NOVITAS-ROYAL
RESEARCH ON YOUTH AND LANGUAGE

2020, 14(2), 38-57
Novitas-ROYAL (Research on Youth and Language)
ISSN: 1307–4733

Research Article

A Corpus Analysis of Support Verb Constructions in British English with a Specific Focus on Sociolinguistic Variables

Ali Şükrü ÖZBAY[1]

[1]Ph.D., Karadeniz Technical University, Trabzon, TURKEY
alisukruozbay@gmail.com
ORCID: 0000-0002-3421-0650

Article information

Submission	01/05/2020	Revision received	31/08/2020
Acceptance	16/09/2020	Publication date	20/10/2020

Keywords: Support Verb Construction, Corpus, Gender, Age, Register

Abstract: English contains a considerable number of lexical combinations with various forms and labels, making it an interesting field of inquiry for researchers. The significance and popularity of support verb constructions (SVC) is that they are used largely by native speakers and include some of the most common words in English but seem to be problematic even for advanced learners. In this study, the British National Corpus (BNC) was used to investigate SVC patterns through sociolinguistic variables. The rationale in doing so is that using sociolinguistics variables through substantive corpus data may give us a better sense and understanding of the nature of the combinations. Whether there are any predictable tendencies between the SVCs and sociolinguistic dimensions was investigated to bring empirical evidence to the areas that merely defy simple generalizations. In total, 39 SVCs were examined, and they displayed similar frequencies for register categories in written and spoken parts. Male speakers and writers use a wider variety of SVCs and more frequently than the females, and SVCs were used more by people 60 years old and older and less frequently by people between the ages of 6 to 14. SVCs were used predominantly by male writers who were writing for mixed audiences, and there was a positive correlation between the age and SVC usage. A sharp increase was observed for SVCs as the age of the target audience increased. Finally, "years of experience," "exposure and previous education" and "familiarity" were found to be other contributing factors for the SVC usage.

Anahtar Sözcükler: Destek Eylem Yapısı, Derlem, Cinsiyet, Yaş, Tür

Sosyodilbilim Değişkenleri Işığında İngiliz İngilizcesinde Kullanılan Destek Eylem Yapılarının (DEY) Derlem Analizi

Özet: İngilizce, çeşitli biçimlerde ve niteliklerde önemli sayıda sözcük gruplarına sahip bir dildir ve bu durum dil araştırmacıları için önemli bir araştırma alanı olmuştur. Destek eylem yapılarının (DEY) önemi ve popüler olması, anadili İngilizce olan kişiler tarafından yoğun olarak kullanılması ve İngilizce'de en yaygın sözcüklerden bazılarını içermesi gibi durumlara bağlı olarak gelişmiştir. Bununla birlikte bu yapıların ileri düzeydeki öğrenenler için bile sorun oluşturduğu görülmektedir. Bu çalışmada, İngiliz Ulusal Derlemi (BNC), bazı Sosyodilbilim değişkenleri üzerinden DEY'nı araştırmak için kullanılmıştır. BNC verileri ve Sosyodilbilim değişkenlerin kullanılmasının, kelime gruplarının doğasını daha iyi anlamamıza yardımcı olabileceği düşünülmüştür. Çalışmada, DEY ile Sosyodilbilim değişkenleri arasında öngörülebilir kullanım eğilimleri olup olmadığı araştırılmıştır. Yine çalışma kapsamında toplam 39 DEY incelenmiş ve bu kelime gruplarının kullanılan derlemin yazılı ve sözlü kısımlarında benzer frekanslar gösterdikleri görülmüştür. Erkek konuşmacıların ve yazarların DEY kullanımı kadınlardan daha fazla ve sık olup, DEY'nın +60 aralığında daha fazla ve 6-14 aralığında daha az sıklıkla kullanıldıkları tespit edilmiştir. DEY ağırlıklı olarak karma kitleye yönelik erkek kullanıcılar tarafından üretilmiş olup ve yaş ile DEY kullanımı arasında pozitif bir korelasyon olduğu görürmüştür. Hedef kitlenin yaşı arttıkça DEY'nda da keskin bir artış gözlemlenmiştir. Son olarak, "deneyim" ve "maruz kalma ile önceki eğitim" ve "aşinalık" değişkenlerinin DEY'nın kullanımına katkıda bulunan diğer faktörler arasında yer aldığı görülmüştür.

To Cite This Article: Özbay, A. Ş. (2020). A corpus analysis of support verb constructions in British English with a specific focus on sociolinguistic variables. *Novitas-ROYAL (Research on Youth and Language)*, *14*(2), 38-57.

1. Introduction

It is a fact that speakers' language use varies in terms of gender, age, previous education, and levels of linguistic mastery, and many researchers have studied these sociolinguistic language-based gender differences extensively (Biber et al., 1999; Nation, 2004; Tao, 2007). In this study, this fact is further supported by difference and dominance theories. According to Uchida's (1992) difference theory, men and women, even those within the same social group, live in different or separate cultural worlds and, as a result, communicate differently. This is in line with the dominance theory, which argues that men and women live different lives in terms of cultural and linguistic aspects (O'Barr & Atkins, 1998, p. 4). Even in the same social class or speech community, there is always a possibility that male and female speakers may employ different linguistics forms, which is supported by Holmes' study (2001) of Amazon Indians' tribal language and how it changes from tribe to tribe and from males to females. The distinction in the linguistic features of male and female speakers range from pronunciation and morphology to vocabulary. Use of different words and word groups with the same meaning, preference for standard language, speech and maintaining relationship, intimacy, status and independence are some of the factors to consider so as to understand the phenomena (Climate, 1997; Holmes, 2001; Kaplan & Farrell, 1994; Leet-Peregrini, 1980; Tannen, 1990). Lakoff's (1975) proposed theories about women's language became a basis for this research. Lakoff's (1975) account of women's language and the set of assumptions about the female language use include "the use of hedges, polite forms, use of empty adjectives, correct grammar, direct quotation, special lexicon, modal constructions, intensifiers" (p. 6).

Roughly defined as "a succession of two or more words that must be used as an integral whole and not pieced together from its component parts" (Palmer, as cited in Kennedy, 2003, p. 468), multi word constructions (MWC) are largely used in various forms and lengths by native speakers who know that seeing a word in a context will bring about a natural expectation for the possible collocates immediately following the node words. Firth (1957) exemplified this combinational nature with the example of the adjective *dark* that collocates with *night* (p. 196). Perhaps, this and many other combinational constructs can best be rationalized with Sinclair's (1991) "idiom principle" model, which governs the choice of words by speakers and writers together with an "open-choice model" (p. 1). The model takes for granted that word combinations or fixed phrases possess a "meaning dimension" in a pre-determined semantic environment (Sinclair, 1991). From a more general dimension, the tendency towards frequent use of word combinations rather than individual words is popular, and the most obvious rationale for this is seen in the prefabricated chunks retained as a whole in the minds of the speakers and "just waiting to be recalled for use" (Lewis, 2000, p. 15).

SVCs can be considered one of the subgroups of a broader category of word combinations. They contain a verb and a noun such as "make an announcement" and "give permission" where the nouns possess the complete meaning, and the remaining verbs act as a complement with almost no meaning (Sinclair et al, 1990). Having different titles, such as "expanded predicates" (Algeo, 1995), "phrasal verbs" (Stein, 1991), "complex verbal structures" (Nickel, 1968), "stretched verb constructions" (Allerton, 2002) or "support verb constructions" (Danlos, 1992; Krenn, 2000), SVCs are used largely and frequently by native speakers, but they are problematic even for advanced learners, and even though the number of SVC is only a small portion of the words in the English language, they include the most common words in English (Sinclair et al., 1990). The criteria for multi word constructions to be counted as SVC are given below:

1. The nouns in SCVs are related to the verb (e.g., have bath - bathe, give permission-permit)
2. No indefinite article between the verb + noun combination (e.g., take action)
3. The noun may be a prepositional object (e.g., take something into consideration)
4. A verb and a noun should not be related (e.g., make an effort) (Labuhn, 2001).

The use of corpus data, on the other hand, for such sociolinguistic parameters is not a new phenomenon. A considerable number of studies based on corpus data have been conducted so far, and one of these was based on the amplifier distribution across different registers in British and American English by Biber et al. (1999, pp. 564-569). In another corpus-based sociolinguistic study of amplifiers in British English by Xiao and Tao (2007), which also inspired the present study, similar extralinguistic parameters on the basis of substantive corpus data were investigated, and the researchers found that the use of amplifiers were more frequent in speech, and there was a correlation between education and a higher frequency of amplifier use. Nesselhauf's (2003) exploratory study focused on the analysis of verb and noun collocations in free written productions by Germans who were English language learners that spoke at an advanced level. Nation (2004) compared the General Service List (West, 1953) and the Academic Word List (Coxhead, 2000) to three lists from the British National Corpus (BNC), only to find out that they contain almost the same words. Pearce (2008) examined the representation of men and women in the BNC by focusing on the collocational and grammatical behaviour of the two nouns "man" and "woman," while Liu (2012) aimed to examine SVC usage patterns by identifying the most frequently used MWCs of various types (e.g., idioms, lexical bundles, and phrasal/prepositional verbs) in general academic writing across academic writing sub-corpora of the Corpus of Contemporary American English (COCA) and the BNC. In another study, Panagiotidou (2015) used the data from the BNC corpus to investigate whether a limited number of linguistic features were conventionally connected to female speech or not in the case of "empty" adjectives. His data were not "supportive of sex differentiation in the use of empty adjectives and therefore contradicts their stereotypical association with female linguistic behaviour" (p. 7). This study seeks to fill a gap in the current literature on sociolinguistic variables likely to bring about certain male and female tendencies and usage-based patterns regarding SVCs through a detailed investigation into the contents of the BNC in terms of SVCs, which has the potential to constitute a norm data set for similar studies in the future.

2. Method

Ethical considerations in any research are critical and should be considered in order to preserve participants' rights and keep privacy in the data collected (Orb et al., 2001, p. 93). In all phases of this study, research and publication ethics are complied with. This study sought to address the popular usage-based patterns by exploring the distribution of 39 SVCs with a specific focus on the sociolinguistic variables such as gender, register and age from the BNC. The research questions are related to the use of SCVs for sociolinguistics variables as mirrored by the native language use in a considerably big representative corpus. The current study aims to answer the following three questions:

1. To what extent is gender an important variable as a quantification measure with SVCs usage as evidenced by BNC?
2. To what extent do register and age play a role in the usage of SVCs in corpus data (BNC)?
3. To what extent do audience gender and age play a role in the present usage patterns of SVCs?

The BNC, a general corpus not specifically restricted to any particular subject field, register or genre, were compiled to serve for academic linguistic research in the areas of lexicology, semantics/pragmatics, syntax, and morphology. The BNC includes 100 million-word samples of written and spoken English and represents both spoken and written British English. The written and spoken components included language samples that were compiled according to strict design criteria based on the three selection features: domain (subject field), time (within certain dates) and medium (book, periodical, etc.). In the written part of BNC, which constitutes 90% of the samples, there are various texts from many sources, including newspapers, journals, books, letters, essays, etc. In the spoken part, which constitutes 10% of the samples, there are various spoken language samples from various contexts (Burnard & Aston, 1998). The sampling frame in BNC is composed of "the language production of the population of British English speakers in the United Kingdom," and this data is represented by the language data in terms of several dimensions such as age, gender and social group. Especially in the "Lectures, talks, educational demonstrations", "News commentaries" and "Company talks and interviews" sections, the gender data is particularly emphasized. The "Age" dimension is also provided in BNC as part of the "author information" and the "demographic parameters to sample the population" data. Table 1 below shows figures for the transcribed material in BNC, grouped by their age and gender.

Table 1.

BNC data from the transcribed material in terms of gender and age (adapted from Burnard, 2007)

Respondents in BNC	**texts**	**w-units**	**%**	**s-units**	**%**
Unknown Gender	5	16245	0.38	2407	0.39
Male	73	1742222	41.14	248241	40.65
Female	75	2475488	58.46	359909	58.94
Age 6-14	26	267005	6.30	41036	6.72
Age 15-24	36	665358	15.71	97993	16.04
Age 25-34	29	853832	20.16	121752	19.94
Age 35-44	22	845153	19.96	126690	20.74
Age 45-59	20	963483	22.75	136530	22.36
Age 60+	20	639124	15.09	86556	14.17

w-units: written
s-units: spoken

Based on the BNC data, all the verb-noun combinations with *make, have, take, give and do* were extracted and listed separately. The number of the SVCs was limited only to the certain most frequently used verbs in de-lexical sense (Akimoto, 1989; Sinclair, 1990). The search was made both from the spoken and written registers. Acceptability criteria was applied from two dictionaries, these being the Oxford Advanced Learner's Dictionary (2000), and BBI Dictionary of English Word Combinations. Then, the selected combinations were checked for accuracy from the written part of the COCA since this corpus is bigger and more representative of the similar genres in BNC. Determining the accuracy of a combination required the correspondence of lexical elements, the number of the nouns, major determiners and the complementation.

3. Findings

3.1. Findings Related to Register Analysis

First of all, register differences between the spoken and written parts of BNC were examined for SVC variation. After a general framework of the comparison data related to the use of

SVCs across written and spoken registers in BNC was formulated, a more detailed examination of the variations was made. In Table 2, the frequencies of the SVCs normalized to a common base of a million, and their log-likelihood (LL) scores are given. Based on the normalized data, it can be seen from the table that there is a considerable difference between the spoken and written registers of SVCs (165.63 and 117.39 instances per million words respectively), which was also validated by the LL scores with an LL score of 8.26 for 1 degree of freedom, $p < 0.001$.

It appeared that individual SVCs were slightly different in terms of written and spoken registers demonstrated in the last column of Table 2, providing the LL scores arranged in a descending order for every single category. Only one SVC, *make (a) comment,* out of 39 yielded a significant difference in its frequency in speech and writing and was observed to be the most prevalent SVC in spoken section (15.08). On the other hand, *give(a) presentation*, *give (an) overview (of)*, *make (an) assumption*, and *make (a) prediction* were used slightly more in spoken discourse modes, but the contrast between the two registers was very small. The analysis of two discourse modes showed that the frequency and variety of SVCs in written texts were significantly higher compared to the spoken register. However, the spoken and written registers are not one piece since each discourse mode differentiates from one another.

Table 2.

SVCs in spoken and written BNC

Category	SVC	Written	Spoken	LL Score
More frequent in writing	have access	15.88	7.2	3.35
	give access (to)	4.76	0.86	2.98
	make (a) contribution	13.14	6.92	1.96
	give (sb an) impression	10.26	5.19	1.69
	give priority (to)	5.03	1.73	1.68
	take precedence	2.42	0.38	1.66
	make (an) impact	5.19	1.92	1.56
	have limitations	1.48	0.10	1.44
	make (a) transition	1.54	0.29	0.94
	make (an) impression	3.79	1.73	0.79
	make (a) statement	11.58	8.17	0.59
	take initiative	3.25	1.63	0.55
	take (an) approach	2.10	0.86	0.54
	give consent	2.09	0.96	0.43
	make adjustments	2.09	0.96	0.43
	make (a) living	3.79	2.21	0.42
	make arrangements	7.24	5.00	0.41
	have (a) tendency	3.30	1.92	0.37
	make(a) judgment	4.64	2.98	0.36
	make (an) assessment	2.10	1.06	0.35
	give emphasis	1.23	0.67	0.17
	give (sb) treatment	1.96	1.25	0.16
	make contact	6.98	5.57	0.16
	give (an) indication (of)	5.44	4.23	0.15
	have (an) obligation	2.48	1.83	0.10
	make (a) recommendation	3.94	3.17	0.08
	give (an) explanation	1.76	1.34	0.06
	make (an) observation	1.81	1.44	0.04
	give consideration	2.72	2.40	0.02
	give guidance	2.24	1.92	0.02

	give information	10.02	9.61	0.01
More frequent in speech	take into consideration	2.81	4.32	0.32
	make (an) argument	0.53	0.77	0.04
	make provision	5.60	5.96	0.01
	give (a) presentation	0.65	0.67	0.00
	give (an) overview (of)	0.61	0.67	0.00
	make (an) assumption	2.95	2.98	0.00
	make (a) prediction	1.43	1.44	0.00
Significantly more frequent in speech	make (a) comment	4.80	15.08	5.58
Total		165.63	117.39	8.26

LL: Log-likelihood

Figure 1 illustrates the distribution of SVCs across written and spoken registers. It can be seen from the graphic that the written register in BNC includes more tokens of SVCs than spoken discourse mode. The frequency of W-misc. is the most remarkable one within both its own register and the spoken register. There are some other varieties dispersed in each discourse modes; for instance, SVCs are more frequent in the context governed speech (S-cg) than in demographically sampled conversations (S-demo).

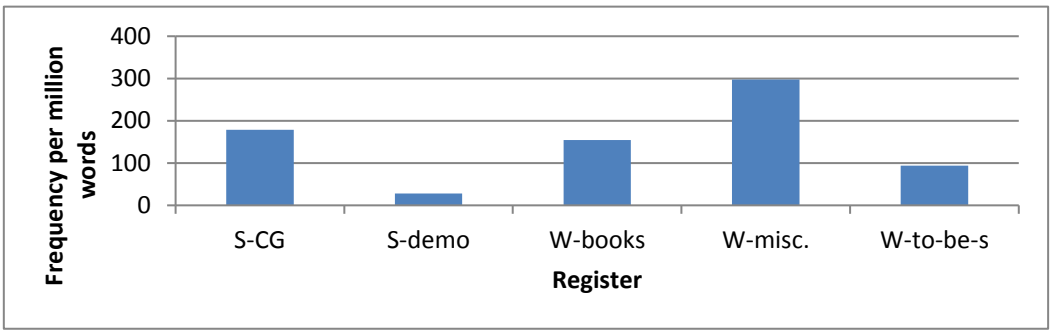

Figure 1. Distribution of SVCs across spoken and written registers

When the written register was considered, it also appeared that books, periodicals (W-books) and the written miscellaneous (W-misc.) consist of statistically more SVCs than written-to-be-spoken register. On the other hand, the comparison of two interaction types in speech (122.26 and 116.34 instances per million words for monologue and dialogue respectively) yielded no significant results with an LL score of 0.15 for 1 d.f, $p = 0.087$.

3.2. Findings Related to Gender Analysis

In this part, the relationships between SVCs and the gender modes in BNC corpus were examined. The findings were categorized according to the written and the spoken modes for both genders. In total, 39 SVCs were examined, and the difference between male and female speakers was statistically insignificant. LL scores under the spoken and written modes and their sub-distinctions between the males and the females remained 1.051025 for the spoken data and 0.344358 for the written data. It was obvious that written mode presented a rather limited variety of SVC uses. The biggest difference in LL scores was seen with the *make arrangement* construction with a LL score of 1.47. This also indicated that there was no significant difference in the use of the SVCs between male and female writers. On the other hand, the data with the spoken form showed a significant difference between the two genders in their use of the *make comment* construction with a LL score of 4.72. Typical male and female usages of the most common SVCs were also found. When LL scores were examined in terms

of speech and writing, the data show males use a larger variety of SVCs and more frequently than females.

Table 3.

Frequent SVCs in speech and writing between male and female users.

Gender	Register	SVCs
Male	Used frequently in speech and writing	have access, give (sb an) impression, make a contribution, make a statement, make contact, make a comment, give information, give (an) indication (of)
	Used frequently in speech only	make a contribution, make a statement, make a comment, give information, make provision, have access
Female	Used frequently in speech and writing	give (sb) an impression, have access, make (a) contribution, make arrangements, give information, make a comment
	Used frequently in writing only	give (sb) an impression, make arrangements, have access, give information

Table 3 contains the most frequent male and female SVCs in speech and writing. Based on the table, it is possible to argue that there is rather a limited variety of SVCs used in both modes and even fewer variety of SVCs with speech and writing separately. It is now clear that males use SVCs more frequently than females in both spoken and written modes. Though the most preferred SVCs used by the males and females are alike, the clear differences between their frequencies of use remain noticeable. For example, *have access* was used by both males and females, but males did so more frequently. When the frequencies of both categories were combined, males had a frequency score of 461, and females had a frequency score of 189. Genre variations in the BNC can shed light on the ways to see the different usages by males and females for it explains the preferred genres. Lee (2001) encoded BNC with 70 genres (24 speaking and 46 writing genres) to make it easier to find the differences between the male and female usage patterns. To simplify the comparison of both genders' usages of SVCs, all the sub-genres were brought together into the upper level genre. In Figure 2, some SVC genres were not used by both gender groups and therefore were omitted. It can be clearly seen that males use SVCs more frequently than females, and this continues with most types despite the fact that the differences between the two genders are small.

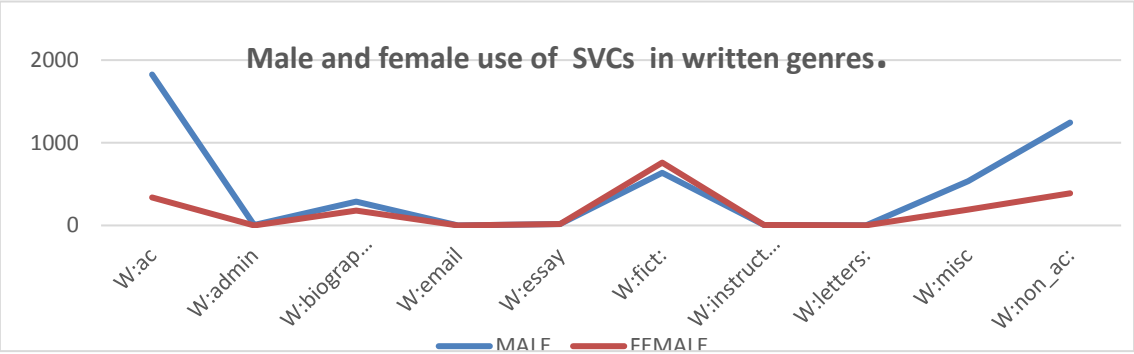

Figure 2. Male and female use of SVCs in written genres

The categories used in Figure 2 are the categories used in the BNC. The categories of writing are academic, administrative, biography, email, essay, fiction, instructional, letters, miscellaneous (i.e., brochures and advertising leaflets) and non-academic. The scale is also designed to show the numbers for frequency per million words. This dominance mainly is the result of two genres, the academic and the non-academic, which suggest that males and females both demonstrate different preferences for individual SVCs. This supports the findings reported in some earlier studies like Stenström (1999) though his study does not involve all genre types.

3.3. Findings Related to Age Analysis

The study also examined the use of SVCs by females and males according to their ages. "Age" is one of the variants affecting the use of language over time and reflects crucial differences between the generations. Our language use is affected by our age, and our age affects the language we use. Age and generation specific language use have created a significant area in sociolinguistic research (e.g., Buchstaller, 2006; Cheshire 1987; Coupland & Coupland, 2001) because there are "pronounced differences in language use over the life-span" (Pennebaker et al., 2003, p. 556). "Age" range in this study is from 6 to 60+. The effect of age on the use of SVCs in speech and writing were examined using the data from spoken and written BNC. Distribution of SVCs across age groups in speech and writing and cross tabulation of age and gender for written genres were examined to reveal the differences in the use of language.

Table 4.

Distribution of SVCs across age groups

Age group	Spoken		Written	
	Per M. words	Rank	Per M. words	Rank
6–14	0.53	8	4.74	11
15–24	0.73	17	3.59	76
23–34	2.56	112	3.31	293
35–44	2.65	111	3.3	885
45–59	2.38	152	7	1047
60+	1.31	58	3.71	806

M: million

Table 4 shows the distribution of SVCs across user age groups. In spoken BNC, the age range of 6–14 are the least frequent users of SVCs (0.53), while the 35–44 (2.65) and 23–34 (2.56) age ranges are the most frequent users of SVCs. In the written BNC, children aged 6–14 are found to be the most frequent users of SVC (4.74). In speech, the second least frequent users of SVCs were in the group of people aged 15–24 (0.73); however, this group frequently used SVCs in writing (3.59). The most striking difference between the spoken and written parts seems to be in the 6–14 age group. Also, those aged 60+ use SVCs more frequently in writing (4.03) than in speech (1.31). The 60 and above group is the third group to show a wide gap in the SVC use between written and spoken genres. Those aged 23–34, 35–44, 45–59 seemed to be more consistent in their use of SVCs in speech and writing. The 45–59 age group used more SVCs in writing than in speech when compared to the 23–34 and 35–44 age groups. Based on the table, it seems that the use of SVCs in speech is increasing throughout the generations up to the 45–59 age group. Then, it decreases to 1.31 in the 60+ age group, while it is 2.38 for those aged 45–59. There is not a huge gap between generations

in written BNC, but the 6–14 age group, and the 60+ age group outnumber the other generations in written BNC.

An explanation for this might be inferred from a cross tabulation of age and gender for written genres to explore the reason why the 6–14 age group are the most frequent users of SVC in writing. As indicated in Table 5, female writers contribute to the high frequency for this age group with a frequency of 5.19.

Table 5.

Cross tabulation of age and gender for written genres

Age group	Female	Per M words	Male	Per M words	Female /male ratio
6–14	Female	5.19	Male	0.00	5.19
15–24	Female	4.16	Male	3.51	1.18
23–34	Female	3.44	Male	2.71	1.26
35–44	Female	2.76	Male	3.53	0.78
45–49	Female	2.66	Male	4.00	0.66
60+	Female	2.53	Male	4.58	0.55

M: million

While female writers contribute to the high frequency of SVCs noticeably, male writers do not, and there is a striking difference in the use of SVCs between males and females in some age groups. Females under 35 years old contribute a considerably larger proportion of SVC use, but males aged 35 and older show a considerably higher frequency than females in this age range, and male SVC usage rises as they get older. People age 60 and above are similar to those aged 6–14 in that they use SVCs more frequently in writing. While female writers contribute to the most frequent use of SVCs in the 6–14 age group, male writers contribute to a considerably larger proportion of SVC use in the 60+ age group. The most noticeable difference in the writings of people 60 and above is that the contributors are males rather than females. Female contribution is 2.53, while it is 4.58 for male writers. Except for children aged 6–14, females are the most frequent users of SVCs in writing when compared to the spoken data.

Age seems to be an important factor in the use of SVCs across age groups. According to data, the lowest frequency users of SVCs in speech are those aged 6–14 and 15–25, while the most frequent users are people aged 35–44 in spoken BNC. It can also be argued that the use of SVCs increases in speech as the users get older. In written BNC, the most frequent users of SVCs are children aged 6–14, and the second most frequent users are people aged 60 and above. The other age groups also use more SVCs in writing than in speech. It is also clear that females use fewer SVC in writings as they get older, while the use of SVCs by males increases with age.

It can be seen in Table 6 that for authors from the age groups 6–14 and 60+, the use of SVCs in spoken BNC is quite infrequent. The 6–14 age group and the 15–24 age groups, do not frequently use SVCs, and the authors' percentage follows a downward trend (10%). In the 25–34 age group, there is only one SVCs use, and there is only *make provision* sample above 10% percentage level with the 35–44 group. *Make (a) comment* is the only SVCs in the 45–59 age group, and there is no SVCs pattern for the 60 and above age group that is above 10%. Contrary to the "spoken high frequency group," the "spoken low frequency group" (below 10%) seems to present several SVC patterns. In the written BNC, the number of

words for both categories (below 10% and above 10%) are not equally divided. For "above 10%," there are only 6 and 2 SVCs in the age ranges from 6-14 to 15-24 respectively. For the 25-34, 35-44 and 45-49 and 60+ age groups, there are 4, 1, 2 and 4 SVC patterns respectively. When it is compared in terms of "above 10%" category, it can be argued that written BNC contain more SVC patterns.

Table 6.

High and low frequency support verb constructions by age groups

Age Group	Per M. word	Spoken BNC	Written BNC
6–14	>10		Make (a) judgment, make (an) impression, give (an) explanation, make (an) impact, give (sb an) impression, take (an) approach
	<10	Make (a) comment, make contact, give consent, give information, make (a) judgment, give (sb an) impression	Give consideration, give consent, give (an) overview (of)
15–24	>10		Make (a) contribution, give (sb an) impression
	<10	Make adjustments, make arrangement, make a living, give priority (to), give (a) presentation, have access	Make arrangements, make provision, make (an) argument, give consideration, give emphasis
25–34	>10	Give information	Give information, give (sb an) impression, have access
	<10	Make (a) contribution, make (a) recommendation, make (a) prediction, make (a) transition, give guidance	Make adjustments, make (an) assumption, make (a) prediction, make (a) living
35–44	>10	make provision	Give (sb an) impression
	<10	take (an) approach, make (an) argument, have access, take (into) consideration	Make (an) argument, make provision, make (a) transition, give emphasis
45–59	>10	Make (a) comment	Make (a) contribution, give (sb an) impression
	<10	Give (sb) treatment, give priority (to), give (a) presentation, take initiative	Give (sb a) treatment, take precedence, have access, have limitations
60+	>10	-	Make (an) arrangement, make (a) contribution, make contact, give (sb an) impression
	<10	Make adjustment, make (an) arrangement, make (a) comment, give (an) indication (of)	Have (a) tendency, have limitations, have (an) obligation, take precedence

Table 7 shows the SVC frequency between the ages of 6–14. A striking difference is seen for *make (an) impression* with its frequency of 33.58 % in written BNC, but no usage was found in the spoken BNC. A similar difference is found with *make (an) impact* which has a frequency rate of 16.79 in written BNC but zero in the spoken part. Interestingly enough, while *give information* has a frequency rate of 7.79 in the spoken BNC, but zero in the written part. Besides, *give consent* and *give information* have frequency rates of 2.6 in the spoken BNC but

again zero in the written part. The most striking one is *give (sb. an) impression* with its frequency of 50.37 in written BNC but only 2.6 in the spoken part. Frequency rates between 15 to 14 are very similar for *make arrangement* in both parts. The frequency of *make (a) contribution* is 25.8 in the written part but zero in the spoken. Written frequencies are high with *make contact*, *make a living*, and *make a comment*, and spoken frequencies are high with *take into consideration* and *give information* in this age group.

Table 7.

High and low frequency support verb constructions by age groups

SVC		Spoken BNC	Written BNC
6–14	Make judgment	2.6	16.79
	Make (an) impression	0	33.58
	Make (an) impact	0	16.79
	Give information	7.79	0
	Give (sb an) impression	2.6	50.37
	Give consent	2.6	0
15–24	Make arrangement	6.73	7.37
	Make (a) contribution	0	25.8
	Make contact	1.68	5.53
	Make (a) living	0	7.37
	Make (a) comment	3.36	9.22
	Give information	3.36	1.84
	Give (sb an) impression	0	25.8
	Take into consideration	5.05	1.84
	Give priority (to)	1.68	0
	Take (an) approach	0	0
25–34	Make (a) recommendation	0	0.44
	Make contact	8.92	7.94
	Give information	13.39	10.15
	Give (sb an) impression	3.57	14.56
	Give (a) presentation	0	0
35–44	Make provision	13.94	2.53
	Make (a) comment	8.37	4.31
	Make (a) statement	7.44	7.73
	Give (sb an) impression	3.72	1.64
	Have access	6.51	12.64
45–59	Make (a) comment	20.75	5.53
	Make arrangement	4.27	8.44
	Have (a) tendency	1.22	4.01
	Give priority (to)	0	4.29
	Give (sb an) impression	3.05	12.31
60+	Make arrangement	7.92	10.73
	Make (a) contribution	2.64	10.14
	Give (sb an) impression	2.64	17.56
	Give (sb) treatment	4.40	1.37

Frequency rates between 25–34 present almost similar frequencies with the exception of *give (sb an) impression* which is four times higher in the written part. Frequency rates between 35–

44 show variations between the two modes. While *make provision* is very high in spoken part with this age group, *give (sb an) impression* and *have access* are very high in the written part. There is an equal distribution for *make a statement* in both modes with a slightly higher percentage for *make a comment* in the spoken part. The 45–49 age group used *make a comment* patterns with a high frequency of 20.75 in the spoken part but 5.53 in the written part. Another pattern is *have a tendency* with a frequency of 8.44 in the written part and 4.27 in the spoken part. The pattern *give priority* is non-existent in the spoken part but 4.29 in the written BNC. The pattern *give (sb an) impression* is seen dominantly in the written part with a frequency of 12.31 as opposed to spoken part having only a 3.05 percentage. Finally, within the 60+ group, the high frequency of SVC patterns is seen in the written part with the exception of *give (sb) treatment* used 4.40 in the spoken BNC.

3.4. Findings Related to Audience Gender

This part of the study deals with analyzing the use of SVC patterns in writing in terms of gender. The target audience is grouped into three categories as male, female and mixed in the written part of BNC. According to our data SVCs were used predominantly by male writers intended for a mixed audience with the LL score of 89.47 and a total usage of 84,20512821. Among the target SVCs, *make contact* is used mostly (LL 151.40 d.f., $p < 0.001$) by male authors for a male audience. Male authors use this SVC much more extensively than that of women in the written register. This is a big gap, and there may be a need to conduct more research to understand the reasons. The least used SVC is *give overview* intended for a mixed audience type (LL 8.2 for 1 d.f., $p < 0.001$). It is quite interesting that among 39 SVCs under consideration, there are 9 SVCs never used by male authors with a male audience. This deficit needs to be investigated with the help of spoken data, but the study shows that SVCs are significantly more frequent in writing intended for male than female audience or mixed audience by male and female authors (413.55, 202.28 and 302.28 occurrences per million words respectively, LL = 50.73 for 1d.f., $p < 0.001$). This fact is probably because not only male authors, but also female authors use SVCs much more frequently in writing with an intended male audience, as shown in cross tabulation in Figure 3.

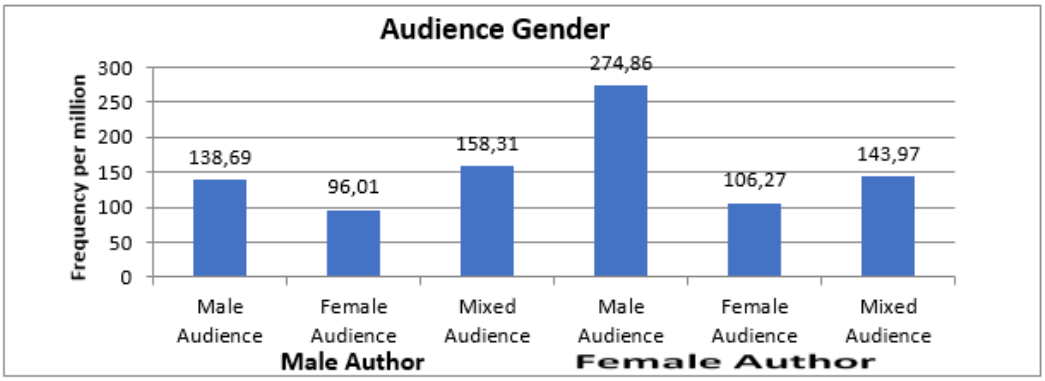

Figure 3. Cross tabulation of author gender and audience type

However, we cannot be sure that all these findings can be applied to spoken English since audience gender is not encoded in the spoken BNC. In terms of target SVCs individually, different items are used for male and female audience in written BNC. From the aspect of SVCs, written material intended for male, female and mixed audiences depicts a widespread categorization. Male and female authors have different tendencies towards using SVC. According to Sinclair (1991), two notions named "open-choice principle and idiom

principle" determine the choice of word. In these two models, the user of language has the opportunity of the choice of words. In this study, it is seen that male authors pay more attention to idiom principle compared to female ones.

3.5. Findings Related to Audience Age

There are four main categories for audience age in written part of BNC, which are child, teenager, adult, and any, but the category any was excluded from our analysis. The number of hits and frequencies per million for written text analyses suggest that the use of SVCs is directly proportional to age of the intended audience. The findings given in Table 7 reveal that there is little or no use of SVCs for children for the most part. The teenager group, on the other hand, received hits almost from all the sample SVCs with the exception of eight SVCs that received no hits. The adult group received hits from all the SVC samples with variations in percentages. Such SCV samples as *make (a) contribution*, *make (a) statement*, *give (sb an) impression*, *have access*, *make (an) impact* and *make (a) comment* received the highest percentages in the adult category. It may be true to say that the number of hits may not give us an accurate picture of the situation, but it does give a general idea about the usage patterns and types of construction used for different audience age groups in BNC irrespective of their rare occurrences.

Table 8.

Total hits and frequencies per million for audience age

	Child		Teenager		Adult	
	Hit	Frequency per Million	Hit	Frequency per Million	Hit	Frequency per Million
make adjustments	1	1.11	4	2.24	176	2.15
make arrangements	6	6.64	8	4.48	596	7.27
make (an) assumption	0	0.00	1	0.56	257	3.14
make (a) contribution	1	1.11	4	2.24	1104	13.48
make contact	6	6.64	14	7.84	567	6.92
make provision	0	0.00	0	0.00	488	5.96
make (a) statement	2	2.21	5	2.80	978	11.94
make (a) recommendation	0	0.00	0	0.00	330	4.03
make (a) prediction	0	0.00	4	2.24	118	1.44
make (a) living	7	7.75	3	1.68	307	3.75
make (an) assessment	0	0.00	0	0.00	184	2.25
make (a) comment	2	2.21	4	2.24	396	4.83
make (an) argument	0	0.00	1	0.56	46	0.56
make (a) transition	0	0.00	0	0.00	130	1.59
make(a) judgment	0	0.00	12	6.72	389	4.75
make (an) observation	1	1.11	5	2.80	153	1.87
make (an) impression	2	2.21	8	4.48	302	3.69
make (an) impact	0	0.00	4	2.24	409	4.99
give access (to)	2	2.21	5	2.80	396	4.83
give consent	0	0.00	1	0.56	179	2.18
give consideration	0	0.00	1	0.56	225	2.75
give emphasis	0	0.00	1	0.56	105	1.28

give priority (to)	0	0.00	1	0.56	426	5.20
give guidance	0	0.00	1	0.56	194	2.37
give information	2	2.21	8	4.48	848	10.35
give (an) explanation	0	0.00	2	1.12	151	1.84
give (an) indication (of)	0	0.00	3	1.68	458	5.59
give (an) overview (of)	0	0.00	0	0.00	54	0.66
give (a) presentation	0	0.00	1	0.56	50	0.61
give (sb an) impression	6	6.64	26	14.56	831	10.14
give (sb) treatment	1	1.11	0	0.00	164	2.00
take initiative	1	1.11	2	1.12	276	3.37
take precedence	2	2.21	3	1.68	202	2.47
take (an) approach	0	0.00	0	0.00	179	2.18
take into consideration	0	0.00	3	1.68	226	2.76
have access	2	2.21	9	5.04	1346	16.43
have (an) obligation	0	0.00	2	1.12	210	2.56
have (a) tendency	0	0.00	4	2.24	276	3.37
have limitations	0	0.00	0	0.00	127	1.55

The average frequencies per million in Table 9 show us a similar picture in that the frequencies per million are 1.25, 2.15 and 4.34 for child, teenager and adult audience groups respectively, and these LL scores indicate that there is a significant difference between the adult audience group and other groups. The sharp increase in the use SVCs as the age of the target audience increases needs further explanation. One plausible explanation for the increase of use of the SVCs according to audience age may be the awareness that years of experience and exposure is necessary for learners of the language to acquire these combinations as Carter (1987) suggests.

Table 9.

Distribution of SVCs across three audience age groups

Age group	Hit	Average Frequency (per million)
Child	44	1.25
Teenager	150	2.15
Adult	13853	4.34

However, we must also be cautious to attribute these results to age alone; education may play a more dominant role than age. The results of the analysis of the 39 SVCs in terms of education level in the previous section reveal that the use of SVCs is more frequent in the higher education level. We are more likely to see these words in academic speeches and writings such as textbooks or lectures, and accordingly, the amount of exposure to academic language may be the prime factor in increasing the familiarity with SVCs and the use of them. In other words, the quality of the input, rather than the quantity may play a more fundamental role.

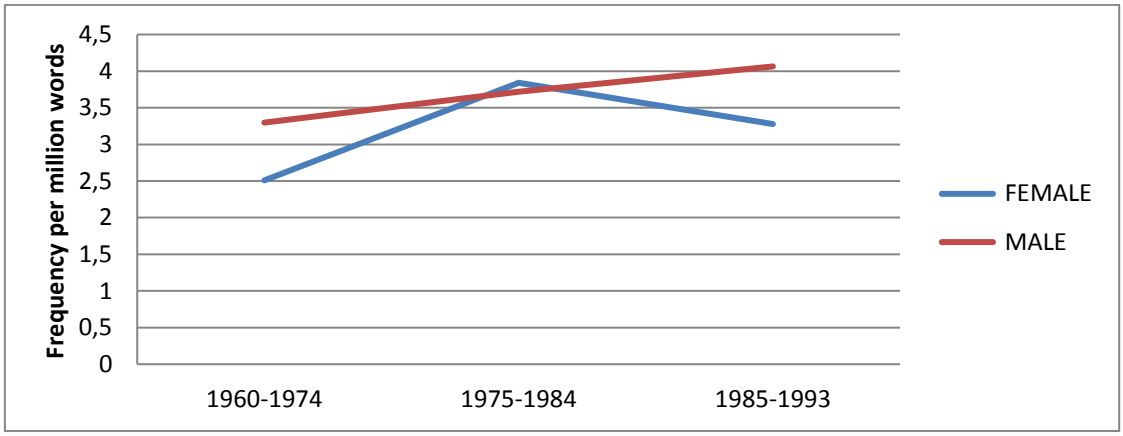

Figure 4. SVCs used by male and female authors across three periods

The change in the number of SVCs among male and female authors between the different periods are obvious. A cross tabulation of author gender with publication proves that there is a great difference between male and female authors. As can be seen in Figure 4, while the female use of SVCs rose and declined steadily across the three periods, the male use bounced in the second period between 3.5 and 4, and after this rise, the number decreased. However, it remained higher than the first period. The male use continued to increase during three periods.

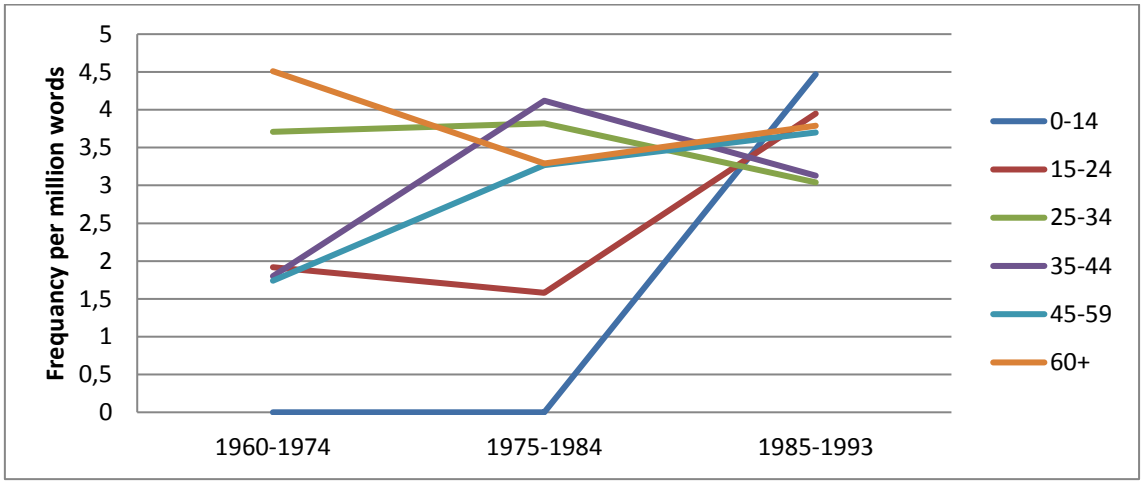

Figure 5. SVCs used by different age groups in three periods

In Figure 5, there is a mix of decreasing and increasing numbers of the SVC used by different age groups. For instance, while there is a sharp increase after 1975 in the use between ages 6–14, there is also a sharp decrease in the use by people over 60 years between 1960 and 1974. The increase with 35–44 age groups continues up to 1975, and then it takes a steady downward trend towards 1993. The possible reasons behind this mix of decreases and increases may depend on factors other than age and gender, and it seems that a larger scale study is needed in order to investigate the validity of other factors at work behind these mixes of decreases and increases in the amounts of SVCs.

4. Discussion and Conclusion

This paper investigated the use of SCVs by native speakers with a focus on such factors as register, gender, age, audience gender and audience age. The analysis was done by referring to the BNC, being the most representative and sizeable corpus of British English today. The results of the analysis in terms of various sociolinguistic variables seem to be complex in nature. Part of this complexity results from some tendencies which are accepted as normal, considering the fact that the use of SVCs is more common in speech compared to writing, and that people with more education are likely to use words with their associative pairs together.

The first analysis was done for register variations in terms of SCV usage. The difference between the spoken and written registers of SVCs 165.63 and 117.39 instances per million words respectively, and a few SVC patterns were observed to be prevalent in the spoken part. Others, however, displayed almost similar frequencies in terms of register categories with higher number of SCVs used in the written register. This finding is concurrent with the Sinclair's (1991) argument that almost half of the English lexicon in academic language is used according to the idiomatic principles in written academic language.

The second analysis was based on gender differences between SVC usages. The findings of the analysis were categorized according to the written and the spoken modes for both genders. Based on this categorization, it seemed that the written mode presented a limited variety of SVC use, but the spoken mode presented a significant difference for some SVCs. In general terms, it is evident that male speakers' and writers' use of the SVCs is generally greater in variety and more frequent than the female use. In spite of the fact that the most common SVCs used by the males and females were similar, their frequencies seemed different. As supported by Stenström (1999), academic and non-academic sections of BNC were the dominant sources behind the preferences of individual SVCs use.

The third analysis was based on age as a sociolinguistic construct. Age is an important variable since research on generation specific language use and existence of huge variations in language use of people in different ages has become very popular in sociolinguistic research (e.g., Cheshire, 1987; Coupland & Coupland, 2001; Buchstaller, 2006; Pennebaker et al., 2003). There were multiple age ranges with the youngest group being 6–14 and the oldest group being 60+ in the study, and female and male use of SVC with a focus on age were investigated to explore the differences in the use of language. In the spoken part of BNC, the category of 6–14 used SVC the least frequently, while the 35–44 and 23–34 categories were the most frequent users. The underuse of SVC by children below 15 might be explained from a developmental perspective, but it falls short to explain their overuse of SVC in written texts. In the written part of BNC, the 45–49 age groups were found to be the most frequent users of SVC. This part shows that age is an important factor in using SVC. According to the data, the most frequent SVC used by all age groups is *give (sb an) impression*; however, the results show that its frequency in the age range 6–14 outnumbers its usage in the age range 60+, besides it is used more in the written part of the BNC. For all the SVCs, the usage of SVCs is least frequent in the age range 6–14 and most frequent in the 60+. The data shows that use of SVCs increases with authors' age. The underlying reason, perhaps, is that authors do not have much SVC in their inventory, and as authors get older and older, they learn or get these SVCs into their inventory via education and start to use them, which means education can also be an affecting factor.

The final group of analyses were based on audience gender and age. There were male, female and mixed audiences in the written part of the BNC, and it was found that SVCs were used predominantly by male writers intended for a mixed audience type. It was also seen that the written parts of BNC depicted a widespread categorization in terms of all three categories with various levels of tendencies toward more idiomatic usage patterns on the parts of the male authors. It is now understood that men and women differ in terms of their communicative behavior (Coates, 1989).

When it comes to the gender difference in terms of the speakers addressing an audience of male dominance, the findings demonstrate some aspects of gender-based speech differences explained by many scholars starting with Lakoff (1975). It can be assumed that women use SVCs more frequently while addressing a male dominant audience can be the result of the effort by women to maintain their subordinate position compared to men's dominance preserved through their linguistic behavior (White, 2003). That is to say, women addressing men adjust their message using more SVCs as an indication of power. Tannen (1990) says 'If women speak and hear a language of connection and intimacy,' a clash of conversation styles can occur when confronted with a men's language concerned with status and independence. According to the dominance approach (Lakoff, 1975) language differences are a reflection of traditional social roles of men's dominance and women's sub-ordinance (White, 2003). According to Lakoff, both hedging and boosting modifiers show a women's lack of power in a mixed-sex interaction (White, 2003). Considering this, it would not be wrong to think that women addressing a male dominant audience use more SVCs in order to compensate for this approach to their status within society. While the hedges' lack of assertiveness is apparent, boosters, she claims, intensify the force of a statement with the assumption that a woman would not be taken seriously otherwise (White, 2003).

On the other hand, the difference approach (Montgomery, 1995) asserts that sex differences are better attributed to relationship oriented contrasting characteristics of men and women rather than power and status (White, 2003). For men the focus is on sharing information, while women value the interaction process. Men and women possess different interactive styles, as they typically acquire their communicative competence at an early age in same-sex groups. There have so far been many scholars working on different linguistic aspects of women's speech focusing on hyper-correct grammar (Cameron & Coates 1989; Coates, 1986; Trudgill, 1983), tag questions (Dubois & Crouch, 1975; Cameron et al., 1989; Holmes, 1986), and commands (Goodwin, 1980; Holmes, 2001;Tannen, 1990). However, literature on the use of SVCs is rather limited, making our study significant to the field. Language used by men and women differ linguistically, which does not actually result from an imbalance between their power relations and status in the society. Maltz and Borker (1983, p. 200) put it saying, "American men and women come from different sociolinguistic subcultures, having learned to do different things with words in a conversation." In contrast to Lakoff, believing that women are deficient of power in their speech, the dual-culture approach believes that men and women are simply different; one is not inferior to the other. From this point of view, women's more frequent use of SVCs can account for an effort of building up communicative bridges across a woman's linguistic realm.

The analysis of the audience age with four main categories was the final part of the study, and it was found that there was a positive correlation between age and SVC use. A sharp increase was observed in the number of SVCs as the age of the target audience increased. As stated before, one reason for this may be the years of experience and exposure and possibly the education since higher education was also a contributing factor for the SVC use. It is also

the case that SVCs were very frequent in academic language, and the users may have been familiar with such combinations thanks to rich input. Age and education level factors may overlap to some extent, and it is difficult to understand the extent of the influence that each one exerts on the use SVCs as people who have spent more years in education will probably be the ones that are older. When they are writing for a specific audience age group, the writers are probably well-aware of these two factors, and they use SVCs accordingly for a given target audience.

In conclusion, based on the data analyzed, it is possible to conclude that a limited number of SVC patterns were used in both spoken and written parts of BNC, indicating that there are register-specific repeated patterns in English and EFL learners need to equip themselves with such patterned data in order to do well in academic writing. Another conclusion may be that male usage of this patterned data seemed somewhat more frequent, suggesting that male speakers and writers tend to use patterned structures as a sign of their communicative behavior compared to female usage. It is also the case that the use of patterned structures slowly but gradually increases with the age of the speakers and writers based on the education and maturity. Dominance, power, social status, and literacy were the other factors at work behind the SVC usages.

The researcher argues that a thorough analysis will give us authentic and reliable data related to the nature of SCVs in various dimensions or sociolinguistics constructs. This understanding is likely to be taken as a norm related to the nature of combinational data in English. It is also the case that future corpus-based contrastive interlanguage investigations related to the comparison data to be obtained from the analysis of lexical structures and multi word combinations can be shaped by the findings of this study. Further studies comparing age and education level variables may give us a better grasp of the issue under inspection.

Ethical Issues

The author(s) confirm(s) that the study does not need ethics committee approval according to the research integrity rules in their country.

References

Akimoto, M. (1989). *A study of verbo-nominal structures in English*. Tokyo: Shinozaki Shorin.

Algeo, J. (1995). Having a look at the expanded predicate. In B. Aarts & C. F. Meyer (Eds.), *The verb in contemporary English*, (pp. 203-217), Cambridge: Cambridge University Press.

Allerton, D. J. (2002). *Stretched verb constructions in English*. New York: Routledge.

Biber, D., Johansson S., Leech G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.

Buchstaller, I. (2006). Diagnostics of age-graded linguistic behaviour. *Journal of Sociolinguistics*, *10*(1), 3-30.

Burnard, L. (2007). Reference guide for the British national corpus (XML Edition). Retrieved from http://www.natcorp.ox.ac.uk/docs/URG.xml

Burnard, L., & Aston, G. (1998). *The BNC handbook: Exploring the British national corpus*. Edinburgh: Edinburgh University Press.

Cameron, D., & Coates, J. (1989). Some problems in the sociolinguistic explanation of sex differences. In J. Coates & D. Cameron (Eds.), *Women in their speech communities: New perspectives on language and sex*, (pp. 13-26). New York: Longman.

Cameron, D., McAlinden, F., & O'Leary, K. (1989). Lakoff in context: The social and linguistic functions of tag questions. In Cameron, D. & Coates, J. (Eds.), *Women in their speech communities* (pp. 74-93). London: Longman.

Carter, R. (1987). *Vocabulary: Applied linguistics perspectives*. London: Routledge.

Climate, C. (1997). *Men and women talking: The differential use of speech and language by gender*. London: Routledge

Coates, J. (1986). *Women, men and language: A sociolinguistic account of sex differences in language*. New York: Longman.

Coates, J. (1989). Gossip revisited: language in all-female groups. In Cameron, D. & Coates, J. (Eds.) *Women in their speech communities* (pp. 94-122). New York: Longman.

Coupland, N., & Coupland, J. (2001). Language, ageing, and ageism. In W. Robinson & H. Giles (Eds), *The new handbook of language and social psychology* (pp.465–86). Chichester, UK: Wiley.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213-238.

Cheshire, J. (1987). Age and generation-specific use of language. In U. Ammon, N. Dittmar & K. Mattheier (Eds), *Sociolinguistics: An Introductory Handbook of the Science of Language and Society* (pp.760-767). Berlin: Mouton de Gruyter.

Danlos, L. (1992). Support verb constructions. Linguistic properties, representation, translation. *Journal of French Linguistic Study*, *2*(1), 1-32.

Dubois, B. L., Crouch, I. (1975). The question of tag questions in women's speech: They don't really use more of them, do they? *Language in society*, *4*(3), 289-294.

Firth, J. (1957), A synopsis of linguistic theory 1930-55: Studies in linguistic analysis, philological society. In F. Palmer (Ed.), *Selected papers of J. R. Firth 1952-59*, (pp. 168-206), Bloomington and London: Indiana University Press.

Goodwin, C. (1980). Restarts, pauses, and the achievement of mutual gaze at turn-beginning. *Sociological Inquiry, 50*(3), 272-302.

Holmes, J. (1986). Compliments and compliment responses in New Zealand English. *Anthropological Linguistics*, *28(*4), 485-508.

Holmes, J. (2001). *An introduction to sociolinguistics*. London: Longman.

Kaplan, N., & Farrell, E. (1994). Weavers of webs: A portrait of young women on the web. *The Arachnet Electronic Journal on Virtual Culture*, *2*(3). Retrieved from http://www.monash.edu.au/journals/ejvc/kaplan/v2n3

Kennedy, G. (2003). Amplifier collocations in the British national corpus: Implications for English language teaching. *TESOL Quarterly, 37*(3), 467-487.

Krenn, B. (2000). *The usual suspects: Data-oriented models for identification and representation of lexical collocations* [Unpublished doctoral dissertation]. German Research Center for Artificial Intelligence and Saarland University

Labuhn, U. (2001). *Von Give a Laugh bis Have a Cry. Zu Aspektualität und transitivität der V+N konstruktionen im Englischen,* Frankfurt: Peter Lang.

Lakoff, R. (1975). *Language and women's place*. New York: Harper and Row.

Lee, D. (2001). Genres, registers, text types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning & Technology, 5*(3), 37-72.

Leet-Peregrini, H. M. (1980). *Conversational dominance as a function of gender and expertise. Language: social psychological perspectives*. Oxford: Pergamon

Lewis, M. (2000). *Teaching collocation. Further developments in the lexical approach*. Hove: LTP

Liu, D. (2012). The most frequently-used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes*, *31*(1), 25-35.

Maltz, D., & Borker, R. (1983). A cultural approach to male–female miscommunication. In J. Gumperz (Ed.), *Language and social identity* (pp. 196-216). Cambridge: Cambridge University Press.

Montgomery, M. (1995). The linguistic value of Ulster emigrant letters. *Ulster Folklife*, *41*, 1-16.

Nation, I. S. P. (2004). A study of the most frequent word families in the British national corpus. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language* (pp. 3–13), Amsterdam: John Benjamins

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, *24*(2), 223-242.

Nickel, G. (1968). Complex verbal structures in English. *International Review of Applied Linguistics*, *6*, 1-21.

O'Barr, W., & Atkins, B. (1998). "Women's language" or "powerless language" ? In J. Coates (Ed.) *Language and gender: A reader* (pp. 401-407). Oxford: Blackwell Publishing

Orb, A., Eisenhauer, L., & Wynaden, D. (2001). Ethics in qualitative research. *Journal of Nursing Scholarship*, *33*(1), 93-96.

Orb, A., Eisenhauer, L., & Wynaden, D. (2001). Ethics in qualitative research. *J Nurs Scholar*, *33*(1), 93-96.

Panagiotidou, T. (2015, July). *Corpus-based gender studies: the case of 'empty' adjectives* [Paper presentation]. Conference: 8th Athens Postgraduate Conference 2015, Aristotle University of Thessaloniki, Greece. https://sites.google.com/site/8thapc/abstracts

Pearce, M. (2008). Investigating the collocational behaviour of MAN and WOMAN in the BNC using Sketch Engine. *Corpora, 3*(1), 1-29.

Pennebaker, J., Mehl, M., & Niederhoffer, K. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, *54*, 547-577.

Sinclair, J. M., Fox, G., Bullon, S., Krishnamurthy, R., Manning, E. and Todd, J. (1990). *Collins Cobuild English Grammar*. London: Collins ELT.

Sinclair, J., & Renouf, A. (1988). A lexical syllabus for language learning. In Carter & M. McCarthy R. (Eds.), *Vocabulary and language teaching*, (pp. 140-160), London and New York: Longman.

Sinclair, J. (1990). *Collins Cobuild English grammar*. London: Harper Collins

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Stein, G. (1991). The phrasal verb type 'to have a look' in modern English. *International Review of Applied Linguistics*, *29*(1), 1-19.

Stenström, A. (1999). He was really gormless – She's bloody crap: Girls, boys and intensifiers. In H. Hasselgård & S. Okesfjell (Eds.), *Out of corpora: Studies in honour of Stig Johansson* (pp. 69-78). Amsterdam and Atlanta: Rodopi.

Tannen, D. (1990). *You just don't understand: Women and men in conversation*. New York, NY: William Morrow and Co., Inc.

Tannen, D. (1990). *Gender and discourse*. New York: Oxford University Press.

Trudgill, P. (1983). *On Dialect: Social and geographical perspectives*. Blackwell, Oxford.

Uchida, A. (1992). When difference is dominance: A critique of the anti-power-based cultural approach to gender differences. *Language in Society*, *21*, 547-568.

West, M. (1953). *A general service list of English words*. London: Longman, Green & Co

White, L. (2003). *Second language acquisition and universal grammar*. Cambridge: Cambridge University Press.

Xiao, R., & Tao, H. (2007). A corpus-based sociolinguistic study of amplifiers in British English. *Sociolinguistic Studies*, *1*(2), 241-273.