debate.

Gender differences in communication styles gained more attention in this decade and is reflected in the book *Men Are From Mars, Women Are From Venus* written by Prof. John Gray (1992) which asserts the notion that men and woman are ‘from different planets’ as each has different needs, goals, values and communication styles (Saal, 2005). Though the differences as stated have no reliable evidence, they justify the actions of men and women in the society by some means. In addition, the general claim above tends to overlap with the findings of a research carried out by another scholar in the Venus-Mars debate, Deborah Tannen. Tannen (1990), in her book: *You Just Don’t Understand: Women and Men in Conversation* describes the miscommunication of men and women as a result of their ‘different words and different world’ (Saal, 2005).

Biber (1995) also discovered that male writing shows greater signs of features identified as ‘informational’ by previous researchers whereas female writing exhibits greater usage ‘involved’ features (cited in Argamon et al., 2003). The feature of ‘involvedness’ is exemplified through female blog writers who are more likely to use personal pronouns and assent words. This is often described as being affective and tentative and female are said to be more prone to use linguistic styles like compliments and modal auxiliaries to solidify relationships (Argamon et al., 2003). On the other hand, men in general reveal their assertiveness with a dominance of explicit commands and directives in their speech (Coates, 1993, p.124). In relation to this, Schler et al. (2006) also noted the higher frequency of the use of articles.
prepositions and hyperlinks in male blog writers to emphasize the ‘information’. All of this confirms and extends findings in Biber’s dimension of ‘involvedness-information’.

In response to the topic of gender difference in language use, an article in The Boston Globe states that men and women write in ways that reveal the gender of the text immediately. According to the Globe article reports, 600 published documents are punched into a computer by a team of Israeli scientists to devise an algorithm that could predict the sex of the author. The test claims an accuracy of 80%, which is reasonably high (Mcgrath, 2003). This particular article and a test in The New York Times Magazine is an inspiration of the birth of the Gender Genie soon after to identify the gender of the author of sample of fiction texts, non-fiction texts and blog entries based on language use.

The Gender Genie accessible at: bookblog.net/gender/genie.html purports to predict the gender of an author from the genre: fiction, non-fiction and blogs. This online analysis tool is based on an algorithm that claims to estimate the gender of an author with 80% accuracy. The application tabulates weighted frequencies of keywords identified as ‘masculine’ or ‘feminine’ and the author’s gender is assigned according to whichever category that scores the highest. From the findings based on over 500 English-language works, Argamon et al. (2003) found out a pattern in gender-based word. They have proposed that male writers tended to write more about specific things, resulting a greater usage of determiner – ‘a’, ‘the’, ‘that’, ‘these’, cardinal numbers and quantifiers – ‘more’ and ‘some’. In contrast, they also discovered that females are apparently far more likely than men to use personal pronouns especially ‘I’, ‘you’ and ‘she’. This could be explained by the tendency of women being more comfortable thinking about people and relationships whereas men prefer quantifying things.
1.3 Research Problem

The statistics provided by Perseus Development Corporation revealed the increasing population in blogsophere with an estimation of 10 million blogs created in the first quarter of 2005 alone (Scheidt, 2006). Previous research has considered the effects of age and gender on blogging (Scher, 2006), gender and its relationship with genre and writing in formal written texts (Argamon et al., 2003), gender and the use of linguistic features (Argamon et al., 2007) and content analysis of weblogs is carried out as well (Herring, et al., 2006). The relationship among language, gender and genre was also drawn attention to in Gender and genre variation in weblogs by Herring & Paolillo (2006). Although the growing presence of blogs in online communication has led to increased academic study, little interest has been drawn to adolescent writing and how gender-linked markers are found as significant in male and female writings. Since blogging is gaining its popularity among the youth, the ways in which adolescents interact through blogs become important considerations in the recent research.

As yet, the studies conducted on gender variation in blog have revolved around the linguistic use of language. Previous works on gender difference in blog-writing are still relatively under-researched. Currently, there are no researches that discuss blogs as an emergent mode of communication, taking into account the linguistic as well as non-linguistic features into analysis.

In view of the immense growth of blogs, it deserves a detailed study to provide a more comprehensive analysis on the variation of gender in blogs. For that reason, several linguistic features are adapted based on various claims and findings that male and female typically employ different linguistic conversational styles (Coates, 1993, p.114). To enhance the communicative ability of blog authors since Internet language is viewed as a novel medium combined of spoken, written and electronic properties, the analysis on language use in blogs could not achieve better
without the considerations of non-linguistic features such as the use of emoticons and repeated letters or exaggerated punctuations, which are prominent adapted forms of linguistic language of the Internet. The inclusion of additional non-linguistic features, which are obviously lacking in the previous studies could help to add new insights or serve as bonus as to how opposite genders write in ways that reveal their identity.

1.4 Aim and Objectives of the Study

The aim of this study is to investigate if and how gender varies with writing style among adolescent male and female blog writers. It attempts to explore Netspeak as a radically new linguistic medium, which will be very much different from other mediums as it is hard to determine at which point along the continuum it is positioned. The growing popularity of publicly accessible blogs offers a valuable opportunity to harvest information from texts authored by diversified authors. Author-provided demographic information has made this type of study convenient and practicable, assuming that most information is genuine.

This study will be guided by the following objectives:

1. To identify the linguistic features and its consistency as found in adolescent male- and female-authored blogs in line with the checklist adapted.
2. To identify the non-linguistic / paralinguistic features and its consistency as found in adolescent male- and female-authored blogs in line with the checklist adapted.
3. To find out how writing differs linguistically between adolescent male- and female- authored blogs.
4. To find out how writing differs non-linguistically / paralinguistically between adolescent male- and female-authored blogs.

The data set obtained attempts to answer the following research questions:

1. What are the linguistic features found in adolescent male- and female-authored blogs and to what extent are these linguistic features similar to those in the checklist adapted?

2. What are the non-linguistic / paralinguistic features found in adolescent male- and female-authored blogs and to what extent are these non-linguistic / paralinguistic features similar to those in the checklist adapted?

3. How does writing vary linguistically between adolescent male- and female-authored blogs?

4. How does writing vary non-linguistically / paralinguistically between adolescent male- and female-authored blogs?
1.5 **Significance of the Study**

By analysing the language use in blogs, gender variation in writing could be explored in a broader sense, taking into account the linguistic and non-linguistic features of the blog language. It is hoped that the findings of this study might shed light on the way adolescent male and female write differently, specifically in blogs.

The findings from this study could offer insights to be drawn upon on other blogs to determine the gender of the author. Aspects which are similar to those in the checklist could suggest the consistency of dominant features displayed as gender-specific while differences noted could be treated as a valuable piece of information for a more refined study as well as indicate a discovery of possible gendered features that are not found in the checklist but realized in the data analysis process.

This study might impart a better understanding of blogs as an emerging mode of computer-mediated communication, encompassing both linguistic and non-linguistic features. Perhaps, more significantly, the present study intends to extend earlier findings regarding gender-linked variation in writing to include the non-linguistic features as possible gender markers that are still under-researched. The current study suggests that linguistic as well as non-linguistic features are closely-related to explain gender variation in language use in blogs.

In addition, this study could also serve as a useful source of reference to conduct similar studies in a different area as the concept of gender differences on language use as discussed might be applicable in other computer-mediated discourse (CMD) such as emails, chatgroups and virtual worlds. The future researchers might see this useful to examine significant or consistent differences in male and female writing style in other contexts of use.
1.6 Operational Definition of Terms

1.6.1 Blogs

Blogs are defined as frequently modified website in which dated entries are written in reverse chronological sequence (Schmidt, 2007). They are websites made publicly accessible and are often updated with links to topics of similar interest for a second visit. Besides, blogs also refer to an outlet where blog authors keep running record of their personal lives in the entries displayed. For the purpose of this study, only personal journals will be selected to narrow down the scope of study. In this study, blogs refer to the 40 randomly selected sample blogs from the Blogflux blog directory under the teen category. They are single-authored blogs, rich in text which contain a large amount of text rather than pictures, and consist of a minimum of five entries that focused on the personalized experiences of an adolescent.

1.6.2 Writing Style

Style is the dress of thoughts that represents the image of character and emanation from one self (Nordquist, 2009). As such, writing style is the manner in which a person expresses a matter in a text that eventually tells their ‘voice’ through the choice of words, sequence of thoughts and language / structures of grammar used in writing. In this study, writing style is referred to the distinct, gender-specific style of writing revealed through the linguistic and non-linguistic features adopted by adolescent male and female blog writers under study.
1.6.3 Gender

Gender as according to Coates (1993, p.3-4) is a term refers to socially constructed qualities or behaviours that are ideally associated with or especially appropriate to man and woman. The expectation on how a man and woman should act accordingly is communicated to us through the culture we live in and brought up with. These are the unwritten rules of what is expected of us in terms of sex. In this study, gender is an important factor that determines its significance on the writing of blogs. The nature of man being more aggressive and woman passive could be identified through their expression in writings by analyzing the linguistic and non-linguistic features of the texts. As for this study, the gender of the blog authors is determined by their indication of gender status (male or female) at their profile page.

1.6.4 Adolescent Bloggers

The definition of adolescent bloggers in this study has been applied quite flexibly to identify individuals ranging in age from 13 to 22 that post frequent entries to their blogs. In the relative study, adolescent bloggers are individual who state their age as 13 to 22 from the Blogfux blog directory.

1.6.5 Linguistic Analysis

Linguistic analysis refers to the study of language use as detected in the blogs created and maintained by the young adult. In this study, the researcher has adapted eleven linguistic features from various claims and findings to distinguish the gender of the blog writer. The linguistic features studied are as followed and its prevalence is traced in the blogs retrieved: taq questions (Coates, 1993, p.119), hedges (Coates, 1993, p.116; Mulac, n.d.).
modal auxiliary verbs (Coates, 1993, p.124), personal pronouns (Schler et al., 2006; Argamon et al., 2003; Herring & Paolillo, 2006), intensifier (Mulac, n.d.), determiners (Argamon et al., 2003; Herring & Paolillo, 2006), quantifiers (Argamon et al., 2003; Herring & Paolillo, 2006), prepositions (Schler et al., 2006; Argamon et al., 2003; Herring & Paolillo, 2006), numbers and blog specific features that includes blog words and hyperlinks (Scher et al., 2006). A detailed definition of each feature will be explained in Chapter 3 of the study.

1.6.6 Non-linguistic Analysis

Non-linguistic analysis refers to the study of non-verbal elements of communication used to modify meaning and convey emotion. In this study, there are two non-linguistic features adapted that signify possible gender markers: the use of emoticons (Huffaker & Calvert, 2005), and paralanguage such as capitalization and exaggerated use of punctuation (Crystal, 2001, p.34). The prevalence and significance of each is examined in thorough in relation to gender. A detailed definition of each feature will be explained in Chapter 3 of the study.
1.7 **Scope of Study**

This study is carried out with the intention to determine whether adolescent male and female write differently in personal blogs by considering two different kinds of potential distinguishing features: linguistic and non-linguistic.

Though this study discusses the linguistic and non-linguistic features that are significant for gender authorship studies, it does not examine the influence of genres to language use among adolescent male and female blog writers. Apart from this, the scope of the study does not include the effects of age on blogging activities. Furthermore, the choice of colour and the layout of the blog are not the determining factors that help to reveal the gender of the authored text.

It is not the scope of the study to compare the writing style of authors from different countries as the participants are randomly selected regardless of geographical factors. Also, it is not the focus of the present study to explore language use in the context of ethnicity and how language is used as an important marker of social status.
1.8 Chapter Review

This chapter has examined the background of the study, the research problem, the aims and objectives of the study together with their respective research questions, the significance of the study, the definition of terms and the scope of the study. This chapter summarizes personal blog as a new communication genre focuses on previous researches that explore gender differences in language use. Besides, the extent to how the language features can be exploited to determine the gender of the author of a sample of written texts is discussed. The next chapter will elaborate further on the review of literature.
CHAPTER 2
LITERATURE REVIEW

2.0 Preview

In this chapter, a review on the major concepts and related studies were carried out to provide a theoretical base for the research so as to determine the direction of the study in order to provide sufficient background needed to carry out the textual analysis. This chapter starts off with a review of literature on language use in different contexts, with an emphasis on gender influence in writing trends within the online setting. The computer-mediated discourse and the nature of Netspeak as the radical linguistic medium are discussed as well. Next, the nature of blog as an emerging mode of computer-mediated communication is also studied.

With the understanding of concepts gained, previous studies were reviewed in order to look examine the relevant key features that were helpful to pose a significant difference in adolescent writings against the gender of the blog writers. Apart from this, the feasibility of the individual studies is considered for practical concerns of the study. Some key features in writing, which were said to be predominantly male or female are explored and brought clearly into view in order to develop a comprehensive checklist for this study.
2.1 Gender and Language Use

Language is a fundamental human activity through which we communicate our particular representation of the world (Thomas & Wareing, 1999). Gender on the other hand, is often discussed as different individual that possesses separate sub-cultures. Thus, the differences in which a particular gender views the world is often reflected in their use of language and these distinction remained the important area of research. Language is, in fact, an inherently social phenomenon and can provide insights into how men and women approach their social worlds.

The differences in which men and women structure their language have been the interest and practical concern in the study of discourse and this inspired Newman, et al. (2008) to conduct an analysis on 14,000 sample texts from 70 separate studies to observe the impact gender impose on the use of language. The explosion of research on gender studies is tremendous and a popular question that aimed to be answered by the researches is the extent to which men and women use language differently.

Lakoff's theory (1975) on women's language started off to mention hedges and tag questions as features significant for females. A related interpretation of female's use of hedges and tag questions are that females are more reluctant to force their views on another person. Consistent with this idea, Lakoff thus claimed that women are more likely than men in the same situation to use modal verbs and extra-polite forms of language (Coates, 1993). Other researcher like Holmes (1984:1987 cited in Coates, 1993) has found further evidence that women use these features to express their uncertainty. Unlike females who express tentativeness in their language, males on the whole avoid self-disclosure and choose to talk about impersonal objects.