

## ABSTRACT

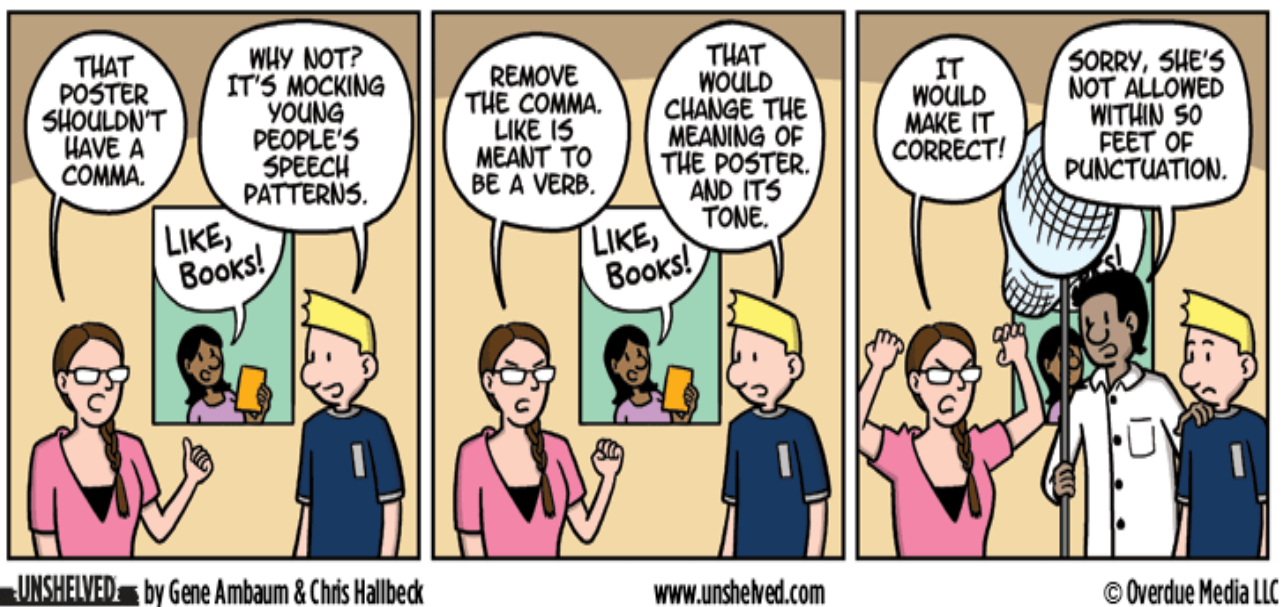
The following thesis in English sociolinguistics consists of a study into the linguistic variable *like*, and the way in which teens and young adolescents of multi-ethnic and multi-lingual backgrounds use this word. The research problem is based on the fact that many people perceive this to be a poor way to communicate; and that people who use the word *like* within their spoken English are often considered to be eroding the English language in some way or are often viewed as being sub-standard speakers of English or generally linked to lacking in vocabulary or education. My thesis therefore attempts to demonstrate that this is not the case, and that the individuals who use *like* are in fact innovators of English, who are creating and frequently using new forms of spoken English in order to express themselves.

In order to test this claim, I interviewed ten individuals between the ages of seventeen and twenty-four in order to analyse their uses of *like* for patterns involving gender, age, and grammatical structure among other things. This was done by transcribing all of the ten interviews and collecting a corpus of data where all the linguistic tokens of *like* were collected and assessed using the statistical analysis program Rbrul. Although *like* can be used in spoken English in a myriad of different ways, the analysis of my interviews focused mainly on the two categories of discourse particle and marker, as these were by far the most frequently used tokens of *like* that were present in my interviewees' spoken English.

The main goal of my thesis is to demonstrate that *like* is in fact a word that has a great deal of structure surrounding the way that it is used, and that it is in fact too simplistic and even dismissive to describe the way it is used as chaotic or haphazard. Indeed, the analysis of my corpus attempts to demonstrate patterns in the way my interviewees use the variable. Perhaps the most salient pattern that was identified in my analysis is the way that the marker is strongly favoured by the male participants, and that this tendency to use the marker increased as the interviewees got older. The marker was therefore found to occur within a well-defined structural pattern whereby it was favoured by older male individuals and the word *like* was usually followed by a declarative clause. The other main finding concerning the discourse particle also runs parallel to this in that the female interviewees strongly favoured the use of the particle followed by a noun phrase. In terms of age, the younger the participant was, the more likely they were to use the particle. This demonstrates that within the spoken English of these young, multi-lingual citizens of Denmark, the discourse marker and the particle are on two different historical trajectories.

I therefore concluded my thesis by stating that although it may be tempting to describe the way people use *like* as some sort of unstructured and incorrect way to use the English language, a linguistically-informed study of the variable in fact shows how structured the word is, and that perhaps many of the preconceptions we have surrounding *like* are in fact misconceptions.

## The Spread of English *like* in the Diaspora: A Grammatically Embedded Phenomenon.



**Number of characters:** 129656  
**Supervisor:** Anne Fabricius  
**Written by:** Richard Peter Gordon-Orr (5311)  
**Contract number:** 8966

# Table of Contents:

<b>1. Introduction.....</b>	<b>7</b>
<b>1.1 Subject/Problem Area.....</b>	<b>7</b>
<b>1.2 Problem Definition.....</b>	<b>8</b>
<b>1.3 Research Questions.....</b>	<b>8</b>
<b>2. Methodology.....</b>	<b>8</b>
<b>2.1 Introduction.....</b>	<b>9</b>
<b>2.2 The Interviewees.....</b>	<b>9</b>
<b>2.2.1 Table 1: The interviewees.....</b>	<b>10</b>
<b>2.3 The Interview Process.....</b>	<b>11</b>
<b>2.4 Transcriptions/Analysis.....</b>	<b>12</b>
<b>3. Theory.....</b>	<b>13</b>
<b>3.1 Grammatical classes of the linguistic variable.....</b>	<b>13</b>
<b>1.Quotative <i>like</i>.....</b>	<b>13</b>
<b>2. Particle <i>like</i>.....</b>	<b>14</b>
<b>3. Adverbial <i>like</i>.....</b>	<b>15</b>
<b>4. Tag <i>like</i>.....</b>	<b>15</b>
<b>5. Discourse marker <i>like</i>.....</b>	<b>16</b>
<b>6.Sentence adverb <i>like</i>.....</b>	<b>16</b>
<b>7.Verbial <i>like</i>.....</b>	<b>17</b>
<b>8.Comparative complementizer <i>like</i>.....</b>	<b>17</b>
<b>9.Other uses (Preposition, conjunction).....</b>	<b>17</b>
<b>3.2 Theories concerning the structure of the marker and the particle.....</b>	<b>18</b>
<b>1.The syntactic position of <i>like</i> .....</b>	<b>19</b>
<b>2. Complementizer phrases.....</b>	<b>20</b>
<b>3. Enumerations/Responses to direct questions.....</b>	<b>20</b>
<b>4. Relative/Subordinate clauses.....</b>	<b>20</b>
<b>5. Determiner phrases.....</b>	<b>21</b>
<b>6. The adjectival domain.....</b>	<b>21</b>
<b>7. The verbal domain.....</b>	<b>22</b>

8. The adverbial domain.....	22
3.3 Theories involving quotative <i>be like</i> .....	23
1. Grammatical person.....	23
2. Content of the quote.....	23
3. Tense/Temporal reference.....	24
4. Quotative <i>be like</i> : a black swan event.....	24
3.4 Misconceptions/Preconceptions concerning <i>like</i> .....	25
1. <i>Like</i> is just a meaningless filler word.....	25
2. People who use <i>like</i> are inarticulate.....	26
3. Women use <i>like</i> all the time.....	27
4. <i>Like</i> is a Valley Girls/Adolescent phenomenon.....	28
4. Historical context.....	29
4.1 Spoken English on a global scale.....	29
4.1.1 Table 2: Categorisation of the interviewees.....	30
4.2 Discourse marker and particle.....	31
4.3 Quotative <i>be like</i> .....	32
4.4 Conclusion.....	33
5. Analysis.....	33
5.1 Introduction.....	33
5.2 Primary Statistical Results Data.....	34
Grammatical classifications of <i>like</i> .....	34
Table 3: Grammatical classifications by gender.....	35
Tables 4: Grammatical classifications in order of frequency	
Most to least frequent class by gender for females.....	35
Table 5: Grammatical classifications in order of frequency	
Most to least frequent class by gender for males.....	36
Table 6: Cross-tabulations of linguistic tokens (without interviewer)	
contrasting classes 2+5 versus the others.....	37

Table 7: Percentages of linguistic tokens 2+5 by gender.....	38
One-level analysis of response class with predictors: Age and Gender.....	38
Table 8: Cross-tabulations for Interviewer's tokens.....	39
Table 9: Cross-tabulations with percentages.....	40
<b>5.3 Secondary Statistical Results Data.....</b>	<b>41</b>
Table 10: Cross-tabulations of marker/particle tokens.....	41
Table 11: All Speakers Together (including the interviewer).....	43
Table 12: Female speakers.....	43
Table 13: Male Speakers (including the interviewer).....	43
Results from the cross-tabulations.....	44
Analysis of the discourse marker.....	44
1.One-level Analysis of Response Class with Predictors: Following	
Group+Age.....	44
Analysis of the discourse particle.....	45
2.One-level Analysis of Response Class with Predictors: Following	
Group+Age.....	45
3. One-level Analysis of Response Class with Predictors: Following	
Group+Gender.....	46
4.One-level Analysis of Response Class with Predictors: Following	
Group+Age+Gender.....	47
<b>5.4 Final Statistical Results Data.....</b>	<b>49</b>
Step-up/step-down analysis of prediction of marker (including interviewer)...	49
Step-up/step-down analysis of prediction of particle (including interviewer)...	51
Step-up/step-down analysis of prediction of marker (excluding interviewer)...	53
Step-up/step-down analysis of prediction of particle (excluding interviewer)...	55
<b>5.5 Summary of the main findings of the analysis.....</b>	<b>57</b>
<b>6. Discussion.....</b>	<b>58</b>
<b>7. Conclusion.....</b>	<b>59</b>
<b>8. Bibliography.....</b>	<b>61</b>

**List of Tables:**

. Table 1: The interviewees.....	10
. Table 2: Categorisation of the interviewees.....	30
. Table 3: Grammatical classifications by gender.....	35
. Table 4: Grammatical classifications in order of frequency Most to least frequent class by gender for females.....	35
. Table 5: Grammatical classifications in order of frequency Most to least frequent class by gender for males.....	36
. Table 6: Cross-tabulations of linguistic tokens (without interviewer) contrasting classes 2+5 versus the others.....	37
. Table 7: Percentages of linguistic tokens 2+5 by gender.....	38
. Table 8: Cross-tabulations for Interviewer's tokens.....	39
. Table 9: Cross-tabulations with percentages.....	40
. Table 10: Cross-tabulation of marker/particle tokens.....	41
. Table 11: All Speakers Together (including the interviewer).....	43
. Table 12: Female speakers.....	43
. Table 13: Male Speakers (including the interviewer) .....	43

# 1. INTRODUCTION

## 1.1 Subject/Problem Area

The following assignment consists of a sociolinguistic investigation of the way in which young multi-lingual speakers use the English language. The linguistic variable I have chosen to examine is the word *like*. Specifically, the way in which young people between the ages of seventeen and twenty-four use the word and how much they do so; as well as the structure of their spoken English involving the linguistic variable in question. My desire to write my thesis on this topic stems from the fact that I feel this to be a worldwide phenomenon that many people seem to have a strong opinion on.

Have you ever for instance been in conversation with someone or perhaps watched a clip on Youtube where someone uses *like* within their oral speech and thought this to be grammatically incorrect? Or some sort of strange unstructured part of the individual in question's spoken English? Or perhaps you have been around a dinner table with family and heard a senior member of the family commenting on the fact that young people today use *like* far too much and in completely the wrong way; or that it is merely an “American thing” or a trait attributed to teenage girls? If so, then I believe this project will be of interest to you.

My desire to work with this topic initially came about when I started reading *Teen Talk: The language of Teens* by Sali A. Tagliamonte and *Discourse-Pragmatic Variation in Context: Eight hundred Years of LIKE* by Alexandra D'Arcy during my Masters studies at university, and these two books have been invaluable throughout my work on this project. What I found particularly fascinating about these two books is the way that they attempt to disprove a lot of the pre-conceptions that we have about *like*, and that the way it is being used by young people today is a form of linguistic innovation as opposed to representing a 'bastardisation' of the English language as so many people believe it to be. The main goal of Alexandra D'Arcy's book for instance, is to show that *like* is systematic, layered and grammatically embedded. It therefore behaves in each function as do all features of language: following rules of usage within a circumscribed (variable) grammar. Her intention is therefore to demonstrate that “a linguistically-informed perspective utterly undermines any claim that *like*, in any of it's uses, is random and meaningless” (D'Arcy, 2017:31).

I decided that the most interesting way to approach my own work would be to attempt to interview people within the community in which I live in order to analyse the way they use *like*. The goal of this being to test whether the variable in question

was random and chaotic as in my experience a large number of people believe it to be; or whether it was in fact something that was structured and innovative as Tagliamonte and D'Arcy are saying. I therefore chose to interview ten people between the ages of seventeen and twenty-four, all of whom use English as either a primary or secondary language. The main reason for the choice of this particular target group is that this is the main point in a person's life where linguistic change occurs in one's oral speech, and this will be elaborated on further within the project. I also chose to interview both male and female members of the community in order to analyse the role of gender on the linguistic variable. My goal therefore was to build up a corpus of work for my analysis, and this was done by transcribing all ten interviews and extracting linguistic tokens of my interviewees' uses of *like*. Over the course of eight and a half hours of authentic recorded speech (the transcriptions of which have been attached in the appendix), exactly 1866 sentences containing *like* were collected and analysed using the statistical analysis program Rbrul. and the results of this quantitative analysis are presented in chapter 5 of this thesis.

### **1.2 Problem Definition:**

With all of this in mind, I would like to present the following problem formulation for my thesis:

**“To what extent does the linguistic variable *like* as a marker and a particle show structure and frequency within the spoken English of my interviewees? And to what extent can the variable be described as unpredictable and haphazard?”**

### **1.3 Research questions:**

- . How does the linguistic variable perform throughout the analysis compared to what is expected of it from the theory section?
- . What are the most frequent types of *like* tokens in terms of class?
- . What role does gender have in the analysis of the variable?
- . What role does age have in the analysis of the variable?
- . Which grammatical functions of *like* can be described as innovative and why?
- . To what extent is this a global phenomenon?
- . How has *like* functioned and performed throughout time?
- . What are the general pre-conceptions concerning *like*?
- . To what degree does the variable show grammatical structure and consistency?
- . Can the variable be described as chaotic and haphazard?

## **2. METHODOLOGY**



## 2.1 Introduction

In order to answer my main problem formulation question, I decided to conduct quantitative sociolinguistic interviews with young adolescents living in Denmark between the ages of seventeen and twenty-four. After some initial discussions with my supervisor, it was quickly decided that ten interviews would be considered a sufficient number in order to acquire enough tokens of the linguistic variable for analysis. The main goal of the interviews then would be to create what was essentially an informal chat between the interviewer (myself), and whoever I was speaking to. Specifically, to elicit genuine, relaxed spoken interactions that would hopefully help encourage and acquire as many tokens of *like* as possible from my interviewees.

## 2.2 The Interviewees

In chapter 3 of Sali Tagliamonte's *Teen talks: The language of adolescents*, Tagliamonte states that age has a huge impact on how a person uses the English language in general, and this is why I initially decided that my target group for interviewees should be between seventeen and twenty-four. Tagliamonte also explains that “as teenagers gain independence, their vocabulary broadens as they are exposed to new forms of language. And when these new uses spread among more and more teens, new expressions enter the English language and can even begin to influence English grammar, which in turn leads to young people becoming a driving force behind language change” (Tagliamonte, 2016:43).

From this statement it is fairly clear that teenagers would be an important group who could potentially be using the variable. But I also wanted to interview some slightly older individuals, such as people in their early twenties. I therefore decided to make the maximum age for my interviewees twenty-four (this number was also based on who was willing to be interviewed, and their age specifically). The reason for this can best be explained by Sali Tagliamonte and William Labov's model of Incrementation (Labov, 2001). The model of Incrementation states that an individual acquires language from their primary caregiver (usually the mother). But as an individual goes to school and becomes part of the community, the frequency of innovating forms tends to increase, right up until late adolescence. At this point, individuals tend to stop increasing their use of incoming forms as both their grammar and vocabulary tend to become more fixed; and this is usually around the age of twenty-five (Tagliamonte, 2016:3-4). And for these reasons I therefore decided that my target group of interviewees should be between seventeen and twenty-four.

In terms of the types of people I wanted to interview; I decided that the target group should reflect the multi-lingual and multi-cultural society of both Copenhagen and

Denmark and was therefore interested in speaking to both native Danes as well as individuals of mixed ethnicities and with a knowledge of multiple languages. I feel that this choice will help to demonstrate to what extent this is a global phenomenon, or whether it is in fact specific to certain individual speakers. If, for instance, the variable was extremely frequent within the oral repertoires of many individuals of multi-cultural and ethnic backgrounds; this would to a certain extent add weight to the argument that what is occurring concerning the variable is happening on a global scale.

I also decided that I wanted to interview both male and female individuals. The main reason for this being that from what I had read on the topic, gender seemed to be a crucial aspect when discussing the word *like*. I therefore decided that my interviewees should be as close as possible to an even split of male and female, depending on who was available and willing to speak to me. After much networking and contact with young adolescents and teens, I managed to acquire and interview 4 boys and 6 girls. The transcriptions of which have been attached in the appendix. Directly below is a list of the names of my interviewees with some general information about them.

### 2.2.1 Table 1: The interviewees:

NAME	AGE	NATIONALITY	OCCUPATION
HARRY	21	AMERICAN/DANISH	UNEMPLOYED
ANDREW	19	DANISH	STUDENT/BARTENDER
JULIE	19	UZBEK	STUDENT
KIRA	18	AMERICAN/DANISH	STUDENT
RITA	24	DANISH/TURKISH	STUDENT/TEACHER
KIERON	19	DANISH	STUDENT
CHRISTOPHER	24	DANISH/ENGLISH	STUDENT/NIGHT CLUB MANAGER
IDA	17	SWEDISH/RUSSIAN/DANISH	STUDENT
MICHELLE	24	DANISH	STUDENT/MEDIA CONSULTANT

LAURA	23	DANISH	STUDENT/MANAGER AT ROYAL ARENA
-------	----	--------	--------------------------------

## 2.3 The Interview Process

As previously mentioned, the interviews were designed in order to elicit informal and authentic forms of English, so the types of conversations that would take place between people who were familiar with one another, and were possibly even friends. The interviews were therefore conducted in the following manner. Each interviewee was asked to give a brief presentation of themselves in terms of who they are and what their occupation was. Once they had done this, they were then asked a series of questions about their lives pertaining to their respective occupations; as well as questions about friends; family members; hobbies; part-time jobs; social media; as well as follow-up questions on each of these. They were also asked to tell anecdotes or stories relating to these topics, as this was believed to be a good way to encourage the use of the variable. However, the structure of each interview varied considerably in terms of what topic was spoken about.

Because all of the individuals I spoke to are quite diverse and different, in a sense everyone had different areas that they were interested in talking about; and so as an interviewer I always tried to focus on the topics they spoke passionately about, and tended not to dwell on the ones that caused them to go quiet. After all I was trying to get them to speak as much as possible without forcing the issue too much, and generally speaking I feel that this was a success as I managed to acquire 1866 linguistic tokens in total.

In terms of overall patterns that the interviews followed, the initial section where they would present themselves tended to be the most formal part of the interview; and probably the part where each interviewee paid the most attention to the way they spoke. Most of them were in fact students at the time of the interview, so they often presented themselves using quite formal and academic language. Once some of the more informal topics were introduced, and also as they began to realise the type of laid-back conversation that was taking place, I would say that a majority of them did relax and start to communicate with me as they would their peers. In this sense I feel that the interviews were successful. Even though I was not a close friend to any of the people I interviewed, many of whom I was in fact meeting for the first time at the time of the interview. Of course the method I used is in a sense limited and could perhaps have been improved upon, had I for instance interviewed people that I was closer to, in that way the tone could have arguably been more authentic. However, given the fact that I am considerably older than my target group and the fact that I do not have many close friends within that specific age group, I feel that my method for conducting the interviews was absolutely as successful as it could have been in the

circumstances.

## 2.4 Transcriptions/Analysis

For ethical reasons, the interviewees have all been anonymised in the sense that their real names have not been used, and each has been given a 'new' name for the sake of my work. Once the interviews had been transcribed, the next step and initial part of the analysis began. This entailed going through every single sentence and utterance and extracting every single instance of *like* from the interviews. For comparative purposes I also decided to include my own examples of the variable as the interviewer. Overall, 1866 linguistic tokens of the *like* variable were gathered throughout the ten interviews, which themselves represented approximately eight and a half hours of conversation between myself and my interviewees.

My initial data collection resulted therefore in a single spreadsheet containing all 1866 linguistic tokens of *like* (i.e. sentences or utterances where the word *like* was used by either interlocutor); where each token was categorised in terms of name, age, gender, and grammatical class. The reason for this being that these were the main factors I was interested in testing the variable in relation to. The next step of this analysis concerned the two most frequently used and innovative grammatical functions of *like*: the marker and the particle. In order to attempt to answer my problem formulation, I decided to gather every token elicited from my interviewees that was either a marker or a particle. A total of 256 discourse markers and 862 discourse particles were collected throughout the course of my interviews. These were then categorised in terms of phrase or clause type in the surrounding environment (i.e. what followed the *like* token from the interview). The spreadsheets for all of these have again been included in the appendix section of this project.

Once all of this data had been collected, it was then analysed using the program Rbrul, which is a program specifically designed for the statistical analysis of linguistic data. In terms of what it can do, Rbrul is capable of performing cross-tabulations; carrying out multiple regression (one-level, step-up & step-down) with unbalanced binary data; it reports effects in log-odds (as well as factor weights); reads common file formats; and has no limit to the number of factor groups or factors per group. It can also support continuous predictors such as age or lexical frequency (both of which I am interested in as far as the variable is concerned), as well as continuous responses like vowel formant measures. It fits mixed models with random effects in order to take into account by-speaker and by-item correlations; and can also estimate between-group effects like gender, and within-group effects like the individual speaker (Rbrul manual, Daniel Ezra Johnson). The results of this quantitative analysis will be presented in chapter 5 of this project.

### 3. THEORY

#### 3.1 GRAMMATICAL CLASSES OF THE LINGUISTIC VARIABLE *LIKE*

In the following section, I would like to start by describing all of the different ways that *like* can be used in a sentence: its grammatical functions. These range from both the standard and fairly unremarkable uses that are generally considered accepted forms of written and spoken English, to the more innovative and remarkable forms that are often commented upon and considered incorrect or questionable when used in either written or spoken English. All of these have been numbered to reflect how they were classed in the analysis section of the project, so if one wishes to look at the classification of them within the analysis they will match and be consistent.

##### 1. Quotative *like*

The quotative *be like/was like* is an extremely interesting and in my opinion innovative function of *like*. As we will see in chapter 4 of my work, it is generally considered to be something that is stereotypically associated with American speakers of English, as well as women in general. The people who use it are often considered as poor or sub-par speakers of English and excessive use of quotative *be like* within oral speech is generally not viewed as something positive or desirable.

A quotative can be described as a grammatical device that is used to mark quoted speech in some languages, in this case English. As such it preserves the grammatical person and tense of the original utterance rather than adjusting it as would be the case with reported speech. It can be equated with "spoken quotation marks". In her book *Discourse-Pragmatic Variation in Context: Eight Hundred Years of Like*, Alexandra D'Arcy states that the grammatical function of *be like* is to introduce constructed dialogue, the reaction of thought, speech, sound and action in the voice of oneself or another. She goes on to describe the quotative function of *like* as one of the most recent and rapidly-growing innovations, and also suggests that historically-speaking the quotative in fact developed from the marker. This occurred when the latter began to co-occur with quotative *be* in the vernacular and will be discussed in the historical context section of my work (D'Arcy, 2017:16).

In terms of how it is used by speakers, the *like* of *be like* does not itself have verbal characteristics. It is *be* that carries tense and other inflectional information such as person, number and aspect, though it does not behave synthetically like a main verb. It instead behaves like an auxiliary. For example it undergoes subject-auxiliary inversion, it cannot co-occur with periphrastic *do*, and it allows adverbs to its right, among other things (D'Arcy, 2017:17).

Below is a list of some examples from my corpus of my interviewees' linguistic

tokens containing quotatives:

1. Once I saw the first documentary I was **like**: “Okay I'm gonna try this”- *Harry (token 23)*
2. And I was **like**: “What are you talking about?”- *Julie (token 316)*
3. So I was **like**: “Okay I'm gonna be a vegetarian.”- *Kira (token 688)*
4. And he was **like**: “It was the only job I was able to like, take!”- *Rita (token 706)*

## 2. Discourse particle *like*

The discourse particle is another interesting and innovative function of *like*, and my analysis of my interviewees' uses of discourse particles will form the main focus of my project in general. The discourse particle is not new, but typically features frequently in the oral speech of young adolescents, and is also stereotypically attributed to female speakers. One of my main aims in interviewing young people between the ages of seventeen and twenty-four of different genders was in fact to test how significant age and gender are in terms of people using discourse particles in their speech. As we will see in the analysis, my interviewees' uses of discourse particles will play an important role for this thesis.

The primary function of the discourse particle is to signal subjective information, which as Alexandra D'Arcy describes, ranges “from the speaker's epistemic stance toward the form of utterance, be it hedging or mitigating authority, non-equivalence, to highlighting or focusing the following information. Discourse particles are also fundamentally interpersonal in that they establish common ground, solidarity, or intimacy between interlocutors” (D'Arcy, 2017:15).

Another important function often attributed to the discourse particle is that of filler, where instead of using small utterances such as *um*, *er*, *mmm*, a person will use *like* instead. Although one would traditionally think that this is a sign of hesitation or a lack of ability to express oneself, interestingly D'Arcy does not believe this to be the case. She in fact states that: “Although *like* sometimes co-occurs with hesitations, false-starts, or pauses, phenomena that indicate indecision, retrieval, or complexity (e.g. processing constraints), such instances are not reflective of the broader distribution of the particle. Indeed, all elements co-occur with these phenomena to some extent. If *like* is sometimes a filler, this function is at best marginal. It is also extremely rare that *like* should be surrounded by pauses. In the majority of cases, particle *like* does not mark disfluency. Rather, it is prosodically integrated in the intonation unit rather than disjunct from it.” (D'Arcy, 2017:15-26)

Below is a list of some examples from my corpus of my interviewees' linguistic tokens containing discourse particles:

1. Erm, I think with **like** my close friends it's not a problem. -*Kira (token 649)*
2. Well yeah I can, but not **like** as consistently! -*Kieron (token 886)*

3. But also the internal **like** communication of the company. -*Christopher (token 1067)*

4. I haven't ever **like** thought about it! - *Ida (token 1186)*

Lastly, the discourse particle does not relate the words around it to other words in the discourse. Its function is instead to bring something out for the benefit of the audience, such as how we feel about a certain thing, or what parts of the utterance we want our audience to focus on. Its syntactic position is nowhere near as limited as that of the discourse marker. The discourse particle can appear in many different parts of the clause, while the marker is only found before the clause.

### 3. Adverbial *like*

In its adverbial function, *like* modifies a verb to signify similarity and/or comparability. This is a fairly standard and accepted form of use for *like* within spoken English, although it is more common in the vernacular and again is often attributed to younger speakers. When used in this way *like* functions as an adverb of approximation and can be alternated with *about*, *roughly*, *approximately*, *around*, and other words of this nature.

Interestingly, Alexandra D'Arcy states that there is a subtle difference between *like* and these other adverbs of approximation; and that they are in fact not quite the same as *like*, as we can see from the following quote: "In other words, whereas *roughly*, *approximately*, and *about* operate at the prepositional level, *like* operates metalinguistically, signalling a lack of commitment. Under this analysis, *like* is not functionally equivalent to adverbs of approximation (i.e. they are not lexical variants)." (D'Arcy, 2017:10)

Below is a list of some examples from my corpus of my interviewees' linguistic tokens containing adverbial *likes*:

1. It takes **like** an hour and a half to get there so I don't really like see him as much now. - *Ida (token 1212)*

2. Cos I think it's **like** 50-50. - *Laura (token 1803)*

3. I played drums for **like**, 12 years or somethin'. - *Kieron (token 953)*

4. I went to a French school in London, when I was **like**, 6. - *Rick (token 182)*

### 4. Tag *like*

A tag *like* functions similarly to a preposition or conjunction and can be described as any use of *like* that comes at the end of a sentence. Again this is a fairly standard and generally accepted use of *like* within spoken forms of English and won't be dealt with in great detail within the analysis of the project. Below is a list of some

examples from my corpus of my interviewees' linguistic tokens containing tag *likes*:

1. I couldn't read it or anything like that, no way! - *Rick (token 187)*
2. "Oh that's fun 'cos he does it and stuff **like** that." - *Andrew (token 150)*
3. Just like the fact that crossing borders and something **like** that? - *Rita (token 730)*
4. "Well we could put up like a stand at a grocery store where you can get smileys and stuff **like** that." - *Laura (token 1856)*

## 5. Discourse Marker *like*

As a discourse marker, *like's* main function is to organise discourse into segments by managing the flow and structure of discourse. It signifies exemplification, illustration, clarification or elaboration. This is usually done by relating the current utterance to prior discourse, or more specifically something that was said in a previous sentence. When used as a discourse marker, *like* can be alternated with *so*, *because*, *and*, *but*, and other words of this nature. Discourse marker *like* is not something that has developed recently, although it has increased in frequency over the second half of the twentieth-century and is widely in use in oral speech today (D'Arcy, 2017:14). Below is a list of some examples from my corpus of my interviewees' linguistic tokens containing discourse marker *likes*:

1. **Like** why have plastic islands around the world... That is just... - *Laura (token 1866)*
2. **Like** the menu changes basically every week or every month so you just have to follow that... - *Michelle (token 1500)*
3. **Like**, New Year's eve, there was a lot of people! - *Christopher (token 1081)*
4. 'Cos I'm like that in a text, **like** I'll give you an example. - *Rick (token 280)*

## 6. Sentence Adverb *like*

As a sentence adverb, *like* provides commentary on the preceding statement and helps to signal to the listener that the proposition being discussed only resembles or approximates reported events. It may also signal the end of old information or mitigation and can be alternated with *so to speak*, or *as it were*. It is not considered widespread across all forms of spoken English, but rather is more likely to be found in regional dialects specific to certain parts of the world such as England, Ireland, and North America (D'Arcy, 2017:12). Below is a list of some examples from my corpus of my interviewees' linguistic tokens containing sentence adverb *likes*:

1. "I've done IB for so many years, like 6/7 years **like**..." - *Julie (token 411)*
2. ...so he was always like, my like, father figure **like**. - *Kira (token 591)*



3. In total **like**. - *Rita (token 734)*
4. ...he tried to get into it as well, **like**. - *Kieron (token 854)*

## 7. Verbial *like*

The use of *like* as a main verb is fairly straight-forward and simple to comprehend. It has been used this way since Old English, where *like* meant “To please/ be sufficient”. Below is a list of some examples from my corpus of my interviewees' linguistic tokens containing *like* as a verb:

1. When I was a kid I didn't really **like** vegetables and stuff like that. - *Rick (token 13)*
2. I do like to occasionally watch funny videos, - *Harry (token 45)*
3. I'm a person who doesn't **like** change even though I've been through so much change and maybe that's why I don't **like** it. - *Julie (tokens 403 and 404)*

## 8. Comparative Complementizer *like*

As a comparative complementizer, *like* introduces a finite subordinate clause, in order to express feelings of similarity or comparison. It usually occurs with appearance/perception verbs such as *seem*, *appear*, *look*, *sound* and *feel* and can be alternated with *as if*, *as though* and *that*. This function of *like* is a fairly recent innovation of spoken English (D'Arcy, 2017, P.7). Below is a list of some examples from my corpus of my interviewees' linguistic tokens containing *like* as a comparative complementizer:

1. Because I feel **like** I don't wanna lose something I already have. - *Julie (token 400)*
2. Well it sounds **like** you won't... - *Rick (token 402)*
3. It may sound a little **like**, not so serious, but it's you know it's like hard! - *Ida (token 1274)*
4. Yeah, so I don't feel **like** it has such an impact on, or like my study has such an impact on what I do now... - *Michelle (token 1441)*

## 9. Other uses (Preposition, conjunction)

The final class category I used for my analysis was an amalgamation of *like* as both a preposition and a conjunction. The main reason this was done was primarily because these two classes were not a focal point for my analysis; but also because the two resemble each other and it seemed logical to group them together. As both a preposition and a conjunction, *like* embodies meanings of similarity and comparison

(D'Arcy, 2017, P.6). The main difference between the two is that conjunctions join words, clauses and sentences together, whereas prepositions express relations between parts of a sentence. As a conjunction for instance, *like* alternates with *as* to link clauses together. Neither are considered remarkable or linguistically innovative and both are frequent in vernacular spoken English. As a conjunction, *like* can be alternated with *but, and, yet, or, because, since, unless, while* among others. Prepositional *like* can be alternated with *about, after, instead of, besides*; as well as words that describe location or placement such as *at, for, in, off, on, over, under*. Below is a list of some examples from my corpus of my interviewees' linguistic tokens containing *like* as a preposition/conjunction (as they were classed and categorised together for the sake of my analysis):

1. So it's more **like** a rave kind of thing? - *Rick (token 1605)*
2. It's kind of **like** riding a bike, so like, I just remember the movements. - *Michelle (token 1616)*
3. So I was just treating him **like** a regular customer. - *Laura (token 1762)*
4. But high key is **like** you're super-interested and you need to know right now! - *Julie (token 311)*

There are of course other grammatical functions that *like* can serve. It can for instance be a noun, an adjective, or even a suffix. However, I will not spend time describing these as they are not significant for the purposes of my work and are in fact quite rare, especially in spoken English. None of them were in fact present in my interviews and although no tokens were collected of any of these, their existence adds weight to the argument that the variable has a well-defined grammatical structure.

### 3.2 THEORIES CONCERNING STRUCTURE OF THE MARKER AND THE PARTICLE

Now that I have described the basic grammatical functions of *like*, I would now like to present some of the main theories pertaining to it. The starting point for this is Alexandra D'Arcy's *Discourse-Pragmatic Variation in Context: Eight hundred Years of LIKE*. As mentioned in the introduction, the main aim of this book is to demonstrate that the way people use *like* in spoken (and written) English is in fact not chaotic and haphazard as so many people believe it to be. And “that a linguistically-informed perspective utterly undermines any claim that LIKE, in any of its uses, is random and meaningless” (D'Arcy, 2017:31).

The main aim of my thesis therefore, will be to test to what extent this statement is true. As we have seen in the previous section, *like* is an extremely versatile word which can display a myriad of different grammatical functions, many of which appear to convey very similar meanings (such as comparison and similarity). I would therefore like to present some of the main theories from the book which concern the

grammatical placement of *like* within a sentence or clause, in order to show some of the factors that encourage and constrain my interviewees' uses of *like*. For the purpose of my thesis, I will be focusing on the two most innovative grammatical functions of *like*: the discourse marker, and the discourse particle.

In Chapter 4 of the book, *Developmental context*, D'Arcy addresses the synchronic development of *like* in apparent time. This is done by drawing on a large corpus of contemporary speech data in order to “illustrate the systematic and rule-governed generalization of the marker and the particle across contexts over the course of the twentieth-century” (D'Arcy, 2017:68). Although the corpus I have collected from my interviewees is much smaller in comparison and focuses on one point in time as opposed to many different ones; I believe that the theories and findings of this study can prove useful in terms of demonstrating the structure of both the discourse marker and particle, and in helping to define the envelope of variation for my own analysis. I would therefore like to present some of the key points from the chapter in the next section of my project.

## **1. The Syntactic Positions of *like*:**

In order to help define the envelope of variation for the linguistic variable *like*, let's now examine some of the syntactic positions of the variable, as discussed in chapter 4 of *Discourse-Pragmatic Variation in Context: Eight hundred Years of LIKE*.

According to D'Arcy, this form of analysis is standard practice in diachronic variation in change and by examining the position of the variable with regard to other elements within the phrase (such as adverbs, particles, pronouns etc...), that these forms serve as a baseline for situating the locus and the nature of variation (D'Arcy, 2017:71). For the sake of my work, I am interested in focusing on the syntactic positions that either encourage or constrain the use of *like* as either a discourse marker or particle. These, along with age and gender, will serve as baselines for situating the locus and nature of variation in my analysis. Once these have been outlined, I will then see to what extent my interviewees uses of *like* correspond or differ from these and attempt to examine why that may be in my analysis section.

As we have already seen in the Grammatical functions section of this project, *like* can occur in a number of syntactic positions; and the three most frequent of these are: clause initial, before a noun phrase and before a verb. According to Alexandra D'Arcy, these distributional patterns reflect diachronic factors and they in fact correspond to the developmental trajectories of *like's* discourse uses. What is perhaps most salient is the way that this chapter explains how the linguistic variable *like* has changed and evolved over time, spreading from outside the clause to inside the clause (D'Arcy, 2017:68). But now let's outline some of the key theories and findings from the chapter pertaining to the syntactical positioning of *like*.

## 2. Complementizer phrases:

According to Keller and Traugott, Any complementizer phrase, regardless of level (i.e. matrix or subordinate), has the potential to host *like*. Although the sentence initial position is considered the most likely place where one would find discourse markers in English, not every complementizer phrase is in fact sentence initial. More specifically, a sentence complementizer phrase is not restricted to the top-most projection of syntactic structure. Not only may it be embedded as a subordinate clause, but various kinds of movement can disrupt the canonical order of sentential elements through topicalization (Keller & Traugott in D'Arcy, 2017:72).

## 3. Enumerations/ Responses to direct questions:

Enumerations and responses to direct questions are two discourse contexts that inhibit and constrain the use of marker *like*. In enumerations for example, the marker *like* can be used on the first clause introducing the unit, but it does not surface on subsequent clauses within the unit. The first *like* clause is not a response to the question but an orientation. In other words, this clause is used to establish the context for the speaker's answer, and *like* signals this broader discourse function. Considered in discourse context, this constraint seems fairly natural given the function of the marker (exemplification), and from its domain of scope, where *like* is placed between sequential units. The belief is that what follows the initial claim in the utterances is a series of examples where *like* will not be present. As such, the speaker can mark the first in the series overtly, with *like* or *you know* or *for example* etc. However to use marker *like* on the clauses within an enumeration would not make sense in the context of the discourse in terms of either the speaker's meaning or intention. In contrast to enumerations, which categorically exclude *like* within within the list sequence, the marker is attested in responses to direct questions, but such uses are rare (D'Arcy, 2017:74-75).

## 4. Relative/ Subordinate clauses:

Two further contexts that constrain and prohibit the discourse marker are relative and subordinate clauses. In some of the earlier data collected in Alexandra D'Arcy's analysis, the marker is categorically absent with non-restricted relatives, and is also exceedingly rare with restricted relatives. D'Arcy theorises that the constraint against *like* in these structures is neither discourse-based nor syntax-based and that it appears to reflect developmental factors (D'Arcy, 2017:76). Specifically, that although these types of structures are not found in the earlier data, examples of both restricted and non-restricted markers in relative clauses can be found in more recent sociolinguistic interviews. She goes on to demonstrate that the ongoing development of this form

raises the possibility for the emergence of new syntactic positions through grammaticalization and states that this is exactly what happened throughout the twentieth-century (D'Arcy, 2017:76). In terms of my analysis, it will be interesting to see to what extent this is true in terms of the presence or absence of both discourse marker and particle *likes* introducing relative and subordinate clauses. If what D'Arcy proposes is true, we can expect to find tokens such as these within my own interviewees spoken English.

## 5. Determiner phrases:

According to D'Arcy, the use of particle *like* is exceptional in determiner phrases, especially for those that form the complement of an adverbial phrase. This in turn helps demonstrate the effect of syntactic structure on the use of discourse *like*, that the particle favours arguments rather than complements (D'Arcy, 2017:P.79). D'Arcy goes on to build the case that arguments, as high frequency constructions, undergo change faster than phrasal complements, which are lower frequency constructions. Therefore, the disfavouring effect of contexts such as these derives from developmental factors rather than a prohibition residing in the structure of language change (D'Arcy, 2017:80).

A further prediction she makes is that *like* will be less frequent in definite determiner phrases, since these will often, but not always, correspond to known information. Indefinite determiner phrases, which include both the indefinite article and the indefinite use of *this*, are also expected to host a higher frequency of *like* tokens. The main observation that she concludes from her results is that “argument, indefinite and modified determiner phrases favour *like* over complement, definite and unmodified determiner phrases in every age cohort” (D'Arcy, 2017:96-98). What Alexandra D'Arcy has basically established is that although younger speakers use the particle more frequently than older speakers, in the case of the determiner phrase, *like* follows a well-established structure that was already visible in the grammar of older members of the speech communities that she interviewed. In essence she is saying that age may in fact not be a massive constraint that prohibits the use particle *like* in this context.

## 6. The Adjectival Domain:

In terms of placement within the adjectival domain, D'Arcy proposes two possible patterns where *like* can be situated between the verb and the adjective. The first being: (adverb) *like* adjective; and the second being: *like* adverb adjective. Overall, she states that the results from the adjectival domain indicate that the overall proportion of *like* is increasing in apparent time, and that an important factor linked to this is the spread of the particle beyond the verb *be* to other verbs that sub-categorise

for an adjective (D'Arcy, 2017:102).

## 7. The Verbal Domain:

As a discourse particle, the linear order of *like* relative to verbs is highly fixed. It categorically occurs to the immediate left of the lexical verb. When functional morphemes such as modal verbs, auxiliary verbs, and infinitival *to* are present, *like* appears between these and the main verb. This is a long-standing observation and not just one that D'Arcy proposes; though this does not rule out the possibility of new adjunction sites as *like* continues to grammaticalize. As discussed above however, *like* is categorically absent before finite inflected *be* within some of the earlier data that D'Arcy collected (such as the TEA) and yet categorically restricted to a post-verbal context with non-finite *be*. (D'Arcy, 2017:102).

A further prediction that emerges from this analysis is that *like* can target the periphery of any verb phrase. Thus, in bi-clausal complexes such as control structures, it should be possible to find instances of *like* adjoining to the higher verb phrase, to the lower verb phrase; as well as to both (D'Arcy, 2017:103). For the sake of my analysis, I will be interested in testing to what extent this structure of *like* matches with my interviewees' spoken English and tokens of *like* with regard to the verbal domain.

## 8. The Adverbial Domain:

As far as adverbs are concerned, D'Arcy states that if *like* has a fixed position in the verbal domain and adverbs can occur in a range of structural projections, then it is entirely plausible that the order of these elements should vary depending on the type of adverb with which *like* co-occurs. Specifically, an adverb that occurs higher in the structure will, in the unmarked case, precede *like*, while one that occurs lower in the structure will follow it (D'Arcy, 2017:104). D'Arcy believes this to be a fairly systematic occurrence with two fairly basic patterns:

1. Speaker- and subject-oriented adverbs tend to appear to the left of *like*, suggesting that they are situated in higher syntax.

e.g: "He actually *like* stood up."

"They honestly *like* threatened me."

"They like it but they never *like* played."

2. Manner and degree adverbs tend to appear to the right of *like*, suggesting they are situated lower in the syntax.

e.g: “But people will *like* slowly get into it.”  
“Some people *like* totally fell into the mould.”

As we can see from all of this, *like* as both a discourse marker and particle does in fact display a degree of grammatical structure, and appears to be evolving and changing quite drastically over time. Although many people still believe it to be random and chaotic, my aim in this section has been to show that this is in fact not the case, and that the envelope of variation for the variable in question is in fact fairly structured and well-defined. Now that I have outlined some of Alexandra D'Arcy's theories and findings regarding the syntactic positions that promote and constrain the use of *like* as both a discourse marker and a discourse particle; let's now examine some other theories pertaining to the other most innovative use of *like*: the quotative.

### 3.3 THEORIES INVOLVING QUOTATIVE *BE LIKE*

In order to help define the structure for quotative *be like*, I would now like to present some of the main theories from Sali Tagliamonte, Alexandra D'Arcy, and Celest Rodriguez Louro's *Outliers, impact, and rationalization in linguistic change* from 2016. As far as the envelope of variation is concerned, the quotative is linked to constructed “dialogue, the reformation of thought, speech, action, non-lexicalized sounds, and gestures in the voice of oneself or another. This type of 'speech within speech' may be introduced by a host of verbs such as *say, think, go, be like, tell, and cry*; or even by nothing at all” (Tagliamonte, D'Arcy, Rodriguez Louro, 2016:828). Below is a list of the three main constraints that have a probabilistic effect on the use of the quotative *be like* within speakers of English:

#### 1. Grammatical person

Grammatical person is a significant predictor for *be like*, and it is favoured with first-person subjects in the reporting clause (Tagliamonte, D'Arcy, Rodriguez Louro, 2016:829). In general, an individual will most often use it to tell a story about themselves as opposed to another person as we can see from the following examples of my interview corpus.

- a. I'm **like**: “This is completely needless... It's unnecessary.” (*Harry, token 20*)
- b. And I was **like**: “what are you talking about?” (*Julie, token 316*)
- c. And I was **like**: “Then university!” (*Rita, token 791*)

#### 2. Content of the quote

Throughout its history, *be like* has commonly been associated with both internal dialogue or attitude, as well as non-lexicalized sounds. Although *be like* is not pragmatically restricted, its association with reported thought has been the strongest predictor with respect to the nature of the quote. Furthermore, current analyses suggest that *be like* entered the repertoire of spoken English in order to encode first-person internal dialogue; as this reporting mode gained prominence for direct quotation in narrative monologues (Tagliamonte, D'Arcy, Rodriguez Louro, 2016:829). Below is a list of such examples from my interview corpus.

- a. 'cos if they yell at you for nothing it's **like**: "I'm 17, like what!" (*Ida, token 1345*)
- b. But it's still weird to know that like nobody's safe because you're **like**: "even if you're a guy you still get drunk" (*Michelle, token 1544*)
- c. So I'm **like**: "Alright I'm looking forward to in 3 years but yeah!" (*Rita, token 752*)

### 3. Tense/Temporal Reference

Since it first rose to prominence and started increasing in frequency, *be like* has been strongly associated with the present tense. However, studies from 2014 onwards have highlighted the importance of distinguishing between surface morphology and temporal reference. In essence, forms that are morphologically and referentially present tense are very different from those that display present surface morphology but refer to past tense situations (Tagliamonte, D'Arcy, Rodriguez Louro, 2016:829). Below are two examples from my interview corpus, the first being referentially present tense, and the second being referentially past tense .

- a. And I'm **like**: "Yeah I'm talking to you!" (*Rita, token 761*)
- b. And they were **like**: "Are you talking to me?!" (*Rita, token 760*)

### 4. Quotative *be like*: A Black Swan Event

The final theory I would like to address concerning quotative *be like*, is attributed to Sali Tagliamonte, Alexandra D'Arcy, and Celest Rodriguez Louro. In *Outliers, impact, and rationalization in linguistic change* they propose that the spread of *be like* throughout the English language is in fact a 'black swan event'. This term is in turn defined as a kind of sudden happening or development such as the internet, social media, September 11, and even the corona virus of 2020. A black swan is an outlier, unlike anything that has preceded it or could predict its occurrence, or as Taleb states: "Black swans are a surprise that change something irrevocably and can only be explained with the value of hindsight" (Taleb, 2010:10, in Tagliamonte, D'Arcy, Rodriguez Louro, 2016:842).



The validity of the black swan event may indeed be a matter of debate. However, what is interesting for the purpose of my work is the way Tagliamonte, D'Arcy, and Rodriguez Louro suggest that language-based phenomena are susceptible to, and show potential to be, such events. Indeed, language is systematic, constrained, regular, and probabilistic; but it is not entirely predictable or deterministic. In the case of quotative *be like*, its development has been shown to be simultaneous, instantaneous, and parallel in multiple urban locations, but yet could not have been predicted. And its emergence and development have completely changed the way people tell their stories to one another. It is now also the majority form for speakers born after the 1970s. The use of quotative *be like* is therefore described as a linguistic black swan event, including both elements of randomness and non-linearity; and shows that language embodies random events as well as predictable structure. It also ultimately demonstrates how language is adapting to the contemporary world situation (Tagliamonte, D'Arcy, Rodriguez Louro, 2016:842-843).

### **3.4 MISCONCEPTIONS/PRECONCEPTIONS CONCERNING *LIKE***

Now that I have outlined the main theories concerning the three most innovative uses of *like*, I would like to present some of the main misconceptions and preconceptions surrounding the word. The main reason I would like to do this is to show that what I have been saying about the way people perceive the word is in fact based on extensive sociolinguistic work and analysis, and is not in fact hearsay or generalisations based on my own personal experience. Alexandra D'Arcy for instance states that: “A particularly striking aspect of the social context of language change is that from a diachronic perspective, the cumulative effects of change are unexceptional, yet in synchronic time individual changes are synonymous with degradation. It is also well established that language change is always most advanced among young speakers” (D'Arcy, 2017:126)

And as I have already mentioned throughout the project, there seems to be a lot of ambiguity and negative attitudes towards individuals who use *like* frequently within their spoken English. The above quote shows that this is often perceived as a form of degradation of the English language which is often associated with younger speakers in general. Let's now examine some of these preconceptions more closely and see to what extent they may in fact be misconceptions.

#### **1. *Like* is just a meaningless filler word**

In chapter 6 of *Discourse-pragmatic Variation in Context: Eighteen-hundred Years of LIKE*, Alexandra D'Arcy describes a tendency that people have of talking about *like* as a monolithic entity, an all-encompassing word that has no grammatical structure surrounding it. However, as we have seen in the syntactic structure sections

of this project, that is categorically not the case. For instance, I have already written quite extensively about the myriad uses of *like* as either discourse-marker, particle, quotative, adverbial, preposition, verb, conjunction and so on so. We can already see that the variable is performing under a fairly rigid set of grammatical rules and can therefore not be described as a mere meaningless filler. D'Arcy states that “There is not one *like*, but a conglomerate, LIKE. In other words, there are multiple variants and each is part of a distinct variable system. Each relates to a specific function and each function has meaning. In most cases this meaning is not concrete and referential but versatile and pragmatic” (D'Arcy, 2017:129).

In this regard, I am very much in agreement with sociolinguists such as Alexandra D'Arcy who dismiss the notion that *like* is merely a meaningless filler word, or some sort of linguistic crutch that signals hesitancy or a lack of ability on the individual's part to articulate correctly. I feel that the word is so diverse and at times ambiguous, which in turn causes confusion. But as we can see in the theory section, it does in fact possess a lot of structure, and can therefore definitely not be described as meaningless.

## **2. People who use *like* are inarticulate**

Later on in chapter 6 of *Discourse-pragmatic Variation in Context: Eighteen-hundred Years of LIKE*, Alexandra D'Arcy presents some of the literature addressing *like*, and states that it “typically characterises it as marking loose talk or vagueness, where vague is typically interpreted as a negative linguistic trait rather than as a skilled interactive practice, and attitudes towards it reflect the belief that LIKE is symptomatic of careless or meaningless speech (Newman, 1974; Schorup, 1985 in D'Arcy, 2017:131-132).

Again, D'Arcy goes on to argue that the reality of the situation is in fact very different from most of these preconceptions. As a discourse marker for instance, *like* signals the sequential relationship between units of discourse, and its use is described as “essential to the rhetoric shape of the argument of the narrative” (Traugott & Dasher, 2002:154 in D'Arcy, 2017:132). As a particle, *like* establishes common ground, solidarity and intimacy between interlocutors. As a quotative verb, it enables the speaker to encode the first person internal monologue as narrative, and is considered a mode of story-telling that has been increasing in spoken English quite drastically since the latter half of the twentieth-century. It is currently extremely common in vernacular English. As an approximative verb, *like* functions as a lexical alternative for other adverbs such as *about*, *roughly*, *approximately*, and *around* just to name a few. As a sentence adverb, *like* is embedded in online discourse often signalling structural cues to guide the listener (Corrigan, 2010 in D'Arcy, 2017:132).

It therefore becomes evident that in all of its grammatical functions and through all of its different uses, *like* requires as D'Arcy rather eloquently puts it: “activation of a skilled set of discourse-pragmatic and language internal constraints. Its use is

strategic, not random. This is the antithesis of inarticulate speech” (D'Arcy, 2017:132). I must admit that the more that I have read on the subject and worked with this topic, the more I begin to agree with Alexandra D'Arcy. Although it is rather tempting to dismiss people who use *like* in its many different grammatical guises as inarticulate, it is in fact a little too simplistic when one examines the many different and quite specific ways and contexts in which *like* can be used.

### 3. Women use *like* all the time

Another widely held misconception concerning *like* is that women use it far more than men do (Daily-O'Cain, 2000, Hesson & Shellgren, 2015, in D'Arcy, 2017:136). Alexandra D'Arcy argues that this view is far too simplistic and the evidence that she has gathered based on all of her sociolinguistic interviews is in fact rather different from this particular preconception. Specifically, that distinct gender patterns are attested across functions of *like*. The quotative and the marker are for instance probabilistically associated with women, whereas the particle is more frequent in men. As we will see in chapter 5 of my work, this was not found to be the case as far as my interviews are concerned. My analysis in fact demonstrates a pattern whereby the discourse marker is used most frequently by the male participants, and is more likely to occur as they increase in age. Both the quotative and the particle on the other hand were forms that were more frequently favoured by the female interviewees. There is also a convergent pattern concerning age whereby the younger female interviewees are more likely to use the particle. The results of my analysis therefore differ from what D'Arcy is saying about gender, because the marker was found to be a form favoured by males and the quotative and particle were favoured by the females, though gender did prove to be a crucial predictor overall for my interviewees.

Whereas the marker, particle and quotative show distinct patterns concerning gender, the approximative adverb and the comparative complementizer on the other hand exhibit no gender conditioning at all. Therefore, even though popular opinion seems to view women as the main users of *like* in spoken English, the truth actually depends on what specific function of *like* we are assessing. This again goes on to add weight to the argument that *like* is not some sort of haphazard filler word with no grammatical structure; as D'Arcy again rather nicely states: “If *like* were just *like*, all manifestations would be similarly constrained by gender. They are not” (D'Arcy, 2017:137).

As we can see the reality behind this preconception that *like* is a word attributed to female speakers of English rather than male ones may in fact be a little more complicated and nuanced than one might imagine. My analysis will in fact demonstrate that factors such as age, gender, frequency, and grammatical structure all help to inhibit and constrain the way *like* is used within my interviewees' spoken English and that this cannot simply be attributed to females in general. It is both too

simplistic and at times completely incorrect.

#### 4. *Like* is a Valley Girls/adolescent phenomenon

One of the main questions that arises from all of these different uses of *like* is where do all of these uses stem from? Who and what started it all? Much of this will be considered in the historical context section of this project, but in chapter 6 of *Discourse-pragmatic Variation in Context: Eighteen-hundred Years of LIKE*, Alexandra D'Arcy observes that: "...popular ideology situates the epicenter of *like* in California, and within North America the Valley Girls are attributed with launching it into the general consciousness (Blythe, 1990:224, Dailey-O'Cain, 2000:70, in D'Arcy, 2017:139). It is also commonly associated as being a trope of young adolescent speakers, and something which is usually outgrown after a certain age once a person reaches adulthood.

D'Arcy goes on to argue that historical evidence in fact paints a rather different picture. The marker and particle for instance are not only attested among elderly speakers across North America but also in cities and towns across England, Scotland and Ireland; as well as in New-Zealand and other settler-colonial settings. In summary, they are therefore not a product of either North America generally or specifically attributable to the Valley Girls, since examples of linguistic tokens such as these have been recorded long before then. What D'Arcy is basically saying is that *like* was already available in the vernacular and its connection with certain groups cultivated the appropriate social context for its discourse functions to be used with more frequency and by a wider social circle (D'Arcy, 2017:140).

The fact that *like* is therefore associated with and attributed to younger speakers, particularly American and female, is more easily explained by the fact that adolescents are at the forefront of linguistic change generally. Even though they may very well use some of these innovative forms of *like* such as the marker and particle more than older speakers, they are in fact not the only members of the speech community using them, and are not solely responsible for their emergence and use within spoken English.

In the case of the Valley Girls for example, the excessive use of certain forms of *like* which have been attributed to this particular group around the 1980's could not have been initiated by them since there were many recorded uses in the data collected by sociolinguists at least twenty to thirty years before this time. A more likely explanation is that this is some sort of recycled phenomenon as can be seen in the following quote: "...the saliency of social categories can be variable across time, and linguistic forms associated with one may later come to be associated with another as each rises to prominence in the cultural landscape of the time" (D'Arcy, 2017:140).

Although many of these innovative grammatical functions of *like* such as the discourse marker, particle, and quotative have stereotypically been attributed to young adolescent and often female speakers, the reality is in fact that the discourse

functions of *like* are not simply an American thing; or a female thing; or even a phenomenon attributable to teenagers. For the most part these forms are present in the everyday spoken English of all speakers from within a community and perhaps the main area of confusion stems from frequency. Maybe the fact that young adolescent speakers (i.e. people of a certain age born after the 1970's) use some of these grammatical forms more than older speakers is what is causing such misconceptions and ambiguity in terms of how they are viewed by the general public. They are, however, clearly not the only speakers using these innovative forms as my analysis will demonstrate.

## 4. HISTORICAL CONTEXT

In the following section of the project, I would like to present a brief history of the word *like*, in order to show how it has changed and evolved throughout time; and also in order to give an idea of how my interviewees will perform when speaking English in the present day of 2020. I will also give a brief overview of the general population currently using English as either a primary or secondary language. This will be done in order to give a broader context in terms of who is using the linguistic variable *like*, and will help to demonstrate the fact that it is not uniquely attestable to primary speakers, but has also spread to individuals who use English as a second language and on a global scale.

### 4.1 Spoken English on a global-scale

*In the role of sociocognitive salience in the acquisition of structured variance and linguistic diffusion: Evidence from quotative be like*, Julia Davydova describes English as the first language in the history of humankind that has expanded as a truly global language. Of course when the numbers are so large and spread over such a vast section of the planet, only rough estimates can be made in terms of speaker numbers. For example, English is spoken as a first language by some 329,140,800 people and as a second language by a further 430,614,500. Additionally, over a billion language users have adopted English as their primary foreign language; which means that there are in fact approximately 2,236,730,000 English speakers throughout the world (Jenkins 2015:2; Galloway & Rose 2015:11-15; in Davydova, 2020:1).

Davydova goes on to describe the categorisation of individuals by making the distinction between ESL (English as a second language), and EFL (English as a foreign language). ESL varieties are described as “part of the colonial heritage of the country and are therefore used in institutionalised national, social, and political settings such as education or the law as well as a primary means of interethnic communication actively shaped through local linguistic practices and daily

interactions” (Davydova, 2020:1).

English as an EFL on the other hand, can be described as “an important means of international communication. A preferred lingua franca in contexts involving interactions between non-native speakers. [...] EFL varieties have a tendency to develop in countries where the majority of the population speak English natively” (Davydova, 2020, P.1). For the purpose of my own work, my interviewees represent both of these groups.

The main difference between the two therefore, is that in education for example, an ESL classroom is one in which English is the primary national language, like the USA for example. Whereas an EFL classroom is one in which English is not the native language. Many European countries such as Denmark are examples of this. Indeed, Denmark is a multi-national as well as a multi-lingual country with a broad mixture of ethnicities, all interacting together and sharing English as a common language of communication.

The majority of my interviewees therefore are speakers of EFL varieties of English since most of them were born, raised, and educated in Denmark. This is the case for seven of the ten interviewees. Kieron, Rita, Michelle, Andrew, and Laura were all born, raised and educated in Denmark. Similarly, Ida has lived in Sweden all of her life and has only studied there, so she can be categorised as an EFL speaker too. Julie has spent time in both Uzbekistan where she was born, and Denmark where she has also spent a majority of her life. She can therefore also be categorised as a speaker of the EFL variety of English. There were only two interviewees who can be categorised as speakers of the ESL variety of English: Harry and Kira. Both of whom were born in the USA and have spent the majorities of their lives living in either the USA or Denmark and also going to school in both of these countries. The final interviewee who I would categorise as a speaker of English as a first language is Christopher. Indeed, Christopher is a unique individual in that he has one Danish parent and one English parent, and speaks both English and Danish perfectly. Although he has been educated in Denmark, I feel it would be incorrect to describe his English as an EFL variety because he was born in England and also educated there from an early age. He also stated in the interview that English was his preferred language of communication on a daily basis and the one he used the most.

#### 4.1.1 Table 2: Categorisation of the interviewees

NAME	CATEGORY OF SPOKEN ENGLISH
HARRY	ESL
KIRA	ESL
CHRISTOPHER	FIRST LANGUAGE
RITA	EFL

IDA	EFL
MICHELLE	EFL
LAURA	EFL
ANDREW	EFL
KIERON	EFL
JULIE	EFL

## 4.2 Discourse marker and particle

Now that I have described some of the general ways that speakers of English can be categorised, and also given some statistics on numbers of speakers of English on a global scale, I would now like to focus more specifically on the discourse marker and particle, and how their use has developed over time.

In chapter 3 of *Discourse-pragmatic Variation in Context: Eighteen-hundred Years of LIKE*, Alexandra D'Arcy describes the majority of the ways in which an individual can use *like* as “longstanding components of the spoken grammar of English” (D'Arcy, 2017:47). She also states that there is a commonly-held assumption pervading many speech communities that *like* is a recent phenomenon, and is commonly viewed as an aspect of young adolescent speakers; as opposed to the older speakers from a given speech community. Interestingly, she believes that the way many individuals view *like* is subject to the *recency illusion* (Zwicky, 2005), which states that “phenomena noticed only recently are in fact recent, when in reality they have been around, with some frequency, for very much longer” (D'Arcy, 2017:47). In the rest of the chapter, D'Arcy draws upon a range of historical, archival, and contemporary corpora from multiple English dialects in order to show that the ways we use *like* are in fact not some sort of recent phenomenon that can be attributed to young adolescent speakers born after the latter stages of the twentieth-century, and that “the pragmatic functions of LIKE are rooted in longitudinal realities and are diachronically regular” (D'Arcy, 2017:47).

In essence, she is attempting to show that the way we use *like* is not some sort of recent development pervaded by the youth of today, but in fact a natural process where linguistic features are transmitted from one generation of speakers to the next. The reason I am including this section in my own work is that I believe that this explanation makes far more sense both on an academic level, and also in my own personal experience, as I will now elaborate on concerning the discourse marker and particle.

As far as the discourse marker is concerned, the historical and archival records presented in chapter 3 of *Discourse-pragmatic Variation in Context: Eighteen-hundred Years of LIKE* provide concrete evidence that its function and use in

contemporary vernacular English is the result of transmission from one generation to the next. In essence, it is diachronically stable when it comes to function, and examples of its use have been collected in D'Arcy's corpora since at least the eighteenth century (D'Arcy, 2017:52). It is clearly not a recent development that has mysteriously been conjured up by young people in the 1960s or 1970s.

As for the discourse particle, the first unambiguous uses of *like* in this way are attested as early as the beginning of the twentieth-century, and again goes some way to show that this is not a recent development. As a particle, *like* has actually been a part of the English vernacular for approximately a century and a half, and perhaps even longer. As with the discourse-marker, many of the examples collected in D'Arcy's corpora are extremely relevant in that they again point to a shared function across speakers born in the late nineteenth and early twentieth-century in very different geographical locations such as western Australia; northern England; south-central inland Canada; as well as the west coast of Canada (D'Arcy, 2017:53).

One therefore cannot help but ask the following: if the discourse particle and marker are not a long-standing feature of the English language, how could one possibly account for such consistent and parallel use on a global scale? It is my belief that they in fact are a long-standing feature of spoken English and that the historical evidence collected in corpora such as D'Arcy's provides irrefutable proof of this generational transmission.

### 4.3 Quotative *be like*

As for the quotative function, Sali A. Tagliamonte describes *be like* as “possibly the most vigorous and widespread change in the history of human language” (Tagliamonte, D'Arcy, Rodriguez Louro, 2016:825). She also considers *be like* to be a fairly recent innovation of the English language (compared to the other grammatical functions of *like*), originating in the United States. The first suggestion that it had spread beyond North America was by Romaine & Lange, who noted traces of a similar development in British English. By the middle of the 1990s, it represented 13% of all direct quotatives in Canadian English and 18% in British English. Although *be like* was clearly increasing in both real and apparent time, in its earliest recorded forms it remained confined to speakers below the age of forty. At that time, Tagliamonte and D'Arcy suggested that North American speakers born in the early 1970s were the first generation of native users of the quotative form of *be like* (Tagliamonte, D'Arcy, Rodriguez Louro, 2016:826).

Tagliamonte, D'Arcy and Rodriguez Louro go on to make some further interpretations based on the analyses of their data. Specifically, that speakers born in the 1960s were the first to use *be like* on a global scale. The next generation, born in the 1970s, took this a step further and accelerated its use; and this was again taken even further by the next generation born in the 1980s. The distributional and statistical models in their work confirm parallel inception in time, parallel



grammatical conditioning, and also parallel developmental trajectories. What they actually demonstrate is that *be like* is a form of linguistic change that has gone from zero to virtual completion with a consistent variable grammar across space in the most telescoped time frame ever documented (Tagliamonte, D'Arcy, Rodriguez Louro, 2016:838).

In essence, *be like* is the only linguistic change that is diffusing on a global scale with temporal simultaneity as well as an attendant suite of parallel internal linguistic constraint; and cannot be explained by traditional sociolinguistic theory such as William Labov's diffusion of linguistic innovation (Tagliamonte, D'Arcy, Rodriguez Louro, 2016:841).

#### **4.4 Conclusion**

I would like to conclude this section by stating that the main aim of this was to attempt to contextualise and place the different forms of the linguistic variable in terms of how they have performed throughout history and time. This has been done both generally; and specifically concerning the discourse marker, the particle, and quotative *be like*. The main reason in doing so was to demonstrate that although the use of the variable has been increasing in frequency in the last twenty years, it has in fact been a fairly consistent and structured part of the English language for much longer. I also wanted to demonstrate that this is in fact a wide-scale global phenomenon and that it should be present (admittedly to varying degrees for each form) and fairly consistent in the spoken English of my interviewees. As I have already outlined in the methodology section, my interviewees are all young adolescents between the ages of seventeen and twenty-four; so my initial hypothesis is that we can expect them to favour the variable in their spoken English. I have also given a brief introduction of my interviewees and categorised them in terms of whether they are EFL or ESL speakers of English. This was done to demonstrate that they are in fact a reflection of the multi-national and multi-ethnic speaker community who all use the English language on a day-to-day basis in Denmark.

## **5. ANALYSIS**

### **5.1 Introduction**

In this section of my thesis, I would like to present all of the data collected from my interviews concerning the linguistic variable. As previously mentioned in the methodology section, 1866 tokens of *like* were collected from my interviews; and this includes examples given by the interviewer as a point of comparison. The linguistic

tokens in question can be defined as any sentence, phrase or utterance where the word *like* was used by any of the interlocutors throughout the course of the interview. These were then placed into a spreadsheet and cross-tabulations were run testing for factors such as age, gender, frequency, and grammatical classification of the variable. The results of these make up the basis of the first section of the analysis.

Once the variable had been tested generally and in all its grammatical functions, the next step involved isolating the two most innovative and frequently used examples of these: the discourse marker and particle. Since these were the two most frequent and arguably the two most interestingly and innovatively used classifications of the variable, my analysis is focused primarily on these two. Indeed, 862 discourse particles were recorded throughout my interviews, making them comfortably the most frequent types of linguistic token of *like* used by my interviewees. An additional 256 examples were discourse markers; meaning that 1118 out of 1866 of all of the tokens were either markers or particles. Again, these two classifications were tested using the Rbrul program in much the same way as the initial set: for patterns primarily involving age and gender.

The third and final section of the analysis consists of a more in-depth examination of the particle and the marker. Specifically, a step up/step down analysis of the two. The main goal of this type of analysis is to find groups which cause the model to change significantly when being added or subtracted. The program therefore tests each factor group and retains the most significant ones, by continually adding groups until no further addition results in a significant change (Tagliamonte, 2006:140). This was also centred around establishing to what extent age, gender, and grammatical form are linked to both the marker and the particle in the examples given by my interviewees.

## **5.2 Primary Statistical Results Data**

The first table that I would like to present consists of the results of my interviewees linguistic tokens classified by gender. In terms of how they were classified grammatically, numbers were used to represent each one and I have enclosed a list below of how they are classified to make them easier to follow and understand.

### **Grammatical classifications of *like*:**

1. Quotative *like*
2. Discourse particle *like*
3. Adverbial *like*
4. Tag *like*
5. Discourse marker *like*
6. Sentence adverb *like*
7. Verbial *like*

- 8. Comparative complementizer *like*
- 9. Other (preposition/conjunction)

**Table 3: Grammatical classification by gender**

CLASS	FEMALE	MALE	TOTAL
1	147	67	214
2	584	278	862
3	65	45	110
4	24	54	78
5	157	99	256
6	10	20	30
7	41	58	99
8	38	45	83
9	39	95	134
TOTAL	1105	761	1866

From the table above, it can be seen that the most frequently used grammatical classification of *like* is the discourse particle (class 2 in table 1), with 862 tokens in total. The female interviewees are leading the way with 584 tokens, and 278 therefore coming from the males. The second most frequent use of *like* within my interviews was therefore the discourse marker (class 5 in table 1) with 256 tokens collected. 157 of these were provided by the females, and 99 by the males. However, it is also interesting to note that the quotative (class 3 in table 1) is in fact the closest to either of these with a total of 214 tokens collected.

In terms of overall tokens, 1105 are attributed to females and 761 to males. Again, my female interviewees are ahead quite significantly in terms of frequency. The next step of my analysis was to categorise the tokens in terms of most to least frequent in terms of gender. The reason for this is because I wanted to look at the males and females as individual groups so they could then be compared and contrasted. The results of this can be seen in tables 4 and 5 directly below.

**Table 4: Grammatical Classification in order of frequency  
Most to least frequent class by gender for females**

CLASS	FEMALE
2	584
5	157
1	147
3	65
7	41
9	39
8	38
4	24
8	10

**Tables 5: Grammatical Classification in order of frequency  
Most to least frequent class by gender for males**

CLASS	MALE
2	278
5	99
9	95
1	67
7	58
4	54
3	45
8	45
6	20

As one might expect given the fact that 1118 of the 1866 tokens are either discourse markers or particles, these categories are most prevalent for both the male and female interviewees. As far as the female interviewees are concerned, the discourse particle is the most frequently used with 584 tokens collected, while the discourse marker is the second most frequent with 157 tokens collected. The table above also shows that

the same is true for men with a total of 278 discourse particles collected, and 99 discourse markers collected.

An interesting point of comparison where the two genders differ regards the third most frequently elicited class for each. The third most frequently acquired type of token among my female interviewees was the quotative variant of *like* (class 1) with 147 tokens. This was however not the case among the males, where class 9 (which was an amalgamation of other uses of *like* such as prepositions and conjunctions) was the third most prevalent usage with 95 examples collected. Only 67 examples of quotative *like* were collected from the males, and this is in stark contrast to the 147 collected from the females. From this we can conclude that the quotative function of the variable is far more frequent amongst my female interviewees than amongst the males.

After these initial cross-tabulations were formulated, I quickly decided that I wanted to test the data again without the interviewer's tokens. The reason behind this decision was that I believe age to be a crucial factor concerning the variable, and therefore wanted to test my interviewees' tokens without my own examples. I am after all thirty-eight at the time of writing this thesis and fourteen years older than the eldest participants from my interviews. I decided the best way to do this would be to separate the particles and the markers, and compare these to all of the other classifications of *like*. Overall, this helped to demonstrate that 67.05% of all the tokens elicited from females were either discourse markers or particles; and 52.04% of all the tokens elicited by males (excluding the interviewer) were either discourse markers or particles. From this one starts to get a picture that the particle and the marker are indeed the most innovative linguistic tokens in terms of frequency.

I believe this to be perfectly in line with what leading sociolinguists have written and what I have attempted to describe in the historical and theoretical sections of this thesis. Specifically, that both the discourse particle and the marker have become common-place and an essential part of the grammar of young adolescents and teenagers in 2020. This is then followed by a one-level analysis of response class which show age and gender as significant predictors and can be seen below.

**Table 6: Cross-tabulation of linguistic token (without interviewer's tokens) contrasting classes 2+5 versus the others:**

CLASS	GENDER		
	FEMALE	MALE	TOTAL
1+3+4+6+7+8+9	364	141	505
2+5	741	153	894

TOTAL	1105	294	1399

**Table 7: Percentages of linguistic tokens 2+5 by gender**

GENDER	PERCENTAGE
FEMALE	67.05
MALE (EXCLUDING INTERVIEWER)	52.04

**ONE-LEVEL ANALYSIS OF RESPONSE CLASS WITH PREDICTOR(S):  
AGE (5.86e-07) + GENDER (1.97e-06)**

\$GENDER

FACTOR	LOGODDS	TOKENS	2+5/2+5+1+3+ 4+6+7+8+9	CENTRED FACTOR WEIGHT
FEMALE	0.32	1105	0.67	0.57
MALE	-0.32	294	0.52	0.42

\$AGE

continuous logodds  
+1 -0.103

\$misc.1

N	DF	INTERCEPT	OVERALL PROPORTION
1399	3	2.53	0.63

\$misc.2

LOG.LIKELIH OOD	AIC	AICc	Dxy	R2
-891.36	1788.73	1788.75	0.19	0.04

Current variables are:  
 response.binary: CLASS (2+5 vs. 1+3+4+6+7+8+9)  
 fixed.factor: GENDER  
 fixed.continuous: AGE

From the above, we can see that the female's have logodds of 0.32 compared to -0.32 for the males. The centred factor weight for the females is 0.57 which is just a little stronger than that of the males which is 0.42. The analysis again shows that my female interviewees are leading the way in terms of frequency of uses of *like* as both a discourse marker and particle (classes 2 and 5). It also shows that age is a predictor for both overall, while the males are not as strong in terms of frequency on either two.

Since I had already removed the linguistic tokens of the interviewer, it seemed a logical step to analyse the interviewer's tokens within each interview compared to my interviewees. Table 8 shows a cross-tabulation of the interviewer's tokens classified by interview and contrasted between whether they are markers/particles against all the other usages of *like*. Table 9 goes a step further by giving the exact percentages of both discourse markers and particles for each interview and both of these can be seen directly below.

**Table 8: Cross-tabulation for interviewer's tokens**

INTERVIEW	1+3+4+6+7+8+9	2+5	TOTAL
1	21	14	35
2	27	20	47
3	18	13	31
4	19	16	35
5	20	5	25
6	37	51	88
7	21	10	31

8	29	47	76
9	23	24	47
10	28	24	52
TOTAL	243	224	467

**Table 9: Cross-tabulation with percentages**

INTERVIEW	1+3+4+6+7+8+9	2+5	TOTAL	INTERVIEW	PERCENTAGE 2 AND 5 OF TOTAL
1	21	14	35	1	40
2	27	20	47	2	43
3	18	13	31	3	42
4	19	16	35	4	46
5	20	5	25	5	20
6	37	51	88	6	58
7	21	10	31	7	32
8	29	47	76	8	62
9	23	24	47	9	51
10	28	24	52	10	46
TOTAL	243	224	467		

From this analysis of the interviewer's tokens, we can see that from a total of 467 examples, a total of 224 have been classified as either discourse markers or particles, while the remaining 243 consist of all the other forms. From table 5 we can see that in interviews such as 6, 8, and 9, over 50% of the linguistic tokens are in fact discourse markers or particles. What we can conclude from this is that even amongst the eldest interlocutor from the interviews (I was thirty-eight at the time the interviews were conducted), the discourse marker and particle are showing an impressive degree of frequency. I feel that this is in line with what has been written in the theory section of this project; that speakers born in the 1980s (I was born in 1981) should also use these types of *like* within their everyday spoken English fairly naturally and systematically.

### 5.3 Secondary statistical Results Data



The next step of my analysis consisted of a more in-depth examination of the discourse marker and particle tokens collected from the interviews. In order to assess these two innovative forms more closely and test to what extent they show grammatical structure, I wanted to examine each token in terms of what came after the *like* when my interviewees said this. I therefore went through the 1118 tokens of markers and particles from the interviews, and classified them all in terms of what followed the word *like* within the sentence or utterance. The main types of grammatical structures that were noted were clauses (for example declarative, interrogative, imperative, subordinate, or not-clauses); phrases (for example verbal, adjectival, adverbial, prepositional, or participial), and the cross-tabulation of these can be seen in table 10 directly below.

**Table 10: Cross-tabulation of marker/particle tokens**

FOLLOWING FORM	DISCOURSE MARKER	DISCOURSE PARTICLE	TOTAL
ADJECTIVE PHRASE	4	30	34
ADVERBIAL PHRASE	2	14	16
APPOSITIVE PHRASE	0	1	1
BARE	10	75	85
DECLARATIVE CLAUSE	149	173	322
IMPERATIVE CLAUSE	1	2	3
INTERROGATIVE CLAUSE	35	34	69
NON-TENSED CLAUSE	0	33	33
NOT-CLAUSE	6	16	22
NOUN PHRASE	37	318	355
PARTICIPAL PHRASE	1	24	25
PREPOSITIONAL	2	27	29

PHRASE			
SUBORDINATE CLAUSE	7	24	31
VERB PHRASE	2	91	93
TOTAL	256	862	1118

From table 10, we can see that the most common grammatical forms that follow my interviewees' usages of *like* are declarative clauses with a total of 322 examples, and noun phrases with 355 examples. The marker also appears to favour interrogative clauses and disfavour a majority of the other forms as it is only present in such small numbers for all the other categories. The particle on the other hand also appears to favour interrogative clauses, verb phrases, as well as the bare form (where nothing is said after *like*); and is also prevalent to a higher degree in almost all other grammatical forms than the marker. Below is a list of examples of some of the most common structures from my interview corpus:

#### **.Discourse marker followed by declarative clause:**

1. **Like** when you practice it's easy enough, but when you speed up the tempo it's difficult. -*Michelle (token 218 from marker spreadsheet)*
2. I, **like** I'm so into working out! -*Ida (token 165 from marker spreadsheet)*
3. **Like** just being able to get anywhere like in a matter of like, hours is nice, like or like less than two hours basically you can go anywhere. -*Kira (token 73 from marker spreadsheet)*

#### **.Discourse particle followed by noun phrase:**

1. In my international school I had American people and British people so like, I learned a few **like**, different words. -*Julie (token 90 from particle spreadsheet)*
2. I think like, in a rich person's house, it's very **like** bohemian style, more rustic. -*Kira (token 265 from particle spreadsheet)*
3. I'm also able to work in **like**, private companies... -*Rita (token 347 from particle spreadsheet)*

The next cross-tabulation that was made involved grouping all the phrase types together, as well as all the clause types together, and the category titled 'bare' where nothing was said after *like*. The main reason for doing this was to show how the three

main classes were performing with regard to the variable. The results of this can be seen in table 11. Once this was done I also wanted to test these grammatical structures in terms of gender, so further cross-tabulations were made for both the male and female interviewees as separate groups, and the results of this can be seen in tables 12 and 13.

**Table 11: All speakers together (including the interviewer)**

		CLASS	
FOLLOWING GROUP	DISCOURSE MARKER	DISCOURSE PARTICLE	TOTAL
PHRASE	48	505	553
BARE	10	75	85
CLAUSE	198	282	480
TOTAL	256	862	1118

**Table 12: Female speakers**

		CLASS	
FOLLOWING GROUP	DISCOURSE MARKER	DISCOURSE PARTICLE	TOTAL
PHRASE	33	332	365
BARE	4	44	48
CLAUSE	120	208	328
TOTAL	157	584	741

**Table 13: Male speakers (including the interviewer)**

		CLASS	
FOLLOWING GROUP	DISCOURSE MARKER	DISCOURSE PARTICLE	TOTAL
PHRASE	15	173	188
BARE	6	31	37

CLAUSE	78	74	152
TOTAL	99	278	377

## RESULTS FROM THE CROSS-TABULATION

### . ANALYSIS OF THE DISCOURSE MARKER

#### 1. ONE-LEVEL ANALYSIS OF RESPONSE CLASS WITH PREDICTOR(S): FOLLOWING GROUP (4.2e-37) + AGE (0.0236)

\$follgrp

FACTOR	LOGODDS	TOKENS	MARKER/MARKER+PARTICLE	CENTRED FACTOR WEIGHT
CLAUSE	1.23	478	0.41	0.77
BARE	-0.43	85	0.11	0.39
PHRASE	-0.79	553	0.08	0.31

\$AGE

continuous logodds

+1 0.023

\$misc.1

N	DF	INTERCEPT	OVERALL PROPORTION
1116	4	-2.13	0.22

\$misc.2

LOG.LIKELIHOOD	AIC	AICc	Dxy	R2

-515.65	1039.31	1039.34	0.49	0.23
---------	---------	---------	------	------

Current variables are:

response.binary: CLASS (Marker vs. Particle)

fixed.factor: following group

fixed.continuous: AGE

From the one-level analysis response above, we can see that clause as a factor and increasing age favour the greater appearance of markers over particles. Therefore, in terms of the types of structures that encourage the use of particles over markers amongst my interviewees, we can conclude that the older the interviewees become, the more likely they are to use markers followed by clause types (predominantly declarative but also others) as opposed to particles.

## . ANALYSIS OF THE DISCOURSE PARTICLE

### 2. ONE-LEVEL ANALYSIS OF RESPONSE CLASS WITH PREDICTOR(S): FOLLOWING GROUP (4.2e-37) + AGE (0.0236)

\$follgrp

FACTOR	LOGODDS	TOKENS	PARTICLE/PARTICLE+MARKER	CENTRED FACTOR WEIGHT
PHRASE	0.79	553	0.91	0.68
BARE	0.43	85	0.88	0.6
CLAUSE	-1.23	478	0.58	0.22

\$AGE

continuous logodds

+1 -0.023

\$misc.1

N	DF	INTERCEPT	OVERALL
---	----	-----------	---------

			PROPORTION
1116	4	2.13	0.77

\$misc.2

LOG.LIKELIH OOD	AIC	AICc	Dxy	R2
-515.65	1039.21	1039.34	0.49	0.23

Current variables are:

response.binary: CLASS (Particle vs. Marker)

fixed.factor: follgrp

fixed.continuous: AGE

When the variable was tested with phrase type and bare as factors, again increasing age favours more frequent appearances of particles from my interviewees. In terms of defining a pattern of structure for the variable, we can see that the older the interviewee is, the more likely they are to use discourse particles in general, so age is not constraining or preventing the use of *like* as a discourse particle amongst my interviewees.

### 3. ONE-LEVEL ANALYSIS OF RESPONSE CLASS WITH PREDICTOR(S): FOLLOWING GROUP (2.47e-37) + GENDER (0.01)

\$GENDER

FACTOR	LOGODDS	TOKENS	PARTICLE/PARTICLE+MARKER	CENTRED FACTOR WEIGHT
FEMALE	0.2	740	0.78	0.55
MALE	-0.2	376	0.73	0.44

\$follgrp

FACTOR	LOGODDS	TOKENS	PARTICLE/PARTICLE+MARKER	CENTRED
--------	---------	--------	--------------------------	---------

			RTICLE+MAR KER	FACTOR WEIGHT
PHRASE	0.77	553	0.91	0.68
BARE	0.47	85	0.88	0.61
CLAUSE	-1.25	478	0.58	0.22

\$misc.1

N	DF	INTERCEPT	OVERALL PROPORTION	CENTRED INPUT PROB
1116	4	1.52	0.77	0.82

\$misc.2

LOG.LIKELI HOOD	AIC	AICc	Dxy	R2
-514.9	1037.8	1037.84	0.48	0.23

Current variables are:

response.binary: CLASS (Particle vs. Marker)

fixed.factor: GENDER follgrp

From this set of results, we can see that female gender, phrase, and bare are all promoting the occurrence of the particle. But when I tried to combine following structure, age and gender as predictors, the effect is too weak and comes out as non-significant, as can be seen in the results set below.

#### **4. ONE-LEVEL ANALYSIS OF RESPONSE CLASS WITH PREDICTOR(S): FOLLOWING GROUP (2.31e-37) + GENDER (0.146) + AGE (0.435)**

\$GENDER

FACTOR	LOGODDS	TOKENS	PARTICLE/PA	CENTRED
--------	---------	--------	-------------	---------

			RTICLE+MAR KER	FACTOR WEIGHT
FEMALE	0.15	740	0.78	0.53
MALE	-0.15	376	0.73	0.46

\$follgrp

FACTOR	LOGODDS	TOKENS	PARTICLE/PA RTICLE+MAR KER	CENTRED FACTOR WEIGHT
PHRASE	0.78	553	0.91	0.68
BARE	0.46	85	0.88	0.61
CLAUSE	-1.25	478	0.58	0.22

\$AGE

continuous logodds  
+1 -0.011

\$misc.1

N	DF	INTERCEPT	OVERALL PROPORTION
1116	5	1.79	0.77

\$misc.2

LOG.LIKELIH OOD	AIC	AICc	Dxy	R2
-514.59	1039.19	1039.24	0.5	0.23

Current variables are:  
response.binary: CLASS (Particle vs. Marker)



fixed.factor: GENDER follgrp  
 fixed.continuous: AGE

## 5.4 Final statistical results data

The final step of the analysis then consisted of a step-up/step-down analysis of prediction for both the particle and the marker. The main aim of this was to find groups which cause the model to change significantly when being added or subtracted. Therefore, the program tested each factor group and retained the most significant ones, by continually adding further groups until no further additions result in a significant change. This method is called step-up/step-down because it consists of two parts that proceed in reversed directions. The step-up analysis begins at 'level 0', where no factors are included, and progresses by adding new groups in the respective levels. The step-down analysis starts by calculating the likelihood of the model when all the factor groups are included in the regression simultaneously, and goes on by abstracting the least significant groups one after another. The results concerning finding the run which is the best fit of the model for the data should be identical in both steps (Tagliamonte, 2006). Now that I have explained how this type of analysis operates, let's examine the data acquired from my own step-up/step-down analysis.

### 1. Step-up/step-down analysis of prediction of marker (excluding interviewer):

**BEST STEP-DOWN MODEL OF RESPONSE CLASS IS WITH PREDICTOR(S): FOLLOWING GROUP (2.54e-22) + AGE (0.0139) + GENDER (0.0403)**

[p-values dropping from full model]

\$follgrp

FACTOR	LOGODDS	TOKENS	MARKER/MARKER+PARTICLE	CENTRED FACTOR WEIGHT
CLAUSE	1.04	393	0.38	0.73
BARE	-0.32	70	0.14	0.42
PHRASE	-0.72	429	0.09	0.32

\$GENDER

FACTOR	LOGODDS	TOKENS	MARKER/MARKER+PARTICLE	CENTRED FACTOR WEIGHT
MALE	0.22	152	0.28	0.55
FEMALE	-0.22	740	0.21	0.44

\$AGE

continuous logodds  
+1 0.078

\$misc.1

N	DF	INTERCEPT	OVERALL PROPORTION
892	5	-2.89	0.22

\$misc.2

LOG.LIKELIHOOD	AIC	AICc	Dxy	R2
-418.07	846.15	846.22	0.45	0.2

**BEST STEP-UP MODEL OF RESPONSE CLASS IS WITH PREDICTOR(S): FOLLOWING GROUP (3.42e-23) + AGE (0.0179) + GENDER (0.0403)**  
[p-values building from null model]

**BEST STEP-DOWN MODEL OF RESPONSE CLASS IS WITH PREDICTOR(S): FOLLOWING GROUP (2.54e-22) + AGE (0.0139) + GENDER (0.0403)**  
[p-values dropping from full model]

STEP-UP AND STEP-DOWN MATCH!

Current variables are:

response.binary: CLASS (Marker vs. Particle)  
 fixed.factor: follgrp GENDER  
 fixed.continuous: AGE

This first part of the step-up/step-down analysis of the discourse marker tested the three main factors of age, gender, and following group on my interviewees alone. The interviewer's tokens were therefore not included in this initial analysis and would be tested afterwards. The above set of results show that age proved a significant predictor for the linguistic variable. Essentially, as the age of the interviewee increases, so too do the logodds for discourse markers for both male and female. This suggests that the discourse marker is a form favoured strongly by the older interview participants as opposed to the younger ones. Gender and following group were also found to be significant predictors. The male interviewees for instance favoured the use of the discourse marker more so than the females, with logodds of 0.22 and a centred factor weight of 0.55. The most likely following group to come after the word *like* was found to be the clause, with logodds of 1.04 and a centred factor weight of 0.73. As discussed in the theory section of my work, Alexandra D'Arcy states that the discourse marker is generally associated with female speakers. My analysis therefore shows that in the case of my interviews, the males were in fact the gender group that favours the use of discourse markers.

Once this had been established, I wanted to run the same test on the discourse particle, again excluding the interviewer's linguistic tokens. The results of this can be seen directly below.

## 2. Step-up/step-down analysis of prediction of particle (excluding interviewer):

Current variables are:

response.binary: CLASS (Particle vs. Marker)  
 fixed.factor: follgrp GENDER  
 fixed.continuous: AGE

**BEST STEP-DOWN MODEL OF RESPONSE CLASS IS WITH  
 PREDICTOR(S): FOLLOWING GROUP (2.54e-22) + AGE (0.0139) +  
 GENDER (0.0403)**

[p-values dropping from full model]

\$follgrp

FACTOR	LOGODDS	TOKENS	PARTICLE/PARTICLE+MARKER	CENTRED FACTOR WEIGHT

PHRASE	0.72	429	0.9	0.67
BARE	0.32	70	0.85	0.57
CLAUSE	-1.04	393	0.61	0.26

\$GENDER

FACTOR	LOGODDS	TOKENS	PARTICLE/PARTICLE+MARKER	CENTRED FACTOR WEIGHT
FEMALE	0.22	740	0.78	0.55
MALE	-0.22	152	0.71	0.44

\$AGE

continuous logodds

+1 -0.078

\$misc.1

N	DF	INTERCEPT	OVERALL PROPORTION
892	5	2.98	0.77

\$misc.2

LOG.LIKELIHOOD	AIC	AICc	Dxy	R2
-418.07	846.15	846.22	0.45	0.2

**BEST STEP-UP MODEL OF RESPONSE CLASS IS WITH PREDICTOR(S):**

folllgrp (3.42e-23) + AGE (0.0179) + GENDER (0.0403)

[p-values building from null model]

**BEST STEP-DOWN MODEL OF RESPONSE CLASS IS WITH PREDICTOR(S): FOLLOWING GROUP (2.54e-22) + AGE (0.0139) + GENDER (0.0403)**

[p-values dropping from full model]

## STEP-UP AND STEP-DOWN MATCH

From the above data we can see that again there is another match, and that age is also an important predictor concerning the discourse particle. In terms of gender, the female interviewees lead the way in frequency with logodds of 0.22 and a centred factor weight of 0.55. The most favoured structure to follow *like* after a discourse particle was the phrase with logodds of 0.72 and a centred factor weight of 0.67. This set of results differs from that of the discourse marker in that it shows a divergent pattern in terms of both age and gender. Firstly, that for each year of age added to the interviewee, the logodds also fall and we can therefore say that increasing age definitely leads to fewer discourse particles from my interviewees. The discourse particle is therefore favoured by my younger interviewees, unlike the discourse marker; which was found to be more and more frequent for each year the interviewee became older. Secondly, in terms of gender, the discourse particle was favoured by the female interviewees which runs parallel to the discourse marker being favoured by the males.

Once the initial step-up/step-down analysis had been conducted with my interviewees alone, the next step consisted of conducting the same set of tests including the interviewer's linguistic tokens of *like*. The results of this can be seen directly below.

### 3. Step-up/step-down analysis of prediction of marker (including interviewer):

**BEST STEP-DOWN MODEL OF RESPONSE CLASS IS WITH PREDICTOR(S): FOLLOWING GROUP (2.47e-37) + GENDER (0.01)**  
[p-values dropping from full model]

\$follgrp

FACTOR	LOGODDS	TOKENS	MARKER/MARKER+PARTICLE	CENTRED FACTOR WEIGHT
CLAUSE	1.25	478	0.41	0.77
BARE	-0.47	85	0.11	0.38
PHRASE	-0.77	553	0.08	0.31

\$GENDER

FACTOR	LOGODDS	TOKENS	MARKER/MARKER+PARTICLE	CENTRED FACTOR WEIGHT
MALE	0.2	376	0.26	0.55
FEMALE	-0.2	740	0.21	0.44

\$misc.1

N	DF	INTERCEPT	OVERALL PROPORTION	CENTRED INPUT PROB
1116	4	-1.52	0.22	0.17

\$misc.2

LOG.LIKELIHOOD	AIC	AICc	Dxy	R2
-514.9	1037.8	1037.84	0.48	0.23

**BEST STEP-UP MODEL OF RESPONSE CLASS IS WITH PREDICTOR(S): FOLLOWING GROUP (1.11e-36) + GENDER (0.01)**  
 [p-values building from null model]

**BEST STEP-DOWN MODEL OF RESPONSE CLASS IS WITH PREDICTOR(S): FOLLOWING GROUP (2.47e-37) + GENDER (0.01)**  
 [p-values dropping from full model]

STEP-UP AND STEP-DOWN MATCH

Current variables are:  
 response.binary: CLASS (Marker vs. Particle)  
 fixed.factor: follgrp GENDER  
 fixed.continuous: AGE

When tested for gender, the logodds for the male participants were 0.2 and the centred factor weight was 0.55 which is significantly higher than for the female

participants, which was 0.44. When tested for following structure (i.e. what came after the word *like*) the clause proved again to be comfortably the most common structure to follow *like*, with logodds of 1.25 and a centred factor weight of 0.77. Both of these numbers are significantly higher than the other predictors tested. The phrase for instance only had logodds of -0.77 and a centred factor weight of 0.31, whereas 'bare' had logodds of -0.47 and a centred factor weight of 0.38. From all of this it becomes fairly clear that the discourse marker is showing a distinct pattern of being used directly before clauses by my interviewees; and predominantly by the males. Unlike the previous set of results concerning the discourse marker, age did not prove to be a significant predictor when tested here.

#### 4. Step-up/step-down analysis of prediction of particle (including interviewer):

Current variables are:

response.binary: CLASS (Particle vs. Marker)

fixed.factor: GENDER follgrp

fixed.continuous: AGE

#### **BEST STEP-DOWN MODEL OF RESPONSE CLASS IS WITH PREDICTOR(S): FOLLOWING GROUP (2.47e-37) + GENDER (0.01)**

[p-values dropping from full model]

\$GENDER

FACTOR	LOGODDS	TOKENS	PARTICLE/PARTICLE+MARKER	CENTRED FACTOR WEIGHT
FEMALE	0.2	740	0.78	0.55
MALE	-0.2	376	0.73	0.44

\$follgrp

FACTOR	LOGODDS	TOKENS	PARTICLE/PARTICLE+MARKER	CENTRED FACTOR WEIGHT
PHRASE	0.77	553	0.91	0.68
BARE	0.47	85	0.88	0.61
CLAUSE	-1.25	478	0.58	0.22

\$misc.1

N	DF	INTERCEPT	OVERALL PROPORTION	CENTRED INPUT PROB
1116	4	1.52	0.77	0.82

\$misc.2

LOG.LIKELIH OOD	AIC	AICc	Dxy	R2
-514.9	1037.8	1037.84	0.48	0.23

**BEST STEP-UP MODEL OF RESPONSE CLASS IS WITH PREDICTOR(S): FOLLOWING GROUP (1.11e-36) + GENDER (0.01)**

[p-values building from null model]

**BEST STEP-DOWN MODEL OF RESPONSE CLASS IS WITH PREDICTOR(S): FOLLOWING GROUP (2.47e-37) + GENDER (0.01)**

[p-values dropping from full model]

#### STEP-UP AND STEP-DOWN MATCH

From the above set of results, we can see that there is another match within the data. The discourse particle is being used most frequently this time by the females, and here the particle favours phrases in terms of following structure. When tested with gender as a predictor, the female interviewees had logodds of 0.2 and a centred factor weight of 0.55. Both of which were higher than the males. When tested for following structure, the phrase was the most dominant with logodds of 0.77 and a centred factor weight of 0.68. As with the previous set of results, the male speakers within my interviews favour the use of discourse markers followed by clauses; while the female speakers favour the use of the discourse particle followed by a phrase. However, again age did not prove to be a significant factor concerning the variable when the interviewer's linguistic tokens were included together with the interviewees.



## 5.5 Summary of the main findings of the analysis

I would now like to conclude my analysis by summarising the main findings concerning both the overall uses of *like* from all the interlocutors from my interviews, as well as the patterns involving discourse markers and particles more specifically.

In the initial phase of the analysis, the overall number of tokens were analysed in terms of frequency, age, gender, and grammatical structure, and a total of 1866 linguistic tokens were collected. 1105 of these were given by the female interviewees, and 761 from the males. The most frequently used forms being discourse particles and discourse markers. From the 862 examples of discourse particles, 584 were elicited from the females and 278 from the males. The second most frequent form of the linguistic variable used by all the interlocutors was the discourse marker. From the 256 examples of discourse markers, 157 were elicited from the female interviewees and 99 from the males. The third most frequent form used was in fact different depending on gender. 147 linguistic tokens of *like* were collected from the females in the quotative form, and 95 were categorised as either prepositional or conjunctive uses from the males. Most importantly, 67.05% of all the uses of *like* from my female interviewees were either discourse particles or markers and 52.04% from the males. In terms of frequency, the discourse marker and particle are by far the most commonly used forms of *like* by all the interlocutors throughout the interviews.

The secondary phase of my analysis helped to analyse the grammatical structure of the variable in greater detail. The main focal point for my cross-tabulations and analysis was to attempt to find patterns concerning what followed the use of *like* within the interlocutor's sentence or utterance. My analysis showed that clauses and phrases were the most prominent structures that followed *like*, with the declarative clause and the noun phrase being the most frequent of these. This second round of analysis helped to demonstrate that the older the interlocutor became, the more likely they were to use the marker *like* followed by a clause. Age was also found to be a factor that inhibited the use of the particles, and was clearly not preventing my older participants from using discourse particles as many tokens were collected from them. Other factors that encouraged the use of the particle were found to be gender and structure. Specifically, the female interviewees strongly favoured the use of the discourse particle *like* followed by a phrase, most frequently a noun phrase.

The third and final step of the analysis consisted of running the numbers one last time without the interviewer's tokens. This was done to test age as a predictor on all of my interviewees. Crucially, it was established that as the age of the interviewee increased, so too did the logodds for markers. And the opposite was in fact true for the particle. For every year added to the interviewee, the logodds decreased, which meant fewer discourse particles were recorded as my interviewees got older. Although the secondary round of analysis including the interviewer's linguistic tokens did not inhibit the use of the discourse particle; when the interviewee's were analysed alone, increasing age was found to constrain and disfavour the use of the discourse particle.

All of this helped to establish an important link in terms of following structure. The male interviewees favoured the use of discourse markers followed by clauses (most commonly declarative clauses), whereas the females favoured the use of discourse particles followed by phrases (most commonly noun phrases). And in terms of age, the older the interviewee became, the more likely they were to use discourse markers and the less likely they were to use discourse particles. I would therefore *like* to conclude my analysis by stating that my interviewees are indeed showing some very strong and well-defined patterns in their uses of *like*; in terms of age, gender and grammatical structure.

## 6. DISCUSSION:

In the following section of my thesis, I would like to discuss some of the delimitations of my work, as well as how I feel this study could be expanded and elaborated upon. As far as the general process of conducting a sociolinguistic project of this size is concerned, I am very pleased on the whole. I managed to interview ten very different and interesting people and acquire around eight and a half hours of relaxed conversation that in turn led to the acquisition of 1866 different forms of *like* that formed the basis of my analysis. However, I do feel that an inevitable constraint surrounding all of this involves time. I think it's fair to say that given more time to conduct interviews and collect data, a larger corpus could have been acquired which could have led to even more interesting analysis and conclusions.

If one for example looks at the work of sociolinguists such as Alexandra D'Arcy and Sali A. Tagliamonte, much of their work is focused on how the linguistic variable performs at different points in time, or trend studies as they are known. A trend study builds up a corpus of speech from speakers at different points in time who can be considered roughly comparable to one another. These studies are called trend studies because the real time lag between the first set of data and later sets allows you to observe how diachronic trends progress through the community (Meyerhoff, 2011:139). Of course this type of study where extensive corpora of spoken English are collected over decades and then analysed would not be realistic for a project that only lasts for five months from start to finish; although I do feel that my thesis could be expanded upon to become a trend study.

I feel that overall, my work has shown that the way my interviewees use the word *like* has a definite and fairly cohesive structure to it. And I do also feel that if this study were to be conducted on a larger scale; but still specific to Denmark and the way its population uses the English language, then this could tell us even more about the variable *like*. If, for example, more interviews were conducted in different parts of the country; they could then be compared in geographical terms as well as over time. I also feel that generally-speaking with this kind of work, the more data available to analyse, the better. If larger corpora could be built up over time and focusing on the

same variable, I think it would be interesting to examine how different language communities within different cities in Denmark use the variable. But again, this is not something that would be realistic over the course of a five month thesis.

Another possible area that I both feel could have been expanded on in my own work, but also on a larger scale, is the study of quotative *be like*. Although much was written in the theory section about the structures and forms that constrain and inhibit the quotative form of the variable, much of this was unfortunately not tested in the analysis due to time constraints. I definitely believe that if one wanted to conduct this kind of work on a larger scale, then a more detailed analysis of the quotative would be another excellent way to do this. The main factors inhibiting the use of the quotative were tense and person, as well as the content of the quote. I think if one were to extract all of the quotative tokens from either my own corpus, or a larger set of corpora; then one could run the same set of cross-tabulations and tests to see to what extent tense and person inhibit the use of *like* as a quotative. One would no doubt also want to examine the other factors such as age, frequency, and gender as well.

## 7. CONCLUSION

I would like to conclude my thesis by discussing to what extent I have answered my problem formulation. The initial problem formulation was: **“To what extent does the linguistic variable *like* as a marker and a particle show structure and frequency within the spoken English of my interviewees? And to what extent can the variable be described as unpredictable and haphazard?”**

Throughout this thesis, I have attempted to demonstrate that the linguistic variable *like* in fact shows a substantial degree of both structure and frequency in all grammatical forms throughout the interviews I conducted. In particular with regard to the discourse particle and marker. Of the 1866 tokens of *like* collected from the interviews, 1118 were either particles or markers. 862 were in fact particles, and 256 were markers. From these numbers we can see that the particle and the marker are clearly the dominant types of *like* being used by my interviewees, and have clearly become an important part of their day-to-day vocabulary of spoken English.

Throughout the analysis section of my thesis, I have also attempted to demonstrate how age and gender are intrinsically linked to frequency, and that some discernable patterns have emerged from my data. Firstly, the fact that over 67.05% of the tokens collected from my female interviewees were either particles or markers shows just how dominant these two forms are within their spoken English overall. And the fact that over 57.04% of the tokens elicited from the male interviewees were either particles or markers shows that although the female interviewees are leading the way in terms of frequency, this is not something that can be solely attributed to women. Specifically within my corpus of data, both males and females favour the use of

discourse particles and markers strongly over all other forms.

Perhaps the most salient pattern that I would like to highlight from my analysis concerns both the age of the interviewee, and the structure of what follows the *like* in each linguistic token. Crucially, the male participants from my interviews strongly favoured the use of discourse markers followed by a clause (most frequently a declarative clause); and for every year that they became older, so too did the chances increase for them to use the marker. The marker therefore is a form that is favoured by the older interviewees, and predominantly the males. This pattern is linked to the particle as well. Indeed, a similar structure was demonstrated in my analysis whereby my female interviewees strongly favoured the use of the particle followed by a phrase (most frequently a noun phrase); and for each year their age increased, the less likely they were to use the discourse particle. From this one can trace a very distinctive pattern in my interviews where as the participants got older, they tended to use less particles and more markers. The younger speakers therefore strongly favoured the use of the particle. Two very distinct and different historical trajectories have been established in my analysis.

Although leading sociolinguists such as Sali Tagliamonte and Alexandra D'Arcy have stated that in the data they collected and in their work generally, the discourse marker tended to be favoured by females in terms of gender and the particle tended to be favoured by males; this was rather interestingly not the case regarding my interviewees. As stated in the previous paragraph the marker was strongly favoured by the older males within my interviews, and the particle by the younger females. Although this particular structure does not match exactly with what has been written in that respect, I feel that this is one of the most interesting and concrete patterns from my work. I feel that this shows very distinctly how age, gender, and structure all play an important role in encouraging and inhibiting the use of these two forms. To describe the variable as chaotic and haphazard is therefore too simplistic.

Throughout the historical and theoretical section of the analysis, I have also attempted to disprove this statement. I feel that all of the historical and theoretical evidence that has been written all points to generational transmission, as opposed to some kind of mystical appearance of the variable overnight amongst young and adolescent speakers. Generational transmission itself suggests a certain structure. That the use of *like* has been well-established for over a century-and-a-half in most of its guises (although admittedly the quotative possibly came a little later), and has steadily been used and re-used by generation after generation of new speakers of English to a point where it is now an undeniable and important aspect of spoken English in the vernacular. The tokens collected from my corpus are unequivocal evidence of this.

Although I do to a certain extent understand that it may be tempting to dismiss the variable as chaotic and unstructured, and even as a sub-standard way to speak English, I feel this is categorically not the case. I think a lot of the negative attitudes surrounding the word *like* are linked to the versatility and ambiguity of the word. Indeed, *like* is an incredibly diverse word that can serve a myriad of functions, many

of which have already been written about in this thesis at great length. And the fact that it is used today with such incredible frequency can lead to a great deal of confusion. However, I would have to agree with Alexandra D'Arcy when she states that “a linguistically-informed perspective utterly undermines any claim that *like*, in any of its uses, is random and meaningless” (D'Arcy, 2017:31), and I hope my thesis goes some way to validating this statement.

## 8. BIBLIOGRAPHY:

. Bitsch Olsen, P. and Pedersen K. ed., (2008). *Problem-Oriented Project Work*. 1st ed. Roskilde: Roskilde University Press.

. D'Arcy A., (2017). *Discourse-Pragmatic Variation in Context: Eight hundred Years of LIKE*. John Benjamins Publishing Company.

Downloaded from: <https://ebookcentral.proquest.com/lib/kbdk/reader.action?docID=5017626>

. Davydova J., (2020). *The role of sociocognitive salience in the acquisition of structured variation and linguistic diffusion: Evidence from quotative be like*. Cambridge University Press.

Downloaded from: <https://www.cambridge.org/core>.

. Meyerhoff M. (2011). *Introducing Sociolinguistics*. Published by Routledge.

. Meyerhoff M., and Schlee E. (2010). *The Routledge Sociolinguistics Reader*. Published by Routledge.

. Tagliamonte S., D'Arcy A., Rodriguez Louro C., (2016). Outliers, impact, and rationalization in linguistic change. Published by Linguistic Society of America.

Downloaded from: <https://doi.org/10.1353/lan.2016.0074>

. Tagliamonte S., (2016). *Teen Talk: The Language of Adolescents*. Cambridge University Press.

### Websites:

. Daniel Ezra Johnson, *Rbrul manual*.

[http://www.danielezrajohnson.com/Rbrul\\_manual.html](http://www.danielezrajohnson.com/Rbrul_manual.html)

.Tagliamonte S., (2006). *Step-up/step-down analysis* Conducting an analysis with the variable rule program.

Taken from: [https://www.uni-bamberg.de/fileadmin/eng-ling/fs/Chapter\\_20/Index.html?Stepupstepdownanalysis.html](https://www.uni-bamberg.de/fileadmin/eng-ling/fs/Chapter_20/Index.html?Stepupstepdownanalysis.html)