Women's and men's language use *and all that stuff*: A corpus analysis of general extenders

Oskar Sellberg

Autumn 2015

Supervisor: Matteo Fuoli

Centre for Languages and Literature

Lund University

Abstract

The purpose of this essay is to find and analyze the discrepancies between women and men when it comes to the use of general extenders. The questions that this essay attempts to answer are: Is there any significant difference in the amount of general extenders used by women and men? What specific types of general extenders do women and men use more frequently? For what purpose do women and men use general extenders in discourse? And, what are the most frequent functions these expressions perform in female and male speech?

Since previous studies have not shown any significant difference between the amounts of general extenders used by women and men and have not really studied the functions women and men tend to use these questions are yet to be answered.

The data in the essay have been collected from the British National Corpus (BNC) and the occurrences are of spoken language. The occurrences from the BNC were then compared and calculated with the help of log-likelihood tests to determine if there were any significant differences. In addition to the log-likelihood tests, I have also attempted to annotate a sub-sample of occurrences from the corpus for their discursive functions.

Overall, the data show that general extenders are more frequently used by women, with some exceptions for specific general extenders. The general extenders range from several different discursive functions (e.g. hedges on the Grice's Maxims and politeness strategies). For *or something,* the process of annotation displays an increased use of hedges on the maxim of quality and negative politeness strategy in the occurrences uttered by women. The occurrences of *and so on* show that positive politeness strategy was more frequently used by men than women.

In conclusion, the functions of general extender in discourse are complicated to determine. If one follow the results presented in this essay, it seems as though women are more likely to cooperate in communication. Since they seem to want adhere to the wants of the hearer's negative face and their use of hedges on the Maxim of Quality this cooperation is evident.

Table of Contents

1. Introduction

The different ways in which people use language in real-world situations have been a subject of interest for some time. Sociolinguists have investigated language use from different perspectives, looking at variables such as age, social class and gender. Over the last couple of decades, interest in the relationship between language and gender has increased. One of the early researchers in the field was Robin Lakoff (1973), who states that "there is a discrepancy between English as used by men and by women; and that the social discrepancy in the positions of men and women in our society is reflected in linguistic disparities" (p. 76). Since then, different researchers have attempted to explain the discrepancy between men and women. This formed ideas such as the notion of male dominance, which means that males use language in order to dominate women, and the idea of gender differences, which involved the argument that women and men viewed conversation in different ways, and therefore spoke differently (Baker, 2014, p. 2).

In order to find these discrepancies and to analyze them in a systematic way, a possibility is to look at data of either spoken or written language in a corpus, which is a widely-used method for this type of analyses. Some studies of that sort have shown that women use intensifiers, modal constructions, tag questions and imperative constructions in question more often than men (McMillan et al., 1977, p. 554). A study conducted by Janet Holmes shows that women apologize more frequently than men and that they also compliment others more frequently than men do (Holmes, 1993, p. 105, 107). Aside from the differences stated above there are several other differences found and discussed by other researchers (e.g. Weatherall, 2002).

One linguistic feature that has been considered from a gender perspective is 'general extenders'. However, the results of studies have not displayed a discrepancy in use between women and men. Overstreet (1999) defines general extenders as "a class of expressions that typically occur in clause-final position and have the basic form of conjunction plus noun phrase" (p. 3). There are two different kinds of general extenders: The "adjunctive", which start with *and* (e.g. *and that, and all)*, and the "disjunctive", which start with *or* (e.g. *or whatever, or something*) (Overstreet, 1999, p. 3-4). The analysis and discussion of these general extenders in language use are the main foci of this paper.

The function of general extenders in discourse is not always clear-cut. In most cases, they are used to indicate more entities in a list, set or category (Overstreet, 1999, p. 11). Martinez (2011) claims that "[t]hey should not only be regarded as simple tokens of vague,

sloppy language or hedges since they can also have other functions in discourse" (p. 2455). There is also the possibility for the general extenders to have an interpersonal function and, according to Overstreet (1999), they can "reflect the speaker's attitude toward the message or addressee" (p. 12).

The terminology used for these types of expressions has varied between several different alternatives, some named after the form and sentence position and others for their function. Dines (1980) calls these expressions 'set marking tags' and this term was later used by Winter and Norrby (2000). In more recent studies by Overstreet (1999, 2005, 2012), Cheshire (2007), Tagliamonte and Denis (2010) and Martinez (2011), the term general extenders is used to identify these expressions. Throughout this essay I will refer to these expressions as general extenders.

The aim of this essay is to compare the language used by women and men, focusing on general extenders. The questions that this paper will attempt to answer are:

1. Is there any significant difference in the amount of general extenders used by women and men?
2. What specific types of general extenders do women and men use more frequently?
3. For what purpose do women and men use general extenders in discourse? And, what are the most frequent functions these expressions perform in female and male speech?

The basis for this essay is a quantitative analysis of the frequency of general extenders used by female and male speakers followed by a more detailed analysis of the discursive functions of these expressions in a smaller sample that concerns the different functions of general extenders. The data used have been collected from the British National Corpus (henceforth BNC). This essay will show that there are differences in the amount of general extenders used between female and male speakers and that the functions used for both of them vary. It will also show that the discrepancies are similar in other parts of language.

The essay is organized as follows. In section 2, I provide a review of previous research on gender and general extenders. In section 3, I account for the method and procedure of my collection of data. In section 4, I present my result of frequency (4.1), the discursive functions of general extenders (4.2) and a discussion of my findings (4.3). I end with a conclusion of the result and discussion in section 5.

2. Background

This section provides an overview of previous research concerning both language and gender and general extenders. The following sub-section will account for earlier findings on how women and men talk and why.

2.1. Language and gender

Several seminal studies that involved the comparison of language use between women and men were conducted in the 1970s. Lakoff (1973) comments that "women are marginal to the serious concerns of life" when compared to men. The marginality or lack of power is evident when one looks at the way women speak but also how they are spoken of (Lakoff, 1973, p. 45). Lakoff (1973) explains that in "appropriate women's speech, strong expression of feeling is avoided, expression of uncertainty is favored, and means of expression in regard to subject-matter deemed 'trivial' to the 'real' world are elaborated" (p. 45). The author comes to the conclusion that changes in language originates from changes in society, and not as much the other way around (Lakoff, 1973, p. 76).

In a study with similar results conducted by McMillan et al. (1977), 98 people were videotaped to find out how they spoke in same sex groups and in mixed groups. The authors argue that, contrary to their initial beliefs, women use more intensifiers in a same sex group than in a mixed group (McMillan et al., 1977, p. 552). Overall, McMillan et al. (1977) show that women use language that "conveys uncertainty" more often than men (p. 558).

Holmes (1993) comes to a different type of conclusion than the one that Lakoff (1973) had about women's language. Instead of considering all tag questions to convey uncertainty there are more complex criteria that determine the function of such expressions. In different forms and contexts, tag questions range from the functions of hedges, strategies of negative and positive politeness, and intensifiers (Holmes, 1993, p. 97-98). Holmes (1993) concludes that women do not show powerlessness in speech. Instead, their speech should be considered positive for communicative purposes. Women also show a greater willingness to take the "face needs of others" into account by using more apologies and compliments than men (p. 111-112).

Weatherall (2002) discusses the problems of early research on language and gender, explaining the different approaches of other studies within the field. She stresses the importance of acknowledging that gender is "produced and sustained through patterns of talk"

(p. 7). The work follows the idea that language is a reflection of and aids in sustaining a social system, and that that system is more beneficial for men. Weatherall (2002) continues on that idea and explains the discursive turn of language and gender research. The discursive turn is based on the assumption that gender is constructed by language and discourse and not an "essential characteristics of an individual's psyche" (p. 76).

Newman et al. (2008) use a collection of language from 70 different published studies, adding up to a total of 14,324 text files. Their results show that women more frequently use language to talk about other people and also words communicating doubt. Other words that women use more often than men according to Newman et al. (2008) include "[t]houghts, emotions, senses, other people, negations, and verbs in present and past tense (p. 229). On the other hand, men use words about "occupation, money, and sports" as well as "numbers, articles, prepositions and long words" (p. 229).

In sum, early research on language and gender shows that women use language that conveys uncertainty more often than men. This view, however, has been challenged by more recent studies, which support the view that certain parts of language are instead used as communicative strategies such as hedges and politeness. Hedges may also convey uncertainty but they may also have other more complex functions in discourse.

2.2. General extenders

When looking at the section above, it is evident that disparities exist between men and women in different parts of language but the focus of this essay is concerned with a specific part of language, namely, general extenders.

General extenders can be divided into either adjunctive or disjunctive. The adjunctive general extenders are the ones that start with the conjunction *and.* According to Tagliamonte and Denis (2010) these expressions "tend to generalize to all members of the set to which the referent belongs" and that the function "signals to the listener that the speaker is talking about more than just the referenced member of the set" (p. 337).

The general extenders beginning with the conjunction *or*, which Tagliamonte and Denis (2010) call disjunctive, are explained as "marked by *or* and refer to only one of any member of the set to which the referent refers" and that the function instead "signals to the listener that the speaker is not necessarily talking about the referenced member, but rather to any member of the set" (p. 337).

The attempt to distinguish and analyze general extenders started in the 1980s. The types of expressions are called 'set-marking tags' by Dines (1980) who looks at data collected from interviews with working-class and middle-class women. The results of the author's study show that working-class women have a more frequent use of these 'tags' (Dines 1980, p. 13). One of the conclusions that the author makes is that "there is nothing to suggest that the occurrence of set-marking tag marks 'vague and inexplicit speech'" (Dines, 1980, p. 30).

In the late 1990s, Overstreet and Yule (1997) observe what they decide to call general extenders to distinguish between their different forms and functions (p. 250). The authors collected their data from telephone conversations and face-to-face interactions in a more informal context. The data consist of ten hours of conversation between eleven female and seven male speakers. Another collection of data was made that also consisted of ten hours of conversation, but in a more formal context (Overstreet & Yule, 1997, p. 252). The result of their research is that familiarity seems to matter in conversation because general extenders are more frequently used in the informal data than in the formal data (Overstreet & Yule, 1997, p. 252-253). Overstreet and Yule (1997) conclude that general extenders are not only used to implicate a category, but they are also connected to "intersubjectivity" which includes some kind of "shared experience" or "social closeness" in conversation (p. 254).

Overstreet (1999) provides an extensive explanation of the definitions, forms and functions of general extenders. The study is based on data from conversations over the phone or in person between people who are familiar with each other (p. 16). The analysis of the data is divided into two different segments, one that covers the "ideational" function of language and one that covers its "interpersonal" function. The chapters that concern the "ideational" function include the "speaker-based", "representational function" of language. It includes functions of general extenders, such as vagueness and the role as list completers. Furthermore, it covers a proposal that speakers can use these to implicate a category. When the general extenders implicate a category it is either used for the hearer to infer something more or something different related to a preceding element (Overstreet, 1999, p. 17). The role as list completers is the claim that the general extender performs the last and third entity in a list, as seen in example (1) from Overstreet (1999, p. 23).

(1) "… So that we can put all the kitchen stuff in the::re, an' all the

heavy stuff, an' just pack out our clo::thes an' tents **an' stuff**."

In the segment of "interpersonal" function, Overstreet (1999) argues that "when viewed within their role in the interactive exchange of talk, general extenders appear to have a function that is interpersonal and tied to the social relationships of the participants" (p. 18). These type of interpersonal functions include for example negative and positive politeness strategies and hedges on the Gricean Maxims of quantity and quality.

The idea of the Maxim of Quality is that you should speak the truth; you should not say anything that you do not think is true and you should not utter anything that you cannot back up with proper evidence (Overstreet, 1999, p. 112). Overstreet (1999) explains hedging on the Maxim of Quality like this: "a speaker may assert something that he or she thinks is potentially inaccurate […] but the speaker indicates in some conventional way a lack of commitment to the necessary truth of the content of the utterance, or part of the utterance" (p. 112).

Two different criteria that make up the Maxim of Quantity are explained by Overstreet (1999) like this: "Grice's proposed Maxim of Quantity is expressed in the following way: 1. Make your contribution as informative as is required (for the current purposes of the exchange). 2. Do not make your contribution more informative than is required. (1975: 45)" (p. 126). Adjunctive general extenders are used to convey a meaning that "there is more" instead of listing everything explicitly in one's utterance, and therefore hedging the Maxim of Quantity.

Negative face is what Brown and Levinson (1987) explain as "the basic claim to territories, personal preserves, rights to non-distraction – i.e. to freedom of action and freedom of imposition" (p. 61). Brown and Levinson (1987) note that the negative politeness strategies called hedges, are used when the speaker wants to follow their rule: "DON'T ASSUME H IS ABLE/WILLING TO DO A" (p. 146). Disjunctive general extenders may be connected to negative politeness strategies, functioning as hedges which reduce the risk of threat to the hearer's face (Overstreet, 1999, p. 104).

Positive politeness strategies are related to the speaker's and hearer's positive face. Brown and Levinson (1987) define this as "the positive consistent self-image or 'personality' (crucially including the desire that this self-image be appreciated and approved of) claimed by interactants" (p. 61). A strategy for maintaining the wants of the positive face of others is according to Overstreet (1999) the use of adjunctive general extenders. They can function as a means of indicating intersubjectivity, conveying a sense of shared knowledge (p. 99).

Winter and Norrby (2000) use the same term as Dines (1980) to refer to general extenders, namely set marking tags. Their study is an attempt to identify the different

functions of these tags. The authors use data from two corpora, one of Swedish tape-recorded informal group interviews with senior high school students. The other corpus consists of 48 interviews with adolescents from different schools in Year 10. Winter and Norrby (2000) find that the set marking tags can have what they call "participation meaning", which includes the notion of shared knowledge linked to positive politeness and that of face saving acts linked to negative politeness (p. 5).

Overstreet (2005), based on her previous data collection, attempts to conduct a comparative analysis between general extenders in English and German. The author finds that the functions of the general extenders are similar in the two languages but they occur in different positions of the clause, and with different variability of form (Overstreet, 2005, p. 1861).

Cheshire (2007) analyzes data from three different English towns. The data consist of speech of 32 adolescents per town. The results show that adjunctive general extenders are more frequent than disjunctives. There are no differences in frequency when it comes to social class or gender in the results (Cheshire, 2007, p. 187).

Martinez (2011) also focuses on the language of teenagers, comparing the use of general extenders in the Bergen Corpus of London Teenage Language (COLT) and the Diachronic Corpus of Present-Day Spoken English (DCPSE) (p. 2457-2458). The analysis reveals that general extenders do not only function as set markers, but they also have interpersonal functions. The expressions are more common in spoken than in written language and they are also more frequent within adult speakers (Martinez, 2011, p. 2467).

Based on the earlier literature, it is clear that general extenders have a range of functions in discourse. However, no previous study has found any significant difference in the use of general extenders between women and men. This essay aims to study if there are any disparities between women and men and also what functions general extenders can have in discourse.

3. Methods

This essay will base its argument on data collected from the British National Corpus. The data were limited to only spoken language, which, in the BNC, adds up to a total of 10,409,858 words. In order to answer the question of whether there is any significant difference in the amount of general extenders used by women and men, the data were also collected with

restrictions of sex. The total amount of words spoken by men in the BNC is 4,949,938 and by women is 3,290,569. The rest of the words in the corpus are either marked as unknown sex or not marked by sex at all. The data have not been restricted to any other social variables (e.g. age).

From this collection of text, 35 different types of general extenders were extracted. (A complete list of the general extenders considered can be found in appendix 1[1]). The data collection process was divided into two phases. First, a list of general extenders was compiled based on the studies by Martinez (2011) and Overstreet (1999). Then, the examples retrieved from the corpus were manually inspected to remove irrelevant items.

When I collected the data for the essay, I chose one study that included the form and characterization of general extenders. In order to try to minimize ambiguities in how general extenders are defined I used an explanation by Martinez (2011) who states that general extenders "generally take the form of a conjunction (*and, or*) plus a noun phrase i.e. *and stuff, or something, and things*." Martinez (2011) also notes that they "tend to combine with one single previous item, although in some cases they may form part of a series list of three or more elements" (p. 2454). General extenders refer to an element that is preceded by it. Many times, these elements seem to be noun phrases but could also be clauses, verbs, prepositional phrases and at times also an unclear reference (Martinez, 2011, p. 2454-2455).

There were several issues related to the accurate identification of general extenders. Following the characterizations of general extenders above, I found that many occurrences with the same form as a general extender did not exactly comply with the definition above. There were occurrences in the data where the forms did not refer to a preceding element, as in (2) and (3). But also examples that not necessarily referred to a preceding element, as in (4).

(2) HE7 303 – "It's a grey area, **and everything** that we do is totally ethical and professional, but obviously what were talking about is obtaining information."

(3) F72 787 – "When you've got an idea **or something** to say, put your hand in the air please."

(4) KD8 670 – "not be done to cook [pause] because of the sausage **and that** splashing."

---

[1] It is important to note that the list of general extenders that I have collected from the BNC is not exhaustive, there are other forms of general extenders and new ones might occur in the future.

In example (2), the conjunction *and* starts a new clause where *everything that we do* functions as the subject. This means that the combination of words in this case does not follow the rules of form that general extenders have. Similarly, in example (3), *an idea* and *something to say* seem to be in an either or relation of when to *put your hand in the air.* These types of utterances are not regarded as general extenders, and have therefore been omitted from the overall raw frequency count.

There were some occurrences from the BNC that had ambiguous meaning and reference. In example (4) above, *and that* could function as a general extender, referring back to the preceding element of *the sausage.* With this type of interpretation, *and that* could potentially mean that it is *the sausage* and perhaps the rest of the sausage casserole that is *splashing.* Another interpretation of this utterance could be that *the sausage* is not *splashing* at all, but it is instead a separate element different from *the sausage* that is doing the splashing. Because of the difficulties in interpreting these ambiguous cases, I have chosen not to include them in my data.

When all of these restrictions were taken into account, a total number of 7.449 general extenders were found in the corpus. These general extenders were then divided into narrower categories such as the distinction between adjunctive and disjunctive general extenders. The frequency of individual forms (e.g. *and so on, or something*) was also taken into account and compared across sub-corpora.

To allow for comparison, the frequencies for the general extenders were normalized by using the formula given in McEnery and Hardie (2012): "*nf*= (number of examples of the word in the whole corpus ÷ size of corpus) × (base of normalization)" (p. 49-50). The base of normalization I used was 1,000,000. In order to test whether the differences found between female and male sub-corpora were statistically significant, and not only due to chance, multiple log-likelihood tests[2] were performed. Comparisons with a p-value lower than 0.05 would display a significant difference.

The quantitative analysis of general extenders was combined with a qualitative analysis. The aim of the qualitative analysis was to determine what functions the general extenders had in different contexts. In order to find discrepancies between women and men I used *and so on,* which was more frequently used by men, and *or something,* which was more frequently used by women. I compiled a sub-sample of 200 randomly selected examples. 50 for per general extenders and gender. Based on similar attempts of this kind (e.g. Overstreet

---

[2] I used the log-likelihood calculator on http://ucrel.lancs.ac.uk/llwizard.html to calculate the p-values. A detailed description of log-likelihood tests, significant difference and p-values can be found in appendix 2.

(1999), Martinez (2011)), I tried to account for some possible functions of these expressions in the discourse. I annotated all of the 200 occurrences based on their potential functions in the different contexts. Depending on the context in which the general extender occurred, I annotated it for one or several potential functions. These functions, based on Overstreet (1999), are listed below.

1. Marking or implicating a set or category.
2. Hedge on the Maxim of Quantity.
3. Hedge on the Maxim of Quality.
4. Positive politeness strategy.
5. Negative politeness strategy.

After the annotation process of the general extenders, I analyzed them quantitatively, counting the number of certain functions that occurred in the sub-sample. The aim of this analysis was to find general tendencies of which functions were more frequently used than others, and to provide an answer to question 3.

4. Results

This section provides the results and discussion of the quantitative and qualitative/functional analyses. First, it accounts for the frequency and statistical analysis of all the general extenders. Then, the results of the analysis of the discursive functions of selected general extenders are provided, followed by an overall discussion of the findings.

4.1. Frequency and statistical analysis

The aim of this section is to provide the basis for the analysis. It shows the normalized frequencies of general extenders in the BNC and the results of the log-likelihood tests. The results are presented on a more general level at first, and then to a more specific level considering the frequency of individual general extenders (e.g. *or something* and *and stuff like that*).
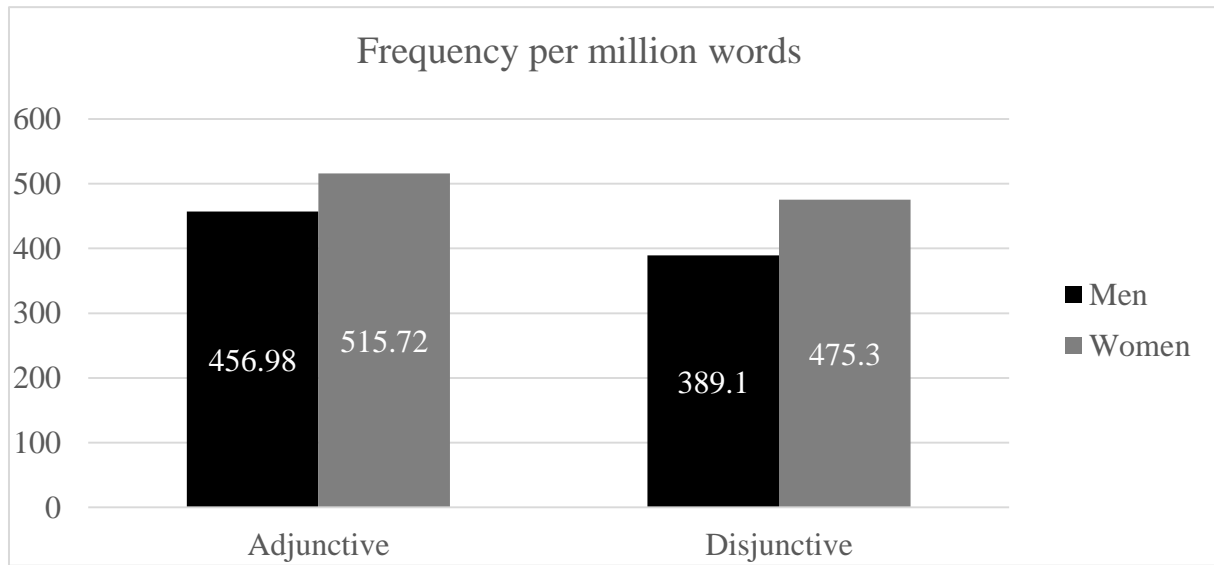
Figure 1. Difference in frequency per million words between women and men for adjunctive and disjunctive general extenders.

As can be seen in figure 1, both the frequency per million for adjunctive and for disjunctive general extenders were higher in women's language use. The number of general extenders in the data provided a significant difference in use for all general extenders (LL= -45.5, $p < 0.001$), adjunctive (LL= -14.09, $p < 0.001$) and disjunctive (LL= -34.27, $p < 0.0001$), with women as the most frequent users. Overall, these results are similar to what other researchers have found. Martinez (2011) finds that adjunctive general extenders are more frequent than disjunctive general extenders, in the language of both adults and teenagers (p. 2460).

General extender	LL	p-value
and stuff like that	12.6	$p < 0.001$
and that sort of stuff	4.8	$p < 0.05$
and so on	195.68	$p < 0.0001$
and so on and so forth	23.82	$p < 0.0001$
or something like that	18.48	$p < 0.0001$
or so	22	$p < 0.0001$
or whatever it is	7.24	$p < 0.01$

Table 1: General extenders that are more frequently used by men and have a significant difference ($p < 0.05$).

Although the frequencies for general extenders were much higher in women's language, I found that some had a much higher frequency for men. *And so on* showed the most significant

difference (LL=195.68, p < 0.0001). Other general extenders that showed a significant difference can be seen in table 1. The collection of these general extenders does not show a specific pattern of discrepancy. There are adjunctive as well as disjunctive general extenders that occur more frequently with men.

When one considers what general extenders women tend to use more frequently than men, the result was that *or something* showed the most significant difference (LL= -108.91, p < 0.0001). Other general extenders that show this significant difference can be seen in table 2 below.

General extender	LL	p-value
and everything	-97.18	p < 0.0001
and that	-103.62	p < 0.0001
and that lot	-6.11	p < 0.05
and all	-5.36	p < 0.05
or something	-108.91	p < 0.0001
or anything	-32.65	p < 0.0001

Table 2: General extenders that are more frequently used by women and have a significant difference (p < 0.05).

The results show that *or something* is not the only general extender that displays highly significant disparity but also *and everything* and *and that*. Similar to what the results for men's most significant forms display, women's most significant forms are not dominated by either adjunctive or disjunctive, but instead consist of both types.

When I looked at women's and men's usage of specific general extenders I found that many did not show any significant disparities (p > 0.05). Because of this result, those occurrences will not be analyzed and discussed further in this essay.

Previous research by others such as Cheshire (2007) show that there were no disparities between the usage of general extenders in language used by women and men and to analyze why would not yield any noteworthy outcome (p. 155). However, since the data this essay have provided show a significant difference both on the more general level between adjunctive and disjunctive but also of more specific general extenders, there might be a reason as to why that is the case. In the next section, this essay will try to account for the potential reasons for the significant differences by studying the discursive functions of general extenders.

4.2. Discursive functions of general extenders

This section includes an attempt to determine what functions the different general extenders can have in different contexts. This type of analysis, in combination with the quantitative analysis, can provide information as to why there are language disparities between women and men in the use of this kind of expressions. The analysis focuses on two general extenders, *and so on* and *or something*, because they display a significant difference in their use between women and men. *And so on* was more frequently used by men and *or something* more frequently used by women. I analyzed 50 randomly selected occurrences from the data for both of the general extenders.

For the disjunctive forms, there is a multitude of functions that are possible in different utterances. These different occurrences have a number of potential functions. If one considers example (5) below uttered by a female speaker, one could argue that the function of *or something* could be that of "set marking", which Winter and Norrby (2000) explain as marking "something … as belonging to a more general category, the set" (p. 2). The set, in this case, could be a general category of what happened to the *ceiling tiles,* with *broken* marking the general category of the state of which they are in. However, to simply state that this is the only possible function of this general extender would not necessarily be right. Instead, one could consider that *or something* functions as what Overstreet (2012) calls "a hedge on the accuracy of what is being said" (p. 3). The speaker might not know if the ceiling tiles got broken or if something else had happened to them. Overstreet (1999) addresses this type of function of *or something* and explains it as "hedges on the Gricean Maxim of Quality" (p. 112).

(5) KC7 1497 – "I remember that we, we had erm a corridor fine in our hall for er just something stupid like they one of the ceiling tiles got broken **or something**."

(6) JP4 1101 – "And, and draw it, and stand up and think how you're going to do it, and maybe you sort of, you know, sit here **or something**, when I've gone and face the window and say, that's North, and I must make sure I get it set up, get this the right way up and [unclear],"

Example (6) above, uttered by a man, allows for another possible function. The speaker wants the hearer to adhere to some kind of instruction. Again, the possibility of *or something* being a set marker in this example is also present. However, instructing someone else to do something like *sit here,* can function as a threat to someone's negative face. If the speaker in example (6) above is interested in maintaining the wants of the hearers face, he can choose to add the general extender *or something* after his instruction *sit here*. By doing so, he uses the negative politeness strategy called hedging. In this case, H is the hearer and A is *sit here*.

Adjunctive general extenders have, as disjunctive ones do, several potential functions in the discourse. The analysis that I attempted to perform for adjunctive general extenders was of *and so on* specifically. These general extenders may also have the function of set marking, but instead of referring to a different entity of the set, they add on more entities that could be included within the specific category. This type of function can be seen in example (7) below.

(7) KRL 46 – "I would just like to [unclear] I know we need to the television programmes on about the Gulf **and so on**, but I've two young children at home, and I find that yesterday was a very long day because there was just nothing on for them at all B B C two at four o'clock."

The general extender *and so on* may in this particular context function as referring to another entity which is similar to *television programmes about the Gulf* (i.e. Programmes not only specifically about the Gulf but of other entities in that category). These entities could include other nature programmes similar to the one spoken of. If one takes the time of the utterance into account, the Gulf could refer to the Gulf war which would be a completely different category.

(8) KRE 397 – "Nobody has the monopoly of wisdom and it's important, it seems to me, that you have economist keeping an eye on the economics of the situation, sociologist looking after the cultural and individual choice sides, psychologists, people's appraisal of their government, geographers saying this sort of development is more appropriate there because of environmental conditions than there, **and so on**."

Instead of functioning as hedges on the Maxim of Quality as the disjunctive general extenders do, these types can function as hedges on the Maxim of Quantity. In example (8) above, the speaker is stating the different areas certain professions should be in charge of when it comes to some kind of knowledge and skills. Since there are a lot more professions than economists, sociologists, psychologists and geographers the general extender *and so on* at the end of the utterance might be a hedge on the Maxim of Quantity. The speaker is not explicitly stating what the other professions are and is therefore hedging the first criterion of the Maxim of Quantity. Another possibility in this particular utterance is that the general extender is referring to what kind of information geographers are supposed to provide. Again, the speaker might be hedging the Maxim of Quantity by not stating all of them.

Some examples of *and so on* in the sample had yet another possible function. Similar to the function of negative politeness hedging, the adjunctive general extender *and so on* provides a possible function involved in positive politeness strategies. A way of maintaining the hearer's positive face is the use of *and so on*. If one considers example (9) below, a potential function of *and so on* could be the common idea that it is a set-marker, indicating that there is something more. The function could also be related to intersubjectivity and shared knowledge. In this situation the speaker might attempt to show that she is aware that the hearer also has the knowledge and understands what she is talking about. An indication that this function is a possibility is that the speaker also uses *you know,* which conveys a similar meaning (Overstreet & Yule, 1997, p. 254-255).

(9) HV5 2 – "I think television would be as much to blame as any you know the bairns the children's programmes are all in English **and so on** you can."

From the 50 examples collected from women's usage of the general extender *or something*, based on my own interpretation of function, I found that 26 of the occurrences functioned as hedges on the Maxim of Quality. When it came to negative politeness strategy, I found that 8 of the occurrences had a possible function of hedging. The men, when I made the same judgement of function, had 22 instances of hedging the Maxim of Quality and 4 instances connected with negative politeness strategy.

The adjunctive *and so on,* because of its function of expressing something more, had a possible function of a hedge on the Maxim of Quantity in all the 50 examples from both women and men. On the other hand, I found that women used *and so on* as a positive

politeness strategy 20 times in contrast to men, who used it 24 times. All of the occurrences for both *and so on* and *or something* had the possibility of functioning as marking or implicating a set or category.

4.3. Discussion

The results of the quantitative and qualitative/functional analysis are discussed in combination with other theories of language and gender as well as theories of general extenders.

The way the general extenders are used in language differs depending on context and other variables. A function that general extenders has is that of vagueness, and the use of vague language seems to have received the negative property of stupidity (Overstreet, 1999, p. 22). Researchers have shown that these types of utterances are not simply fillers in conversation but have certain functions. One of the functions is explained by Dines (1980), who states that "in every case their function is to cue the listener to interpret the preceding element as an illustrative example of some more general case" (p. 22). This type of function can be interpreted in example (10) below.

> (10) KBD 7168 PS040 – "Well Cajun stuff **and all that** I thought, and then there's them you can do some long kebabs **and all that**, that's why if we get one of them charcoal grills that you can pa shoo, shee shoo."

For both of the occurrences of *and all that* in example (10) the general extenders may implicate the category of *food* that generally accompanies *Cajun stuff* and *long kebabs.* Both of the general extenders in (10) have an element of vagueness which correlates to the difficulty in knowing what the speaker exactly means with *and all that*.

If one only assumes the most common function of these general extenders, the one marking a set or indicating a category, a general sense of expressing vagueness and tentativeness would reflect on the speaker. But, if we take the multifunctionality of these expressions into account, that idea need not be true anymore. The use of general extenders can instead be considered to be closely related to cooperation in communication. Women's language, or the culture of feminine language tends to take interpersonal aspects into their means of communication (McMillan et al, 1977, p. 554). This could explain their increased use of general extenders with the functions of hedges on the Maxim of Quality and Quantity.

They display a greater willingness to cooperate with their interactant by attempting to not break these maxims.

Different positions held by women and men in society could explain their discrepancy in the use of negative politeness strategies. Men are more likely to threaten the negative face of their addressee because they are more often in a position to give directives, instructing or even ordering others (Weatherall, 2002, p. 87). Therefore, men are less likely to use negative politeness strategies. On the other hand, women tend to use language, as McMillan et al (1977) state, to "be polite and permit others to have different perspectives or desires" (p. 555). Hence, women could be said to adhere to the wants of other people's negative face.

The overall frequency and the interpretation of function of *and so on* suggest that men use language with the function concerning positive politeness strategy more often than women. Contrary to what the results display, Holmes (1993) notes that "there is evidence that women's concern for others extends to the frequency with which they use politeness strategies" (p. 112). Holmes (1993) further develops on the idea that women apart from men use these strategies, stating that "[t]here are clearly definable female norms which put the addressee's interests and needs first" (p. 113).

The general tendency is that women have an increased use of both adjunctive and disjunctive general extenders. For some specific general extenders, men show a more frequent use. These results in combination with the discursive functions of adhering to Grice's Maxims and politeness strategies contradict the notion that general extenders only convey vagueness and tentativeness. Instead, the discursive functions are closely connected to cooperation in communication. Holmes (1993) argues that women are "ideal speaker-hearers" in the way that they use not only politeness strategies, but a range of other strategies that aid communication (p. 111).

5. Conclusion

This essay studies the use of general extenders by female and male speakers. In order to compare the use of general extender in the BNC, normalized frequencies for the general extenders were calculated. Then, a log-likelihood test was performed to account for the significant differences between the use of these expressions for women and men. After the more quantitative analysis of the essay, a qualitative analysis of the functions of general

extenders was conducted by means of manual corpus annotation. The qualitative analysis and annotation were attempts to clarify how and why the functions occur in communication, and provide further insights into the observed differences between female and male speakers.

Overall, the results display a more frequent use of general extenders by women, both for the adjunctive and disjunctive forms. In a more narrow division of the general extenders, significant differences were found for the expressions. Some general extenders were used more frequently by men and others were used more frequently by women.

The detailed analysis of the discursive function of general extenders has shown that general extenders do not only have a function as a set marker or indicating a more general category of a preceding element, but are instead multifunctional in their use. The possible functions of these expressions can also be interpersonal in nature, such as expressing solidarity, shared knowledge and face-saving functions. They range from hedges on the Maxim of Quality (disjunctive general extenders) and Quantity (adjunctive general extenders) to politeness strategies. The disjunctive general extender can have the function of negative politeness hedges, whereas the adjunctive general extenders can have the function of positive politeness hedges.

Both women and men seem to adhere to the Maxim of Quality in certain utterances containing *or something*. However, the women were more frequent in their use. The men, if we consider the normalized frequencies, used more hedges on the Maxim of Quantity when using the general extender *and so on*.

The negative politeness strategy connected to *or something* was significantly more common in the language of women but the positive politeness strategy connected to *and so on* was more frequent in the language of men. The reason as to why men do not use the negative politeness strategy as frequently could be that they are more often than women in a position to give orders or directives. That women are more interested in the needs and faces of others is also evident in other studies and could be an explanation as to why they use general extenders relatively more often.

The discursive functions of general extenders discussed in this essay allow for a different conclusion than earlier studies have reached. Even though the tentativeness and vagueness remain as functions or attributes connected to the use of general extenders, the analysis in this essay strongly contradicts the argument that these are the main or only functions of general extenders. Instead, the use of general extenders should in a great deal of cases be considered a means of cooperating in communication. This cooperation is related to the notion of ideal speaker-hearers and that people care about the interests of others.

In conclusion, some limitations of the present study should be acknowledged. First, the list of general extenders analyzed in this essay is not exhaustive, a complete list of all general extenders might have displayed different results. Furthermore, there are some problems when considering the general extender *and that.* When used in spoken communication, this general extender could easily be used when pointing at an object in close proximity of the interactants. It is also important to note that the interpretation of functions for these general extenders are tentative, because of the difficulties in understanding the context and the participants of the spoken occurrences one can only make attempts to understand how and why they are used. A reliability test would be useful in order to assess if my annotations of functions are replicable. Another limitation to the study is that the general tendency that I have found need not be true for all women and men. In addition to this, the BNC is also limited to only include a number of people from Britain alone, not necessarily representing the entire English-speaking world.

Contrary to other investigations on the subject such as the one by Cheshire (2007), this study has revealed differences in the use of general extenders between the genders and that opens up for further research. Studies that go into more depth when it comes to the context and familiarity of the speakers and addressees could improve our understanding of how social parameters affect women and men differently, and therefore also their language. In order to fully comprehend the functions of general extenders in discourse several variables have to be taken into account. If a more comprehensive theoretical approach to this type of research can be developed and used, further studies can determine the functions with higher accuracy.

References

Baker, P. (2014). *Using Corpora to Analyze Gender.* London: Bloomsbury.

Brown, P., Levinson, S. (1987). *Politeness.* Cambridge: Cambridge University Press.

Cheshire, J. (2007). Discourse variation, grammaticalisation and stuff like that. *Journal Of Sociolinguistics, 11*(2), 155-193.

Dines, E. R.. (1980). Variation in Discourse: 'And Stuff Like That'. *Language in Society, 9*(1), 13-31.

Holmes, J. (1993). New Zealand women are good to talk to: An analysis of politeness strategies in interaction. *Journal of Pragmatics, 20*(2), 91-116.

Lakoff, R.. (1973). Language and Woman's Place. *Language in Society 2*(1), 45-80.

Martinez, I. (2011). "I might, I might go I mean it depends on money things and stuff" A preliminary analysis of general extenders in British teenagers' discourse. *Journal of Pragmatics, 43*(9), 2452-2470.

McEnery, T. & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice.* Cambridge: Cambridge University Press. Retrieved from https://www.dawsonera.com/readonline/9781139155250/startPage/9

McMillan, J. R., Clifton, A. K., McGrath, D., & Gale, W. S. (1977). Women's Language: Uncertainty or Interpersonal Sensitivity and Emotionality?. *Sex Roles 3*(6), 545-559.

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender Differences in Language Use: An Analysis of 14,000 Text Samples. *Discourse Processes: A Multidisciplinary Journal, 45*(3), 211-236.

Overstreet, M. (2005). And stuff *und so*: Investigating pragmatic expressions in English and German. *Journal of Pragmatics, 37*(11), 1845-1864.

Overstreet, M. (2012). Pragmatic expressions in cross-linguistic perspective. *Applied Research in English, 1*(2), 1-13.

Overstreet, M. (1999). *Whales, Candlelight, and Stuff Like That: General Extenders in English Discourse.* New York: Oxford University Press.

Overstreet, M., & Yule, G. (1997). On Being Inexplicit and Stuff in Contemporary American English. *Journal of English Linguistics, 25*(3), 250-258.

Tagliamonte, S.A., Denis, D. (2010). The Stuff of Change: General Extenders in Toronto, Canada. *Journal of English Linguistics, 38*(4), 335-368.

Weatherall, A. (2002). *Gender, Language and Discourse.* Hove: Routledge.

Winter, J., & Norrby, C. (2000). Set marking tags – 'and stuff'. In *Proceedings of the 1999 Conference of the Australian Society,* 1-8.

Appendix 1

General extender	Raw frequency		Frequency per million		
	Men	Women	Men	Women	LL
And things	259	193	52.32	58.65	-1.43
And things like that	193	129	38.99	39.2	0
And stuff	138	115	27.88	34.95	-3.18
And stuff like that	72	21	14.55	6.38	12.6
And everything	240	359	48.49	109.1	-97.18
And everything like that	7	7	1.41	2.13	-0.58
And everything else	50	37	10.1	11.24	-0.24
And that	333	458	67.27	139.19	-103.62
And that sort of thing	58	45	11.73	13.68	-0.6
And that sort of stuff	4	0	0.81	0	4.8
And that type of thing	1	2	0.2	0.61	-0.87
And that kind of thing	13	7	2.63	2.13	0.21
And that lot	7	14	1.41	4.25	-6.11
And so on	544	93	109.9	28.26	195.68
And so on and so forth	43	4	8.69	1.22	23.82
And all	118	107	23.84	32.52	-5.36
And all that	149	96	30.1	29.17	0.06
And all that stuff	4	1	0.81	0.3	0.91
And all this stuff	0	1	0	0.3	-1.84
And all that crap	2	1	0.4	0.3	0.06
And all this/that business	12	4	2.42	1.22	1.58
And all these things	5	0	1.01	0	5.1
And all like that	10	3	2.02	0.91	1.66
Or something	795	881	160.61	267.73	-108.91
Or something like that	235	94	47.48	28.57	18.48
Or something of this/that kind	5	1	1.01	0.3	1.53
Or so	227	85	45.86	25.83	22
Or anything	182	215	36.77	65.34	-32.65

Or anything like that	93	62	18.79	18.84	0
Or anything of that kind	0	0	0	0	0
Or anything of that sort	3	1	0.61	0.3	0.4
Or whatever	334	204	67.48	62	0.91
Or whatever it is	50	16	10.1	4.86	7.24
Or whatever you call it	2	4	0.4	1.22	-1.74
Or whatever they are	0	1	0	0.3	-1.84
Total	4188	3261	846.1	991.01	-45.5
Total adjunctive	2262	1697	456.98	515.72	-14.09
Total disjunctive	1926	1564	389.1	475.3	-34.27

Appendix 2

The results for the log-likelihood calculations give different values which I have decided to name LL. Values of LL correspond to p-values that determine whether or not there is significant difference. Positive LL values correspond to an overuse by men and negative LL values correspond to an overuse by women. The different p-values that are used in this essay can be seen below. Retrieved from: http://ucrel.lancs.ac.uk/llwizard.html.

- "95th percentile; 5% level; $p < 0.05$; critical value = 3.84
- 99th percentile; 1% level; $p < 0.01$; critical value = 6.63
- 99.9[th] percentile; 0.1% level; $p < 0.001$; critical value = 10.83
- 99.99[th] percentile; 0.01% level; $p < 0.0001$; critical value = 15.13"

No specific p-values were calculated between these critical values and that resulted in only four different p-value outcomes.

McEnery and Hardie (2012) explain that "when we do multiple significance tests, we expect some of them to give a false result, just by chance" and that "95% is standard in statistics" (p. 52). The 95% cut-off point means that one expects that one in twenty of all the occurrences is not valid and only occurs by chance.