

A Diachronic Analysis of the Intensifier *very* in British English

Osman Mamoon Osman Mohamed^a and Tokhirjon Mirzajonov^b

^aUndergraduate Student at Carnegie Mellon University - Qatar

^bUndergraduate Student at Carnegie Mellon University - Qatar

Plain Language Summary

As an intensifier, the word *very* is used to increase the force or emphasis of other words. Little research exists on how the word *very* is used differently based on an individual's age, gender, or social class. This study examines how the frequency of the word *very* has evolved across these groups over time, using data from a corpus of spoken English samples known as the Spoken British National Corpus. A 1994 set and a 2014 set were compared, and the findings showed that the frequency of *very* grew over time. Individuals 60 and older used *very* more frequently than people of other ages. Men used *very* more often than women in 1994, but their frequencies nearly converged in 2014. Retired people used it the most in 1994, and in 2014 middle-class individuals used the word *very* most frequently. These results are important because they demonstrate how language use can vary between various social groups and change over time, giving us a better understanding of how society and culture affect how we speak.

Publication Category

Course-based Assignment

Academic Context

In *Introduction to Corpus Linguistics*, students complete a high-stakes assignment in which they analyze a corpus of texts using some combination of quantitative and qualitative analyses. The balanced combination and the specific computational methods applied depend on each student's research question(s). As a methods course, close attention is paid to the intentionality of analytical choices and the ability to communicate those intentions in the project.

Introduction

Human language has undergone a great deal of change over time, and linguists have been interested in studying those changes because they provide insight into societal trends and cultural shifts, contributing to our understanding of language as a social phenomenon. These changes may come from changes in sentence structure or in the use of words. In linguistics, studying language change over time is called diachronic variation and is often studied using corpus tools to analyze large data sets.

In this study, we analyze the diachronic variation of the intensifier *very* in the Spoken British National Corpus (BNC) 1994 and 2014. Bolinger (1972) describes intensifiers as "the chief means of emphasis for speakers for whom all means of emphasis quickly grow stale and need to be replaced" (p. 247). Some standard intensifiers in the

English language are *extremely*, *incredibly*, and *great*. They all emphasize the word(s) they modify.

The significance of this research lies in its contribution to linguistic studies, specifically its potential to shed light on the sociolinguistic aspects of language use. By studying the variation of the intensifier *very* across time, age, gender, and social class, we can gain insights into how these factors relate to language use and change. This essay provides a literature review that analyzes past research on the diachronic variation of intensifiers, then details our methods, which include comparing the Spoken British National Corpus (BNC) between 1994 and 2014. The findings are presented in the results section, followed by a discussion that takes into account the limits of our study and makes recommendations for future research. Our data shows that the use of the intensifier *very* has

significantly increased through time, with variances seen across different age groups, genders, and social classes.

Literature Review

Many researchers have investigated the diachronic variation of intensifiers in different English-speaking countries. For instance, Aijmer (2018) studied the diachronic change of intensifiers such as *super*, *real*, and *nice* in the BNC 2014, while Park (2016) examined the variation of the intensifier *literally* in the Corpus of Contemporary American English (COCA). Stratton (2020) traced the use of the intensifier *well* from Early Modern English to Present Day English using various sources. These studies have significantly advanced our understanding of the diachronic variation of intensifiers, revealing shifts in frequency over time and grammatical changes. Some of these studies have begun to show the influence of sociolinguistic variables such as age, gender, and social class on the use and changes in these intensifiers over time. Research has shown these variables to be very important in many dimensions of language and social practice, so there exists a continued need for understanding how these variables interact with language use and change over time.

Researchers like Deborah Tannen (2007), Marjorie Harkness Goodwin (2006), and Deborah Cameron (2008) offer insights into how language use varies by gender, age, and socioeconomic status. According to Tannen's (1990) research, men and women communicate in fundamentally different ways, which may affect how often the intensifier *very* is used. Cameron (2008) presents an alternative perspective on the influence of gender on language use, challenging current notions about the gender gap in verbal skills and behavior by utilizing sociolinguistic evidence and modern ideas of gender and identity. Cameron's (2000) study of language use in British call centers revealed, for example, that the preferred style of speech imposed on workers of both sexes was a symbolically "feminized" manner, indicating that the language used in these professional contexts is not just neutral or random. Instead, they are closely related to society expectations for gender roles and standards. The 'feminized' speech pattern, which is marked by courtesy, empathy, and emotional expression, is frequently connected to conventional female roles and is seen to be more suitable for the customer-focused nature of call center job. Although the use of intensifiers is not specifically discussed in Cameron's (2000) work, it is crucial to remember that intensifiers are frequently linked to a more expressive, sympathetic, and polite style of speech. This phrasing could be seen as a component of the strategically

"feminized" speech patterns that Cameron examines in the study of language use in British call centers. In our study, we extend Cameron's work by specifically examining the use of the intensifier *very* across gender, as well as age and social class.

Goodwin (2006) offers insight into how socioeconomic status affects language use, particularly the choice and application of intensifiers. Goodwin's study has mostly focused on the embodied language practices that children employ to construct their social worlds as they engage with others while playing on the playground or on the street. In our study, we aim to expand upon that into also looking into other age ranges in the BNC 1994 and 2014, as well as social class, to investigate their influences the use of the intensifier *very*.

By analyzing the diachronic variation of the intensifier *very* in the Spoken British National Corpus (BNC) 1994 and 2014 concerning age, gender, and social class, this study aims to contribute to literature on language variation and change. It seeks to shed light on gender norms, age-related language trends, and the impact of socioeconomic class on language use, thereby highlighting the potential social implications of these linguistic patterns.

Methods

Data Collection

This research aims to investigate the general relative frequency, absolute frequency, and the overall count of the intensifier *very* in British English. The Spoken British National Corpus (BNC) 1994 and 2014, which is widely available for linguistic research through BNCLab (Brezina, Gablasova, & Reichelt, 2018), was used in this study. "The British National Corpus (BNC) is a 100-million-word collection of written and spoken language samples from a wide range of sources, designed to represent a wide cross-section of British English from the later part of the 20th century" (*British National Corpus*, 2022). The Spoken BNC 1994 and 2014 corpora are subsections of the BNC, consisting of spoken British English in the late 20th and early 21st century. The Spoken BNC 1994 and BNC 2014 corpora contain transcripts of recorded conversations gathered from the UK public.

We selected the BNC 1994 and 2014 corpora in order to offer a window into social history and a snapshot of language, as they enable a comparison of changes to language and society over a 20-year period (Aijmer, 2018). The BNC project was officially launched in 2014, 20 years after the original British National Corpus was collected and pub-

lished. The median year of the data for the BNC 2014 spoken corpus, gathered from 2012 to 2016, is 2014.

Data Analysis

To study the diachronic variation of the intensifier *very* in the Spoken BNC 1994 and BNC 2014, we examined if there have been any changes in the general relative frequency, absolute frequency, and the overall count of the intensifier from 1994 to 2014 regarding the following sociolinguistic variables: age, gender, and social class.

The analysis was conducted using R programming language and its various packages. This allowed us to determine the frequencies of the intensifier *very* across the different sociolinguistic variables. We then compared these frequencies between the two time periods (1994 and 2014) to identify any significant changes. This approach provided a quantitative measure of the diachronic variation of the intensifier *very* in relation to age, gender, and social class. The results of this analysis are presented in the following section.

Results

The following section reports the diachronic variation of the intensifier *very* in terms of the following sociolinguistic variables: age, gender, and social class between the Spoken British National Corpus (BNC) 1994 and 2014. The search for the token *very* in the BNCLab's Spoken British National Corpus (BNC) 1994 and 2014 returned 13,778 results (1,005.95 per million) in speech.

General Change Over Time

Using quantitative text analysis on datasets obtained from BNCLab utilizing the R programming language and its various packages, we examined how frequently the intensifier *very* was used in both the BNC 1994 and BNC 2014. It is essential to underline that in this study, we focused on absolute counts of the intensifier *very* rather than normalized frequencies.

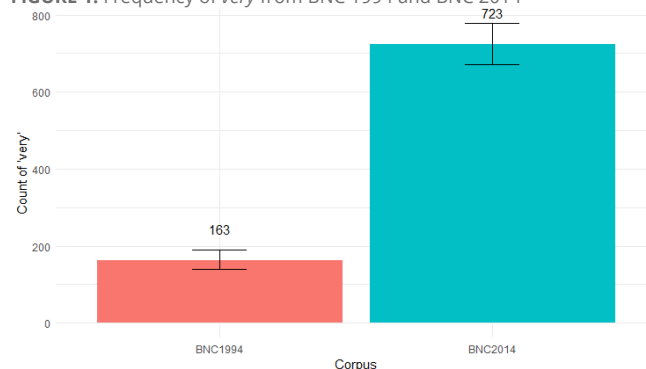
In our analysis, we used a non-parametric approach to calculate the 95% confidence interval based on the chi-square distribution. This approach provides a trustworthy range of the anticipated true count and offers robustness versus potential skewness in the distribution of word frequencies.

The results show an increase in the usage of *very* from BNC 1994 to BNC 2014. In the BNC 1994, the observed count of *very* is 163, with a 95% confidence interval estimated to be from 138.94 to 190.03. In contrast, the BNC 2014 corpus exhibits a substantial rise in the usage of *very*, with an

observed count of 723 and a corresponding 95% confidence interval ranging from 671.25 to 777.68.

These findings are represented in Figure 1, which exhibits the count of *very* in both corpora along with their respective confidence intervals. This significant increase in usage frequency necessitates a deeper language investigation to comprehend the semantic changes and societal influences that may have influenced this trend.

FIGURE 1. Frequency of *very* from BNC 1994 and BNC 2014



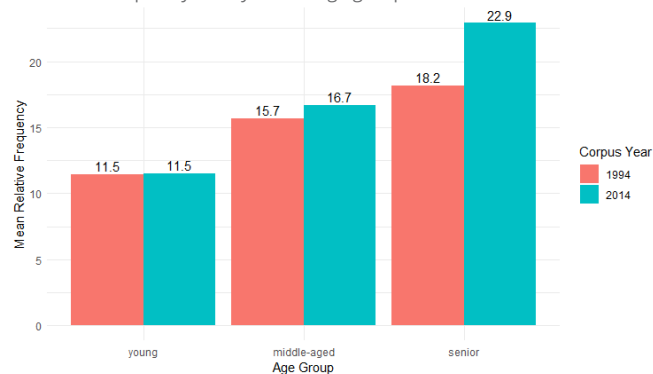
Change Over Time and Age

In the exploration of the use of the intensifier *very* among various age groups, a focused statistical analysis was conducted on the datasets obtained from BNCLab for the corpora of speakers in the BNC 1994 and 2014. This analysis, utilizing the R programming language and its various packages, enabled us to gauge the absolute frequency and relative frequency of *very* regarding the ages of the speakers.

A categorical analysis was conducted on both corpora. This analysis is represented in Figure 2, which categorizes the age of the speakers into three groups: Young (5–29), Middle-aged (30–59), and Senior (60 and over). For each age group and corpus year, we calculated the mean relative frequency of the word *very*. The analysis indicated that the Senior age group used the intensifier *very* most frequently, as suggested by the mean relative frequency values in BNC 2014 (22.9) and BNC 1994 (18.2).

The combined BNC 1994 and 2014 corpus data were split into two separate data frames, each one corresponding to a different year. The absolute frequency of the word *very* was the dependent variable for both corpora, while the speakers' ages were the independent variables in our regression models. Age and the absolute frequency of the word *very* are statistically significant, according to the regression model for the BNC 1994 corpus (p-value:

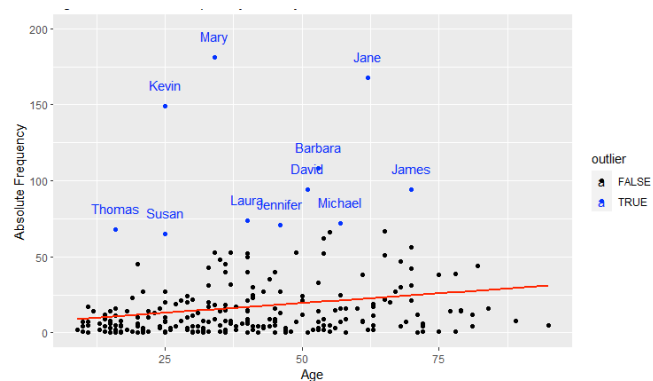
FIGURE 2. Frequency of *very* across age groups in BNC 1994 and BNC 2014.



0.001791). The link between age and the absolute frequency of *very* usage in the BNC 2014 corpus was not statistically significant (p-value: 0.17503). However, the studentized residuals approach allowed for the identification of clear outliers. To investigate further, we examined studentized residuals to identify outliers—individuals who used *very* more frequently than average for their respective age groups, displayed in Figures 3 & 4.

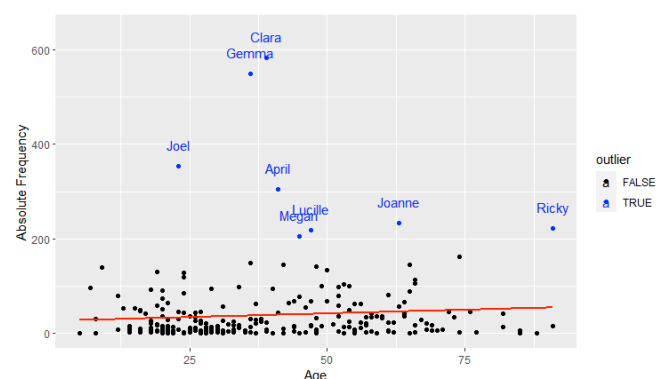
Figure 3 displays the absolute frequency of *very* in terms of age in the Spoken BNC 1994 corpus. The figure shows that we found 11 outliers: Kevin (149 uses), Jane (168 uses), Laura (74 uses), David (94 uses), Thomas (68 uses), Mary (181 uses), James (94 uses), Barbara (108 uses), Michael (72 uses), Susan (65 uses), and Jennifer (71 uses). These individuals used *very* more frequently than the average for their respective age groups.

FIGURE 3. Absolute frequency of *very* across age groups in BNC 1994



Similarly, for the Spoken BNC 2014 corpus shown in Figure 4, which also displays the absolute frequency of *very* in terms of age. The graph illustrates 8 outliers were identified: Gemma (548 uses), Clara (582 uses), Joel (354 uses), Ricky (223 uses), Megan (205 uses), Joanne (234 uses), April (304 uses), and Lucille (218 uses). These individuals showed higher usage of *very* in comparison to the average within their respective age groups.

FIGURE 4. Absolute frequency of *very* across age groups in BNC 2014.

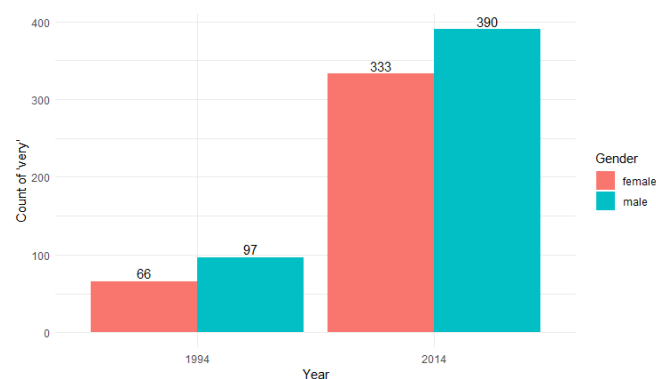


Change Over Time and Gender

Using the R programming language, we conducted a quantitative study to examine variations in the usage of the intensifier *very* among males and females between the BNC 1994 and 2014 datasets accessible in BNCLab. The raw data were imported into R, where a function was constructed to determine the overall count of *very* within the various subgroups (i.e., speakers who identified as male or female in 1994 and 2014). Male and female were the only available gender categories in the corpus.

Figure 5 illustrates the change in the absolute count of *very* across the years for both genders. For males, the count of *very* ascended from 97 in 1994 to 390 in 2014. Females' usage also rose from 66 to 333.

FIGURE 5. Frequency of *very* across gender in BNC 1994 and 2014.



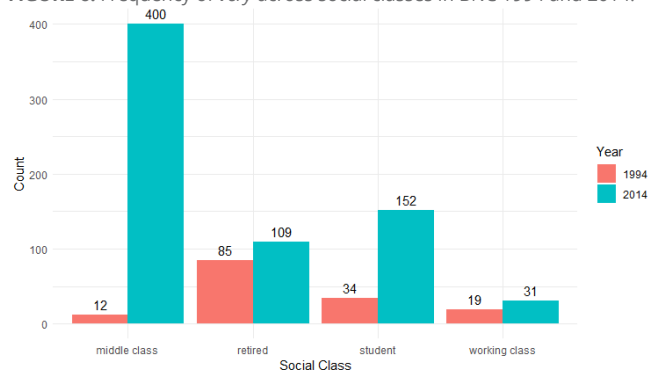
For each year independently, an exact binomial test was run to determine the significance of these shifts. To evaluate if the observed changes are statistically significant, this test—a particular kind of a two-proportion z-test—compares the "success" frequency (count of *very*) to the total trials (total counts). These analyses revealed that the proportions of the intensifier *very* used by males and females varied considerably over time (p-value = 0.01851 for 1994 and p-value = 0.03721 for 2014), showing a significant change in the intensifier's usage over time.

Change Over Time and Social Class

Utilizing the "R" statistical programming language, we analyzed the usage of the intensifier *very* across different social classes in the BNC 1994 and BNC 2014, from datasets which are accessible in BNCLab. It is important to note that we excluded the "unknown" social class corpus category from our analysis due to their unidentified status.

The bar plot in Figure 6 shows this data, giving a complete picture of how often the intensifier *very* was used between BNC 1994 and BNC 2014 across social classes. For the year 1994, there were 12 instances of *very* from the middle class, 85 from the retired class, 34 from students, and 19 from the working class. For 2014, the instances were 400 for the middle class, 109 for the retired class, 152 for students, and 31 for the working class.

FIGURE 6. Frequency of *very* across social classes in BNC 1994 and 2014.



According to Figure 6, the social class retired group used the intensifier *very* the most in BNC 1994. However, in the BNC 2014, the middle class group used this intensifier the most frequently.

We performed a proportion test for each social class to determine the statistical significance of the observed differences in word usage between 1994 and 2014. The null hypothesis stated that there was no difference in the usage of *very* between the two years within each class.

The test results showed that the middle class, retired, and working class all had significant shifts in the usage of *very* between 1994 and 2014, with p-values of less than $2.2e-16$. However, the student class had a p-value of 0.8511, indicating no significant change in the usage of *very* over the course of the two decades.

These results provide strong evidence that the usage of *very* changed significantly across different social classes between 1994 and 2014, with the notable exception of students. Additional qualitative research may shed light on the factors influencing these changes in language usage.

Discussion

The quantitative analysis of the intensifier *very* in the years of 1994 and 2014 using BNC data suggests that the relative frequency, absolute frequency, and the overall count of this intensifier increased over time. The parameters used in this research to determine the diachronic variation of the intensifier *very* were age, gender, and social class.

The analysis of different age groups showed that in both corpora, the Senior age group, defined as those 60 years and older, used the intensifier *very* more frequently. This finding suggests a pattern that could be indicative of a linguistic preference or style among this demographic. For instance, phrases such as "mm very good," "it was very funny," and "that was very timid" were more commonly observed in the speech of individuals aged 60 and above.

The study of intensifiers used by different genders remains a debated topic with some studies suggesting that the use of intensifiers is more common for female or vice versa. According to the present study's findings, men had a larger absolute count of *very* than women. In light of these findings, Deborah Tannen's theory that men often use a more assertive speaking style deserves some consideration. Guri et al. (2022) indicated that women in a male-dominated industry in Malaysia would change to a more assertive report-talk style while speaking with their male coworkers, which lends support to this argument. Men's usage of the intensifier *very* may be an example of assertive language. But other academics, like Deborah Cameron and Marjorie Harkness Goodwin, contend that there are not as many obvious disparities between men and women when it comes to language use. Particularly relevant in this context is Cameron's (2005) work on the role of language in defining gender and gender equality. Cameron (2005) argues that gender disparities in language use are socially produced and context-dependent rather than intrinsic. This viewpoint aids in the analysis of the present findings by posing the possibility that observed gender disparities in the usage of *very* may be impacted by social and contextual variables rather than being a defining feature of male or female speech. Our findings may support this work, particularly Cameron's theory of context-dependent gender differences in language use. Female-reported increased use of *very* from BNC 1994 to BNC 2014 may reflect shifting societal norms and gender roles over time.

One reason for the lower count of the intensifier *very* by females might be their preference for other intensifiers with similar meanings such as *really*. Notably, the absolute count for females increased from the BNC 1994 to BNC

2014 reaching 333, and the males absolute count in BNC 2014 is 390, which might suggest that in the future the absolute count in use of intensifier *very* might become equal for both males and females.

The present research adds to what Goodwin's (2006) study on the impact of socioeconomic status on language use. Goodwin's research largely focuses on children's language use, but our study extends this to adults and shows that socioeconomic factors have an impact on how often the intensifier *very* is used.

Our study adds to the larger conversation on language diversity and change by concentrating on a particular linguistic trait and examining its use over time, ultimately showing that age, gender, and socioeconomic status surface in these variations, and warranting further studies on language use and the socializing practices and contexts across these populations.

Limitations and Future Research

This study's focus on the years 1994 and 2014 prevents it from analyzing minute variations or slow shifts in the frequency of the intensifier *very*. Additionally, the research is based on Spoken British National Corpus (BNC) data from 1994 and 2014, which has a small sample size, particularly for 1994. As a result, the gradual shift in the frequency of this intensifier among males and females could not be visualized clearly.

Future studies could benefit from a more thorough analysis using additional corpus data sources that may produce different results. Future research could also benefit from a more concentrated focus on the social and identity factors of the users of the intensifier *very*, an aspect this paper did not extensively explore. This could involve conducting a more comprehensive analysis of the literature on the relationship between language use and socioeconomic class, gender, and age all together.

Conclusion

The present research found an increase in the intensifier *very*'s overall relative frequency, absolute frequency, and count from the spoken BNC 1994 to BNC 2014. Seniors who are 60 or older tend to use *very* more frequently. Usage increased for both males and females between 1994 and 2014. The middle social class used the intensifier the most in the BNC 2014 corpus compared to the retired social class who used it the most in the BNC 1994 corpus. These results underscore the need for more study into these sociolinguistic characteristics and add to the larger

discussion on how language use differs and changes across age, socioeconomic status, and gender.

Acknowledgements

We would like to express our gratitude to Professor Maria Pia Gomez Laich for her guidance in this project during our *Introduction to Corpus Linguistics* course. We would also like to thank the three anonymous reviewers for *Majāl*, who offered substantive and impactful feedback.

References

- Aijmer, K. (2018). Corpus approaches to contemporary British speech. In V. Brezina, R. Love, & K. Aijmer (Eds.), *Corpus Approaches to Contemporary British Speech: Sociolinguistic Studies of the Spoken BNC 2014*. Routledge. DOI: 10.4324/9781315268323
- Bolinger, D. (1972). Degree words. In *Janua Linguarum. Series Maior*. De Gruyter Mouton. DOI: 10.1515/9783110877786
- Brezina, V., Gablasova, D. & Reichelt, S. (2018). *BNClab*. Lancaster University. <http://corpora.lancs.ac.uk/bnclab>
- British National Corpus*. (2022). Oxford University. <http://www.natcorp.ox.ac.uk/>
- Cameron, D. (2000). Styling the worker: Gender and the commodification of language in the globalized service economy. *Journal of Sociolinguistics*, 4(3), 323–347. DOI: 10.1111/14679481.00119
- Cameron, D. (2005). Language, gender, and sexuality: Current issues and new directions. *Applied Linguistics*, 26(4), 482–502. DOI: 10.1093/applin/ami027
- Cameron, D. (2008). *The Myth of Mars and Venus*. Oxford University Press.
- Goodwin, M. H. (2006). *The Hidden Life of Girls: Games of Stance, Status, and Exclusion*. Wiley-Blackwell. DOI: 10.1002/9780470773567
- Guri, G., Md Zolkapli, R. B., & Mohd Kenali, S. F. (2022). Female language used in a male-dominated workplace in Malaysia. *Malaysian Journal of Social Sciences and Humanities*, 7(10), e001840. DOI: 10.47405/mjssh.v7i10.1840
- Park, Semi. 2016. Literally does not always mean literally: A corpus- based diachronic study on literally as an

intensifier. *SNU Working Papers in English Linguistics and Language* 14, 124-142.

Stratton, J. M. (2020). A diachronic analysis of the adjective intensifier well from Early Modern English to Present Day English. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, 65(2), 216–245. DOI: 10.1017/cnj.2020.6

Tannen, D. (2007). *You Just Don't Understand: Men and Women in Conversation*. William Morrow Paperbacks.