Journal of English Language Teaching and Applied Linguistics
ISSN: 2707-756X
DOI: 10.32996/jeltal
Journal Homepage: www.al-kindipublisher.com/index.php/jeltal

A Corpus-based Study of the Use of Pause Fillers Among British English Speakers

Hamad Abdullah H Aldawsari
Prince Sattam Bin Abdulaziz University, Saudi Arabia
✉ **Corresponding Author:** Hamad Abdullah H Aldawsari, **E-mail**: H.alnawwar@psau.edu.sa

ARTICLE INFORMATION

Received: 15 October 2021
Accepted: 02 November 2021
Published: 17 November 2021
DOI: 10.32996/jeltal.2021.3.12.2

KEYWORDS

Pause fillers, sociolinguistic variables, corpora

ABSTRACT

Many people use *pause fillers* such as *um*, *erm*, and *er* in order to signal to the other person that they have not finished speaking yet. This paper aims to investigate pause fillers and their relationship with the two sociolinguistic variables of age and gender. The data-driven analysis is based on the British National Corpus (BNC). The results show that the sociolinguistic variables of age and gender influence the use of pause fillers among British English speakers, which is proposed to be linked to the advancement of age and an improved fluency among female speakers.

1. Introduction

When people speak to each other, they sometimes pause for a moment before continuing. In order to signal to the other person that they have not finished speaking yet, many people use a word or a sound to indicate this status. These words or sounds are identified as *pause fillers*. This paper will attempt to investigate three of these pause fillers (*um*, *erm*, and *er*) and their relationship with the two sociolinguistic variables of age and gender. The paper will base its analysis on the British National Corpus (BNC), and, more specifically, will utilize the online corpus manager BNCweb to perform the analysis. The approach to this analysis will be data-driven, meaning that no assumptions are held before the beginning of the investigation. Before engaging in the topic, a brief overview of pause fillers, sociolinguistic variables, and the BNC will be presented.

2. Overview

2.1 *Pause fillers*

Pause fillers describe 'all of the non-silence devices that can be deployed after the current word has been brought to completion to delay the next word due' (Amiridze, Davis and Maclagan, 2010, p.2). These words or sounds are used to signal that the speaker has not yet finished what he/she wants to say (Juan, 2010). Each language uses different words or sounds to signal this pause. In English, typical ones include *um, erm,* and *er*, which this paper intends to investigate through the two sociolinguistic variables of age and gender.

2.2 Sociolinguistic variable

Focus on language variation in relation to sociology began with William Labov's work on New York community speech in the 1960s (Holmes, 2008). Many studies have since looked at how the use of language changes based on background, age, gender, and other variables (Trudgill, 2001). The emphasis in this paper will be on the two sociolinguistic variables of age and gender and how they affect the use of pause fillers among British English speakers.

2.2.1 *The sociolinguistic variable of age*

Age variation relates to changes in the language used by people in the same community but in different age groups (Trudgill, 2001). A more specific type of age variation focused on in this paper is age-graded variation, which relates to single linguistic changes within a stable community (Labov, 2010). An idealized pattern of age-graded variation can be described as a U-shaped curve, where 'young' speakers (0-29 years) and 'old' speakers (51+ years) use language with less formal constraints, while 'middle-

Copyright: © 2021 the Author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) 4.0 license (https://creativecommons.org/licenses/by/4.0/). Published by Al-Kindi Centre for Research and Development, London, United Kingdom.

aged' speakers (30-50 years) use more formal language (Holmes, 2008). A question remains, however, as to whether this idea would be reflected in the analysis of pause fillers.

2.2.2 The sociolinguistic variable of gender

In addition to changes in language use based on age, scholars have also determined that language use changes according to the speaker's gender (Coates, 1986). Though at first controversially evaluated as a sign of women's inferiority (Lakoff, 1975), later studies argued that these changes are similar to 'cultural' changes in which male speakers aim to be more factual in their communication, while women are more concerned with relationships in their communication (Tannen, 1990). Concentrating on the use of pause fillers, it would be interesting to examine whether one gender is more likely to use them than the other.

2.3 *British National Corpus*

The British National Corpus (BNC) is an electronic collection containing 100 million words in British English, consisting of about 90 million written words and 10 million spoken words (Aston and Burnard, 1998). For this paper, the online corpus manager BNCweb is used to carry out an analysis of three pause fillers: *er*, *erm*, and *um*. BNCweb allows for distribution analysis of queries according to several criteria, including speaker age and gender (Hoffmann et al., 2008). Not only that, but BNCweb is also capable of performing a cross-tabulation function, which allows two criteria to be cross-analyzed collectively. This function is particularly useful in this research, as it allows the age and gender variables to be examined together.

3. Methodology

This paper intends to compare and contrast the three linguistic pause fillers *er, erm,* and *um* in relation to the two sociolinguistic variables of age and gender while utilizing BNCweb to conduct the analysis. To accomplish this, several steps were taken so the analysis could be thoroughly conducted. These steps include retrieving the frequencies of the three linguistic fillers, performing an age and gender distribution analysis on them, and investigating the results both individually and collectively.

3.1 Frequencies retrieval

To analyze the three pause fillers selected for this study, raw frequencies and normalized frequencies were first retrieved from the BNC. This was simply done using the BNCweb interface (Hoffmann et al., 2008). After registering, logging into BNCweb allows the user to search for any term using the query page. Using the query page, each pause filler was researched separately. An effort was made to make sure all instances of the pause filler researched were functioning as pause fillers and not as something else, such as backchannels, that is, as words used to show listener response (Yule, 2010). To ensure that, it was found that restricting the analysis to UNC (unclassified words) as recognized in the CLAWS-5 tagset[2] yielded more accurate results for pause fillers. The results of each query were recorded to serve as grounds for testing any potential statistical significance.

Focusing on the raw frequencies and normalized frequencies of the three linguistic pause fillers, it was found that the pause filler *um* returned significantly fewer results than the pause fillers *er* and *erm*. More specifically, the pause filler *um* only returned 282 hits in 33 different texts (or 27.09 instances per million words), while the *er* returned 88354 hits in 856 different texts (or 8487.53 instances per million words) and the *erm* returned 62352 hits in 811 different texts (or 5989.71 instances per million words). Initially, looking at these results made the writer consider focusing on only the two pause fillers *er* and *erm*, as their frequency figures appeared to be more valid and reliable for the analysis. However, considering the diversity of the texts the pause filler *um* appeared in as well as the number of hits, it was ultimately decided that all three linguistic fillers would be examined regardless of the difference in frequency inside the BNC. The pause filler *um* might not be as frequent as the pause fillers *er* and *erm*, but it would still be investigated.

3.2 Distribution analysis

In order for the two sociolinguistic variables of age and gender to be tested, a distribution analysis on the three linguistic pause fillers was performed using BNCweb. First, the individual raw and normalized distributions of each of the sociolinguistic variables of age and gender were obtained in order to perform a specific investigation of the three pause fillers in relation to these variables. This meant that any hits obtained from making a query about any pause filler were distributed according to how old the speaker was and was then distributed again according to the sex of the speaker. After recording these results, the cross-tabulation function in BNCweb was used to cross these results together so collective results could be gathered. So, rather than having separate results distributing the pause filler according to speaker age and speaker gender, these two variables were blended together to produce a result that would show both the age and the gender of the speaker for each pause filler.

[2] For a description of the CLAWS automatic grammatical tagging system, see Garside (1987).

Though the gender distribution was simply divided into two categories, male and female, the age distribution was divided into six categories: 0-14, 15-24, 25-34, 35-44, 45-59, and 60 and above.

The results of the distribution analysis are represented in tables rather than graphs to allow for more information to be included, for example, dispersion over speakers.

4. Results and discussion

4.1 Pause fillers and the sociolinguistic variable of age

Focusing on the sociolinguistic variable of age, it was found that results differed between the three linguistic pause fillers *um, erm,* and *er*. Starting with the pause filler *um*, it was noted from the analysis that younger speakers used fillers more often than older speakers. However, due to the low number of hits, as well as the low percentage of speakers who used this pause filler, this result cannot be reliable. Table 4.1 illustrates the results of the distribution analysis of the age variable, ranking the age categories using the same methodology as BNCweb, which ranks them according to the number of hits per million words.

Table 4.1. BNCweb age distribution analysis of pause filler *um*

Age group	No. of words	No. of hits	Dispersion (over speakers)	Frequency per million words
25-34	1,120,516	220	5/351	196.34
0-14	385,234	7	6/258	18.17
45-59	1,638,364	19	3/436	11.6
15-24	594,400	4	2/302	6.73
35-44	1,075,749	6	3/335	5.58
60+	1,137,433	3	3/318	2.64
total	5,951,696	259	22/2,000	43.52

Table 4.1 shows that speakers aged 25-34 and 0-14 used *um* more than any other age category, with frequencies of 196.34 and 18.17 per million words, respectively, while speakers aged 35-44 and 60+ used *um* less than any other group, with frequencies of 5.58 and 2.64 per million words, respectively. It is also worth noting that people aged 45-59 and 15-24 scored in the middle of the table, with the age group 45-59 scoring higher than 15-24.

However, as noted previously, there were fewer hits for *um* than for the other two pause fillers investigated in this paper. Not only that but the results are dispersed among very few people. Considering how these results came from only a small percentage of the speakers recorded in the BNC, no real conclusion can be drawn from it other than to imagine that the use of *um* was perhaps more of a personal trait in some individuals rather than a general feature. However, there may not be enough evidence from this study to support that claim.

The pause filler *erm* showed slightly different results than *um*, but with much more frequency in the BNC. These results are shown in Table 4.2.

Table 4.2. BNCweb age distribution analysis of pause filler *erm*

Age group	No. of words	No. of hits	Dispersion (over speakers)	Frequency per million words
25-34	1,120,516	7,126	261/351	6359.57
45-59	1,638,364	9,031	340/436	5512.21
15-24	594,400	3,205	217/302	5391.99
35-44	1,075,749	5,283	237/335	4911
60+	1,137,433	5,218	193/318	4587.52
0-14	385,234	1,767	161/258	4586.82
total	5,951,696	31,630	1,409/2,000	5314.45

The results show fluctuations in the use of the pause filler in relation to the age group. The speaker age group 25-34 used *erm* the most (with a frequency of 6359.57 per million words), followed by the age group of 45-59 (with a frequency of 5512.21 per million words). Slightly behind that was the age group 15-24 (with a frequency of 5391.99 per million words) followed by two older age groups of 35-44 and 60+, with frequencies of 4587.52 and 4586.82 per million words, respectively. Lastly, the analysis shows the youngest age group of 0-14 used the pause filler *erm* the least.

Unlike the pause filler *um*, *erm* appears to be spread over more speakers, which might allow for a more reliable consideration of the results. These results are discussed together with the following results from the analysis of *er*.
The investigation of the pause filler *er* revealed a consistent descending order through speaker age groups. It shows that older speakers tended to use *er* more than younger ones. Table 4.3 presents the results of the analysis.

Table 4.3. BNCweb age distribution analysis of pause filler *er*

Age group	No. of words	No. of hits	Dispersion (over speakers)	Frequency per million words
60+	1,137,433	15,138	264/318	13308.92
45-59	1,638,364	12,361	360/436	7544.72
35-44	1,075,749	7,283	279/335	6770.17
25-34	1,120,516	6,896	284/351	6154.31
15-24	594,400	2,929	213/302	4927.66
0-14	385,234	1,536	163/258	3987.19
total	5,951,696	46,143	1,563/2,000	7752.92

Table 4.3 displays the 60+ age group (with a frequency of 13308.92 per million words) as the highest user of *er*, then systematically descends through the younger groups 45-59, 35-44, 25-34, 15-24, and 0-14 with slightly sharp drops in frequencies per million words through each category (13308.92, 7544.72, 6770.17, 6154.31, 4927.66 and 3987.19, respectively), clearly demonstrating a tendency for *er* to be used more by older people.

Looking at the results of *erm* and *er* together, a number of inferences can be drawn. First, it is clear that people over age 60 use the pause filler *erm* and *er* the most (with a combined frequency of 17,896.44 per million words), with an obvious tendency to use the pause filler *erm* more than *er*. This might be related to the slow rate of speech of older people (Mitzner and Kemper, 2003), which might lead them to use these pause fillers in their speech. Second, people aged 0-14 are the least frequent users of *erm* and *er* compared to other age groups suggesting perhaps fewer pauses during speech. Lastly, the use of pause fillers among people aged from 15-59 varies noticeably, but there is still a tendency to use the pause fillers *er* and *erm*. It could be suggested that generally, the older the speaker is, the more pause fillers he/she uses.

In summary, the three pause fillers showed varying results regarding how often they were used by different age groups. The anticipated U-shaped curve of the age-grading hypothesis did not appear to apply to the pause fillers examined here, perhaps because it is more closely related to 'words' rather than 'sounds' (Labov, 2010). Overall, it appears that the pause filler *er* and *erm* are used more often by older speakers and least often by younger speakers. No conclusion can be drawn from results on the pause filler *um* for lack of evidence. Such varying results are problematic in regards to making any generalizations about pause filler use by speakers of different ages. For that reason, perhaps no solid conclusion can be drawn from this outcome.

4.2 Pause fillers and the sociolinguistic variable of gender

In contrast to the previous age distribution analysis in which results for the three pause fillers *um, erm,* and *er* appeared to be inconsistent with each other, results from the speaker gender distribution analysis were consistent throughout the three pause fillers. In the three fillers, the analysis showed that male speakers tend to use pause fillers much more than female speakers. In the case of *um*, although the number of hits is not as large as the other two pause fillers and the number of speakers is low, the analysis still showed that male speakers use *um* more than female speakers, as shown in Table 4.4.

Table 4.4. BNCweb gender distribution analysis of pause filler *um*

Gender	No. of words	No. of hits	Dispersion (over speakers)	Frequency per million words
Male	4,949,938	250	17/2,448	50.51
Female	3,290,569	18	12/1,360	5.47
total	8,240,507	268	29/3,808	32.52

As displayed in the table, male speakers were recorded to use the pause filler *um* 250 times (with a frequency of 50.51 per million words) in the BNC, while female speakers used it only 18 times (with a frequency of 5.47 per million words). The difference in distribution is striking and probably suggests that *um* is a male-gendered trait. However, as with the age distribution analysis of *um*, no concrete claim can be made from these numbers other than to suggest that perhaps the use of *um* is a personal trait of some individual, mostly male, speakers.

The gender distribution analysis of the pause filler *erm* revealed a somewhat similar result to *um*, albeit with more instances per million words and also more dispersion over speakers sampled in the BNC. Table 4.5 illustrates the results of the analysis.

Table 4.5. BNCweb gender distribution analysis of pause filler *erm*

Gender	No. of words	No. of hits	Dispersion (over speakers)	Frequency per million words
Male	4,949,938	30,940	1,515/2,448	6250.58
Female	3,290,569	15,796	889/1,360	4800.39
total	8,240,507	46,736	2,404/3,808	5671.5

Looking at the distribution, it is clear that male speakers use *erm* more frequently than female speakers. Although the number of hits for male speakers is about double that of female speakers, the normalized frequencies adjusted to the number of words spoken by males identified in the BNC is somewhat less than that, scoring 6250.58 instances per million words for male speakers and 4800.39 instances per million words for female speakers.

Examining the final pause filler *er* yielded results fairly similar to those of the other two pause fillers. Gender distribution analysis of the filler *er* showed that male speakers use *er* significantly more often than female speakers, as presented in Table 4.6.

Table 4.6. BNCweb gender distribution analysis of pause filler *er*.

Gender	No. of words	No. of hits	Dispersion (over speakers)	Frequency per million words
Male	4,949,938	49,949	1,833/2,448	10090.83
Female	3,290,569	16,421	910/1,360	4990.32
total	8,240,507	66,370	2,743/3,808	8054.12

Table 4.6 reveals a substantial difference in frequency of using *er* between male and female speakers. Even in normalized frequencies, the tendency for male speakers to use the pause filler appears to be more than double that of females.

To sum up, the gender distribution analysis indicates that all three pause fillers are used by male speakers more often than by female speakers. It might be argued that female speakers have fewer pauses in their speech than males and are thus more fluent. However, this study examined three pause fillers only. Perhaps female speakers use fewer of these three pause fillers because they use other pause fillers or rely more on 'discourse markers' such as 'like' and 'you know' (Laserna, Seih, and Pennebaker, 2014). Either way, a more comprehensive analysis is needed before any reliable generalization can be claimed. What is left now is to highlight the results drawn from attempting to cross the two previous distribution analyses of age and gender together.

4.3 Pause fillers and cross analyses of age and gender

After looking at the three pause fillers through each of the two sociolinguistic variables of age and gender individually, this section will highlight results compiled from looking at the three pause fillers through the two variables of age and gender collectively. These results indicate that speakers of all ages and both sexes use the three pause fillers with varying frequencies. Beginning with the pause filler *um*, BNCweb's cross-tabulation analysis of age and gender highlighted that males aged 25-34 and 45-59 used *um* more often than males of other ages, while females aged 0-14 and 35-44 used it more often than females of other ages. In addition, the analysis highlighted some divergence in male and female use of *um* according to age, with the female age groups 0-14, 35-44, and 60+ using *um* more than their male counterparts and the female age groups 15-24, 25-34 and 45-59 using *um* less frequently than their male counterparts. These results are specified in Table 4.7.

Table 4.7. BNCweb age and gender cross-distribution analysis of pause filler *um*

Gender: Male				
Age group	No. of words	No. of hits	Dispersion (over speakers)	Frequency per million words
25-34	549,763	219	4/201	398.35
45-59	1,086,180	16	2/303	14.73
15-24	237,033	3	1/143	12.66
0-14	224,388	2	2/152	8.91
35-44	558,419	1	1/190	1.79
60+	599,631	1	1/201	1.67
total	3,255,414	242	11/1,190	74.34
Gender: Female				
Age group	No. of words	No. of hits	Dispersion (over speakers)	Frequency per million words
0-14	160,846	5	4/106	31.09
35-44	517,330	5	2/145	9.67
45-59	550,563	3	1/132	5.45
60+	537,802	2	2/117	3.72
15-24	357,367	1	1/159	2.8
25-34	570,544	1	1/149	1.75
total	2,694,452	17	11/808	6.31

As it can be seen, even though the analysis categorized male and female speakers according to their ages, the frequency, as well as the dispersion over speakers, are still relatively low, preventing any significant interpretation. The other two pause fillers, on the other hand, yielded more results.

The age and gender cross-analysis of the pause filler *erm* (shown in Table 4.8) revealed varying degrees of frequency among male and female age groups. Most interestingly, male speakers aged 25-34 and 35-44 used *erm* more than any other male age group (with frequencies of 7288.6 and 6419.91 per million words, respectively), while males aged 0-14 and 60+ used it the least (4514.5 and 4414.38 instances per million words, respectively). On the other hand, female speakers aged 25-34 and 15-24 used the filler more often than any other female age group (with frequencies of 5457.95 and 5271.89 per million words, respectively), whereas females aged 45-59 and 35-44 used it the least (3919.62 and 3282.24 instances per million words, respectively).

Table 4.8. BNCweb age and gender cross-distribution analysis of pause filler *erm*.

Gender: Male				
Age group	No. of words	No. of hits	Dispersion (over speakers)	Frequency per million words
25-34	549,763	4,007	150/201	7288.6
35-44	558,419	3,585	134/190	6419.91
45-59	1,086,180	6,860	241/303	6315.71
15-24	237,033	1,321	98/143	5573.06
0-14	224,388	1,013	92/152	4514.5
60+	599,631	2,647	101/201	4414.38
total	3,255,414	19,433	816/1,190	5969.44
Gender: Female				
Age group	No. of words	No. of hits	Dispersion (over speakers)	Frequency per million words
25-34	570,544	3,114	110/149	5457.95
15-24	357,367	1,884	119/159	5271.89
60+	537,802	2,571	92/117	4780.57
0-14	160,846	754	69/106	4687.71
45-59	550,563	2,158	98/132	3919.62
35-44	517,330	1,698	103/145	3282.24

total	2,694,452	12,179	591/808	4520.03

In addition to examining the ranking of age groups for males and females individually, collectively examining them in Table 4.8 shows that female speakers aged 0-14 and 60+ used the pause filler *erm* more than male speakers in the same age groups. On the other hand, male speakers between the ages of 15 to 59 used *erm* more than female speakers of the same ages. The degree of difference varies between females and males of the same age group. Some appear to have large differences (e.g. frequencies of 6419.91 to 3282.24 per million words in the age group 35-44), while others appear to have small differences (e.g. frequencies of 4687.71 to 4514.5 per million words in the age group 0-14).

Conducting a cross-distribution analysis of age and gender on the pause filler *er* indicated that, generally, older male and female speakers used *er* more often than younger ones. Table 4.9 illustrates the distribution analysis.

Table 4.9. BNCweb age and gender cross-distribution analysis of pause filler *er*.

Gender: Male				
Age group	No. of words	No. of hits	Dispersion (over speakers)	Frequency per million words
60+	599,631	10,271	162/201	17128.87
35-44	558,419	5,408	161/190	9684.48
45-59	1,086,180	9,663	264/303	8896.32
25-34	549,763	4,883	169/201	8882.01
15-24	237,033	1,461	100/143	6163.7
0-14	224,388	1,100	96/152	4902.22
total	3,255,414	32,786	952/1,190	10071.22
Gender: Female				
Age group	No. of words	No. of hits	Dispersion (over speakers)	Frequency per million words
60+	537,802	4,867	102/117	9049.8
45-59	550,563	2,653	95/132	4818.7
15-24	357,367	1,468	113/159	4107.82
35-44	517,330	1,875	118/145	3624.38
25-34	570,544	2,010	114/149	3522.95
0-14	160,846	436	67/106	2710.67
total	2,694,452	13,309	609/808	4939.41

Examining the table shows that with the exception of female age groups 15-24 and 35-44, *er* seems to be used more by older female and male age groups and less by younger age groups. Furthermore, it appears that male speakers of all ages use *er* significantly more often than their female counterparts.

Results from the age and gender cross-analysis of the two pause fillers *erm* and *er* showed clear variance in their uses by male and female English speakers of different ages. Investigating the two variables together revealed a number of surprising results that would otherwise be very difficult to infer from performing individual age and gender analyses. Perhaps the most unanticipated of these results was that, contrary to the gender analysis conclusion that male speakers use pause fillers more than female speakers, the cross-analysis of age and gender showed that female speakers in certain age groups actually used the pause filler *erm* more than their male counterparts. In the age groups 0-14 and 60+, females appeared to use *erm* more than males in the same age groups. This demonstrates the strong usefulness of cross-analysis of two sociolinguistic variables, as it sometimes reveals information about the data that normally would not be perceived.

5. Conclusion, implications, and limitations

This study investigated how three pause fillers *um*, *erm*, and *er* vary in speech across the sociolinguistic variables of age and gender in the BNC. The study showed that the sociolinguistic variables of age and gender influence the use of pause fillers among British English speakers. It demonstrated that, in general, older speakers use more pause fillers than younger speakers, and male speakers use pause fillers more often than female speakers. The study also explored the interaction between age and gender in relation to the use of the three pause fillers. It discovered that, although male speakers generally use the three fillers more, female speakers use the pause filler *erm* more than male speakers in the youngest and oldest age groups (0-14 and 60+). Interpretations of some of the results were presented, including the slow speech rate of older people as a probable cause of more frequent pause fillers (Mitzner and Kemper, 2003), as well as the hypothesis that female speakers pause less when speaking while male speakers probably

pause more often, suggesting that females are either more fluent or perhaps use different pause fillers or discourse markers (Laserna, Seih and Pennebaker, 2014).

It is hoped that this study presented a reasonable account of the use of the three pause fillers *um*, *erm*, and *er* in consideration of the two sociolinguistic variables of age and gender. Nevertheless, much more research is needed in this area. This would include evaluating other sociolinguistic variables such as social class and education, investigating other forms of pause fillers and also discourse markers, compiling results from different and more updated corpora, comparing corpora representing different geographical areas (e.g. UK vs US English), and analyzing other lexical or grammatical features of English using the same methodologies.

This study is not without its limitations. One of the biggest limitations is the lack of instances of the pause filler *um*, which meant no real inferences could be drawn from its analysis. Another limitation relates to using the somewhat dated BNC corpus, which was completed between 1991 and 1994 (Aston and Burnard, 1998). Thus, perhaps the results should be first compared with a more up-to-date British corpus. In addition, the BNC is only 10% spoken (about 10 million words); hence, perhaps a larger speech corpus would produce different results. One more drawback of this analysis concerning spoken language in the BNC is that it did not take into consideration the other part(s) of the conversations. It seems likely that the way a person speaks would vary depending on who he/she is interacting with. Nonetheless, it is hoped that despite these limitations, this study provided some insight into the research on pause fillers in relation to age and gender variables.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

[1] Amiridze, N., Davis, B. and Maclagan, M. (2010). *Fillers, pauses, and placeholders*. Amsterdam: John Benjamins Pub. Co.
[2] Aston, G. and Burnard, L. (1998). *The BNC handbook*. Edinburgh: Edinburgh University Press.
[3] BNCweb. (2013). *BNCweb (CQP-Edition)*. http://bncweb.lancs.ac.uk/
[4] Coates, J. (1986). *Women, men, and language*. London: Longman.
[5] Garside, R. (1987). The CLAWS Word-tagging System. In Garside R., Leech G. and Sampson G. (eds), *The Computational Analysis of English: A Corpus-based Approach*. London: Longman.
[6] Hoffmann, S., Smith, N., Lee, D. and Berglund-Prytz, Y. (2008). *Corpus linguistics with BNCweb - a practical guide*. Frankfurt am Main: Peter Lang.
[7] Holmes, J. (2008). *An introduction to sociolinguistics*. Harlow, England: Pearson Longman.
[8] Juan, S. (2010). *Why do we say 'um', 'er', or 'ah' when we hesitate in speaking?* http://www.theregister.co.uk/2006/05/06/the_odd_body_language_fillers
[9] Labov, W. (2010). *Principles of linguistic change*. Chichester: Wiley-Blackwell.
[10] Lakoff, R. (1975). *Language and woman's place*. New York: Harper & Row.
[11] Laserna, C., Seih, Y. and Pennebaker, J. (2014). Um . . . Who Like Says You Know: Filler Word Use as a Function of Age, Gender, and Personality. *Journal of Language and Social Psychology*, *33*(3): 328-338. https://doi.org/10.1177/0261927X14526993
[12] Mitzner, T. and Kemper, S. (2003). Oral and Written Language in Late Adulthood: Findings From the Nun Study. *Experimental Aging Research*, *29*(4): 457-474. https://doi.org/10.1080/03610730303698
[13] Tannen, D. (1990). *You just don't understand*. New York, NY: Morrow.
[14] Trudgill, P. (2001). *Sociolinguistics: An introduction to language and society*. 4th ed. London: Penguin.
[15] Yule, G. (2010). *The study of language*. 4th ed. Cambridge: Cambridge University Press.