The discourse marker LIKE: a corpus-based analysis of selected varieties of English

Dissertation zur Erlangung des Grades des Doktors der Philosophie bei der Fakultät für Geisteswissenschaften

Fachbereich Sprache, Literatur, Medien & Europäische Sprachen und Literaturen der Universität Hamburg

Überarbeitete Version

vorgelegt von

Martin Schweinberger

aus Bad Karlshafen

Hamburg, Juni 2014

Hauptgutachter:	Prof. Dr. Peter Siemund
Zweitgutachter:	Prof. Dr. Markku Filppula
Datum der Disputation:	30.09.2011
Abgenommen von der Fakultät für Geisteswissenschaften der Universität Hamburg am:	02.11.2011
Veröffentlichung mit Genehmigung der Fakultät für Geisteswissenschaften der Universität Hamburg am:	03.09.2014

Acknowledgments

This book is the result of research conducted during my work at the University of Hamburg. My warmest thanks for his support go to Prof. Peter Siemund, my supervisor – without him this book would have been impossible and to whom I am deeply indepted for his support. The many people within the University who have made this an extraordinary fruitful work environment through their collegial support, fruitful exchange of ideas, and their friendship are too numerous to mention, but among them, Florian Dolberg, Georg Maier, Thomas Berg, Suzanne Flach, Tayo Takada, Patrick McCrae and Svenja Kranich stand out.

Much of the book is based on data taken from the *International Corpus of English* (ICE) and the sociolinguistic approach taken here would not have been possible without having been granted access to the wealth of extra-linguistic information not yet available to the wider scholarly community. My gratitude is thus due to all those ICE teams which supported my research. I am also indebted to the Hamburgische Wissenschaftliche Stiftung for kindly granting additional financial support, which enabled me to take part in conferences and work-shops which I would otherwise not have been able to attend.

In addition, I am grateful to *The Skeptics Guide to the Universe*, *Richard Dawkins*, *Thunderf00t*, *Pharyngula*, *Point of Inquiry*, *The Atheist Experience*, and *Aronra* for promoting science and skeptical thinking. They have entertained me during endless hours of coding data, running statistical models and editing this book, as well as educating me on the scientific method and illustrating how easily one can be deceived, not only by others but also by oneself.

Last but not least, my family, in particular my parents Erika and Klaus, my brother Enno, my grandmothers Charlotte and Anna, and my daughter Zoe Milena deserve special thanks for their support and love, which has carried me along what has at times been a difficult passage.

Table of Contents

List of Figures

List of Tables

List of abbreviations

ICE	International Corpus of English
CIE	Corpus of Irish English
COLT	The Bergen Corpus of London Teenage Language
LCIE	Limerick Corpus of Irish English
AmE	American English
AusE	Australian English
CanE	Canadian English
EngE	British English
EAE	East African English
IndE	Indian English
IrE	Irish English
JamE	Jamaican English
NIE	Northern Irish English
NZE	New Zealand English
PhiE	Philippine/Filipino English
SctE	Scottish English

1 Introduction

The discourse marker LIKE[1] as in (1) is one of the most salient features of present-day English (cf. D'Arcy 2005:ii). Despite being deemed archaic (Underhill 1988:234), dismissed as meaningless and considered symptomatic of careless speech (Newman 1974:15), this non-standard feature has received scholarly attention and attracted interest in the public media (Diamond 2000; Johnson 1998; Levey 1999; Peters 2008). In spite of being met with derision, its functional versatility and global presence make LIKE an ideal object for cross-varietal, sociolinguistic analyses of ongoing change and socially motivated variation. In fact, vernacular uses of LIKE are "rapidly increasing in the speech of the younger generation, particularly in Western English speaking countries" (Tagliamonte 2005:1898).

(1) a. And took the stairs and a lot of people were going out but the thing is *like* maybe half of the people were still stayed in there. (ICE Philippines:S1A-007$B)

b. Because they had this sort of <,> *like* uhm <,> you know <,> that kind of flooring tile up the stairs and… (ICE Canada:S1A-007$B)

c. But a lot like Mike's *like* quite a trustworthy guy and might just *like* let us drink it all afterwards. (ICE GB:S1A-030$A)

d. I I think it's 'tis geared for death you know *like* <S1A-055$A> Mm (ICE Ireland:S1A-055$B)

The discourse marker LIKE is well worth scholarly attention, because of its high frequency and salience in present-day English(es), and because it may help to understand mechanisms and processes of local implementation of globally available innovative forms[2]. In addition, the global spread of LIKE may

1 In the following, LIKE in capital letters refers to the discourse marker LIKE, while *like*, written in lower case italics refers to instances of *like* which do not fall into the category 'discourse marker'. This does not apply, however, to instances of the discourse marker LIKE which occur in examples.

2 This terminology is adopted from Buchstaller (2008), and Buchstaller and D'Arcy (2009) who have approached the analysis of *be like* as an opportunity to study how globally available features are adapted during implementation in local systems. As the

serve in evaluating the stability of sociolinguistic mechanisms in distinct contact scenarios. Hence, the present investigation addresses questions such as: How do linguistic features spread throughout speech communities? Who promotes linguistic change? Which groups adopt new features more readily? Are some varieties of English integrating incoming features more quickly than others, and if so, why? Do certain communities reject LIKE due to its ideological association with the United States?

Vernacular uses of LIKE are ideal for answering these kinds of questions. The discourse marker LIKE is highly frequent and almost universal, particularly among younger speakers; it is also syntactically optional and thus flexible, allowing rapid change. These features are valuable for the purpose of observing and describing the globalization of vernacular features, as they also allow an evaluation of the strength and consistency of recurring patterns in ongoing language change in diverse and multilingual settings. These attractive characteristics coincide with features of prototypical discourse markers and, hence, determine LIKE's status as a part of speech. In addition, these characteristics render LIKE a perfect testing ground for studying the globalization of vernacular features.

Despite the ever-growing amount of literature, the discourse marker LIKE has so far not been investigated from a cross-varietal, variationist perspective which systematically surveys usage patterns across varieties of English. This is remarkable, considering that a cross-varietal analysis of multifunctional vernacular forms "provides a unique opportunity to assess the complex

present investigation is similar to these studies with respect to its theoretical outlook and methodological approach, the wording is adopted to avoid terminological confusion.

interaction of social and discourse-pragmatic correlates across the world's many varieties of English"[3] (Tagliamonte & Denis 2010:28).

The current study aims to resolve this shortcoming by focusing on the sociolinguistic patterning in distinct regional varieties of English in order to unearth the sociolinguistic mechanisms of local adoption and embedding of global innovations – exemplified by vernacular uses of LIKE. To address these issues and related questions, this analysis employs a multi-method approach (i.e. combining quantitative and qualitative methodology) based on large-scale, comparable data sets (the ICE family of corpora) and advanced statistical evaluation.

Given that LIKE has become a salient feature in contemporary spoken English around the world, it is rather surprising that LIKE has so far not been investigated from a cross-varietal perspective which systematically surveys usage patterns across varieties of English. Consequently, the aim of the present study is to survey the variety-specific usage patterns of the discourse marker LIKE and to retrace its spread across the world. Furthermore, this study will re-assess the pragmatic functions associated with LIKE and it will test claims regarding the use of LIKE by speakers differing in age and gender.

Although the study of the linguistic behavior is based on a "sizeable amount [sic.] of cases, [they are] scattered across the few communities that have been selected for a sociolinguistic study" (Labov 2001:284). Although more recent studies have taken a less anglo-centric perspective, ongoing change has so far only rarely been analyzed from a global perspective. In fact, most sociolinguistic studies have focused on monolingual settings (cf., for example, Labov 2001:518), while the more common case of multilingual settings has for the most part been neglected (Sture Ureland 1989:242-245). The overall perspective is predominantly monolingual, and it stresses inner-linguistic

[3] This assertion by Tagliamonte and Denis (2010) originally refers to general extenders and not to discourse markers, but it is also viable with respect to other discourse-pragmatic features.

processes, assuming that contact is secondary and of minor importance. The present study addresses this shortcoming by providing a detailed analysis of ongoing language change within eight geographically distinct varieties of English.

With respect to the structure of the current investigation, the present chapter introduces the overall context and presents the issues addressed in this analysis of vernacular uses of LIKE, while chapter 2 is dedicated to describing the theoretical underpinnings of this research. Accordingly, chapter 2 introduces the basic variationist concepts, provides a brief overview of contemporary sociolinguistic theory, and explains essential mechanisms of language change and variation. Thus, it provides the theoretical framework employed to summarize, discuss, and interpret the findings in chapters 6, and 7.

Chapter 3 focuses on discourse markers in general and LIKE in particular. Hence, this chapter deals with definitions and typical features of discourse markers to ascertain whether it is justifiable to classify specific instances of LIKE as a discourse marker. Although this may seem trivial, it is not: for example, if the instances in (2) show, LIKE does not always behave like a prototypical discourse marker: in contrast to other typical discourse markers, such as *you know*, *I mean*, *well*, *so* etc., it appears to be more deeply integrated into syntactic structure and may at least in certain contexts be regarded as a borderline case between discourse marker and quotative complementizer as in (2b) or, according to Andersen (1997:379), between discourse marker and adverbial (as in (2c)).

(2) a. Uh there is there was *like* another company that did ribbons stripped ribbons. (ICE Jamaica:S1B-072$C)

b. [A]nd I was *like* forget it. (ICE Canada:S1A-022#B)

c. I ran away for *like* five days for almost a week I stayed at my friend's house... (ICE Philippines:S1A-047$B)

Chapter 4 presents theories which deal with the historical development of LIKE. In particular, this chapter discusses models of grammaticalization and

pragmaticalization put forth in the respective literature. It surveys the relevant literature on LIKE and depicts functionally distinct uses of LIKE as discussed particularly in discourse-pragmatic analyses. Although it may appear that all uses of LIKE are realizations of a single, underlying form, a fine-grained analyses provide a more detailed picture (e.g. Andersen 2000; D'Arcy 2007). Indeed, LIKE appears to be of a multifaceted nature (cf. D'Arcy 2007:391-397) comprising a heterogeneity of functionally distinct uses which occur under specific conditions and in rather well circumscribed contexts (e.g. D'Arcy 2005:ii; Tagliamonte 2005:1897). Depending on the linguistic context, LIKE fulfills a variety of (pragmatic) functions. Section 4.7 exemplifies the functional and positional diversity of LIKE and provides a classification which allows the systematization of seemingly unrelated instances of LIKE.

Chapter 5 concerns itself with issues related to the data and methodology of the present analysis. In particular, the database deserves additional attention: it illustrates how the ICE can be utilized to serve as a valuable resource for sociolinguistic analyses of ongoing change.

In order for the ICE to be helpful for variationist studies, it had to be computationally processed. The resulting edited version of the ICE matches the requirements of fine-grained sociolinguistic research, as the exact word counts for each individual speaker are extracted. In contrast to previous studies based on the respective ICE components, these word counts allow the calculation of the raw frequency of innovations in the speech of each speaker. In turn, these raw frequencies can be transformed into normalized frequencies (e.g. per-1,000-word frequencies) which guarantee maximal comparability. While previous studies have mainly been concerned with regional variation, elaborate computational processing expands the versatility of the ICE data and enables research even on the level of individual speakers.

The edited version of the ICE consists of text files comprising only utterances of one particular speaker. As the teams compiling the ICE components provided extensive information of the individual speakers, each

individualized sub-component has been assigned to various sociolinguistic attributes of the speaker. When retrieving the instances of LIKE, it was thus possible to retrace the age, gender, occupation, and L1 of the speaker producing this instance. In addition, it was possible to retrieve whether the informant speaks other languages, e.g. German or Spanish, where he or she has been brought up, and exactly where the speaker now lives. Accordingly, it has been possible to assign each token of LIKE to a multitude of sociolinguistically relevant variables. For example, a certain occurrence of LIKE can now be attributed to a speaker aged 26 to 33 who is pursuing an academic career, lives not in Northern Ireland but in the Republic of Ireland, and speaks English as his first language. Although this approach is extremely intriguing, particularly with respect to the analysis of the sociolinguistic distribution of certain forms in regionally and culturally diverse settings, it has a notable deficiency: the speaker information did not always cover the entire spectrum of features, or was even entirely missing for certain speakers. In such cases, the instances of such speakers had to be removed from the analysis, in some cases leaving only a relatively small number of speakers in the data set.

This represents a valuable innovation, since the resulting data may offer intriguing opportunities for studying ongoing change on the micro-level, i.e. on the level of individual speakers. Hence, the increase in versatility offers a highly accurate depiction of regionally distinct usage patterns. In this sense, the regionally distinct ICE components represent ideal resources for cross-varietal analyses and satisfy the need for both matching data sets and comparable methods as expressed by Buchstaller and D'Arcy: "What is needed, therefore, are reliable and comparable methods applied rigorously and uniformly across datasets to uncover which constraints hold both across and within varieties of English worldwide" (2009:298).

Nonetheless, the analysis of LIKE in geographically distinct locales may serve as a case in point for how ICE components may serve as resources for future research in sociolinguistics. Indeed, they represent ideal databases as

they offer a wide variety of extra-linguistic variables and represent various registers paired with a matching design.

Chapter 6 surveys the relative frequencies of LIKE in the combined data, focusing on and comparing the use of LIKE across varieties of English. This cross-varietal survey displays differences of both the overall frequency and the functionally distinct uses of LIKE. In addition, this chapter provides a preliminary analysis of the global gender and age distribution of LIKE to gain insight into general tendencies of its use.

Chapter 7 represents the core of the present analysis, as it evaluates correlations between functionally distinct uses of LIKE and extra-linguistic social variables such as age and gender. The relevant aspect for the present study is, however, not the mere existence of such correlations, but the fact that they are indicative of specific stages of ongoing language change. As specific phases of change are prototypically associated with distinct degrees of gender differentiation (Labov 2001:307-308) and age stratification (Labov 2001:449), the distribution of innovations across age groups and their degree of gender differentiation, inform about whether a certain variant is currently undergoing change, and if so, at which stage of change it can be located. Hence, based on correlations between the use of LIKE and extra-linguistic variables, it is possible to provide a fine-grained account of the sociolinguistic distribution of LIKE within regionally distinct varieties of English. Indeed, contemporary sociolinguistic theory utilizes the sociolinguistic profiles of innovative variants to draw inferences about the trajectory of the ongoing change. The crucial point here is that the trajectory of change provides information not only about the type of change, but also about the subsection of speakers responsible for "the fascinating spread of types of uses of *like* (Labov, personal communication)" (Fox & Robles 2010:716).

This assumption that the distribution of innovative forms allows far-reaching inferences to be drawn about trajectories of change is, however, not without problems. This criticism is particularly relevant with respect to

analyses of language change which are based solely on the apparent-time construct which has been essential to sociolinguistic analyses for over 50 years (Bailey et al. 1991:241). The basic assumption underlying the apparent-time construct is that the linguistic behavior of older speakers reflects earlier historical stages of the linguistic system. This premise is not trivial. Although a monotonic pattern (a near-linear recess in use with increasing age) commonly reflects ongoing change, emerging monotonic patterns in apparent-time analyses require additional inspection from a real-time perspective. Without real-time confirmation, issues relating to the exact type of change – age-grading, generational change, communal change (Labov 2001:76) – remain unresolved. In other words, whether a given distribution represents age-grading or change in progress "can only be determined by comparing the usage of speech communities at two points in time. Only then can we tell if contemporary variation, or what we might call 'change in apparent time' is a stage in long term change, or 'change in real time'" (Romaine 2005:1702).

This problem arises because the apparent-time construct relies on the assumption that once the use of a certain form is fully acquired, it remains "essentially fixed or static over the course of the lifetimes of […] individuals" (Tagliamonte & D'Arcy 2009:61). However, Tagliamonte and D'Arcy point out that "[t]here is an increasing body of research […] documenting ongoing change throughout the lifespan" (2009:61). To address this difficulty, the present study complements the apparent-time results with an additional real-time analysis, when the data permits.

Chapter 8 is dedicated to the interpretation of the findings with respect to re-tracing both the spread of LIKE across the world, and its diffusion through local speech communities and global spread. In addition to discussing the findings in light of previous research on LIKE, chapter 0 sheds light on the interaction between (universal) mechanisms of language change and the cultural diversity of local practices.

Finally, chapter 9 presents the conclusions drawn from the present analysis and provides an outlook for further research.

In summary, the present analysis focuses on systematic correlations between the use of the discourse marker LIKE and extra-linguistic variables in geographically distinct locales. Thus, the investigation sets out to analyze the sociolinguistic mechanisms underlying use of LIKE, and employs both apparent-time and real-time data to retrace LIKE's historical development. To account for the differences in the usage patterns of LIKE across varieties of English, the study employs a multi-method approach, i.e. combining quantitative and qualitative methodology. The cross-varietal section employs multifactorial statistical evaluation of the frequencies of LIKE occurring in various grammatical environments. This quantitative, empirical analysis is based on large matching data collections – the ICE family of corpora.

2 Language change and variation

The following chapter provides and discusses the theoretical framework of this study and familiarizes the reader with crucial concepts of modern sociolinguistic theory of language variation and change. In addition to presenting these concepts, the focus of this chapter lies on discussing findings which form the empirical basis of these concepts. The chapter will proceed from more general issues, such as what language change refers to and how it is affected by globalization, to concepts specific to a Labovian understanding of mechanisms underlying language variation and change such as age-grading and communal change. Furthermore, the chapter will discuss prominent approaches used to detect ongoing change, such as apparent-time and real-time analyses, and extra-linguistic factors that affect ongoing change such as gender, age, and prestige.

2.1 Introduction

The study of LIKE is neither an end in itself, nor is it primarily intended to provide a synchronic profile of LIKE use across and within varieties of English. Rather, it serves to exemplify and test more general models and mechanisms of language change. So far, studies have focused almost exclusively on change within single communities in rather limited geographical areas. In order to address this shortcoming, the current investigation attempts to recreate how LIKE entered the various varieties of English. Thus, it takes a global perspective on the sociolinguistic patterns of language change, with the aim of evaluating the features which reflect universal patterns of change and contrasting these with features which are variety-specific and display restrictions on universal mechanisms by a variety of specific conditions.

2.2 Language change

Language is by nature dynamic, and thus change is ever present (Hickey 2001:1). It follows that language is not a homogenous entity, but constitutes heterogeneous forms. Nevertheless, this heterogeneity is not random; it is an "orderly heterogeneity" (Weinreich, Labov & Herzog 1968:100), allowing the scientific endeavor to unearth general patterns within seemingly irrational or unpredictable processes underlying linguistic behavior (cf. Labov 1994:10). The search for stable patterns of linguistic variation constitutes the aim of variationist approaches to language. The object of models for language change is thus uncovering the mechanisms through which variation and its general trends arise.

In view of this endeavor, the study of the discourse marker LIKE does not intend to provide merely a synchronic profile of LIKE use across and within varieties of English, but it exemplifies and tests more general models and mechanisms of language change. Hence, this investigation provides a global perspective on language change and variation, and addresses questions such as how linguistic innovations spread around the world by identifying which social groups and factors advocate ongoing change. So far, studies have focused almost exclusively on change within either single communities, or a moderate set of varieties within rather limited geographical boundaries. In order to address this shortcoming, the present investigation takes a global perspective on the sociolinguistic patterns of language change with the aim of evaluating which features reflect universal patterns of change and which are constrained by variety-specific conditions.

The aim of discovering general laws of linguistic change and separating these from features of change best attributed to chance is by no means a recent development. Throughout the past century, starting with Gauchat's (1905) study of ongoing phonological change in the Swiss French village of Charmey, a growing body of research has provided a detailed account of trends and principles underlying seemingly random changes of linguistic behavior. Such

general principles would allow the prediction that change, once initiated, "will move through the speech community in a uniform fashion" (Labov 2010:184).

The mechanisms by which new forms are introduced and promoted within speech communities are inextricably linked with the identification of the social location of innovators. Knowledge of the social location of these speakers allows one to investigate the role of factors such as socioeconomic status, gender, age and ethnicity on the transmission, incrementation, and continuation of change (cf. Labov 2001:xiv). Hence, the analysis of variables which allow linguistic behavior to be reconstructed and subsequently predicted has become one of the key aims of recent sociolinguistic approaches towards understanding the underlying general mechanisms of language change throughout the last century.

The fact that language, like everything else in nature, is constantly undergoing change, transformation and renewal is, however, often considered a destructive force resulting in decay or corrosion and leading to an inferior, less perfect state of language by both scholars and laymen (Aitchison 2001:4-14).

> A language may become greatly altered and excessive prevalence of the wearing out processes, abandoning much which in other languages is retained and valued. It is necessary that we take notice of the disorganizing and destructive workings of this tendency. (Whitney 1904:75)

The 'Golden Age' principle, i.e. the assumption that language is continuously decaying, has led to the view that language change is "an unmixed evil" (Labov 2001:30), and encouraged linguists to blame proponents of change for their corrupting influence and demanded that social action be taken.

> Such phonetic changes […] are inevitable and creep in on themselves; but this is only another way of saying that we do not know who in particular is to blame for them. Offenses must come, but there is always that man by whom they come, could we but find him out. (Whitney 1904:34)

The negative evaluation of language change appears to be part of the human condition, as lamentation about the decay of language and morals of every younger generation is probably as old as language itself. It has sparked many attempts to uncover its mechanisms and provide explanatory models. Only in the latter half of the twentieth century have scholars emphasized positive explanations of language change – Chambers (1995), for example, has stressed the positive aspects of variation (cf. Labov 2001:191).

2.3 Globalization, local practice and the diffusion of LIKE

The most fundamental issue addressed in the present research relates to adoption and embedding of globally available innovations into localized speech communities. In other words, this investigation explores the implications of globalization for contemporary concepts of sociolinguistic theory: "globalization forces sociolinguists […] to rethink itself as a sociolinguistics of mobile resources, framed in terms of trans-contextual networks, flows and movements" (Blommaert 2010:1). Globalization in the present context is, hence, defined as:

> a multidimensional set of processes that create, multiply, stretch, and intensify worldwide social interdependencies and exchanges while at the same time fostering in people a growing awareness of deepening connections between local and the distant (Steger 2003:13).

Indeed, over the past few years, speech communities in culturally diverse settings have increasingly gained attention among scholars in general, and sociolinguists in particular (cf. Buchstaller & D'Arcy 2009:293). However, this globalized setting and the impact of supra-local flows on local systems require adjusting methodological and theoretical concepts.

Meyerhoff and Niedzierlski (2003) were among the first to address the relation between effects of globalization (or more specifically, Americanization), i.e. supra-local trends, and their implementation in local

systems. What they observed was, however, not a uniform trend towards standardization, as e.g. Hjarvard (2004) suggests, but that "globalization has been found to be accompanied by increased localization" (Meyerhoff & Niedzierlski 2003:535). This finding contrasts with previous studies which suggested that "[l]anguage itself is seen as essentially unaffected by globalization (culture, society, and so on), and globalization is seen as just another context in which language is practiced, a new one at best" (Blommaert 2010:2).

Commonly, sociolinguistic analyses have viewed the diffusion[4] of linguistic variants as spreading outward from a major economic or cultural epicenter. While this poses less of a problem when regionally adjacent speech communities are concerned, diffusion across non-continuous geographic settings is more complex (Buchstaller & D'Arcy 2009:291). Meyerhoff and Niedzierlski (2003) point to a very intriguing aspect of trans-national spread: following Audretsch (2000:73) they hypothesize that complex meanings of variables spread "only if there is face-to-face, quality contact between individuals" (2003:537), while only fairly superficial aspects of the innovation are transferred if personal contact is notably limited (cf. Meyerhoff & Niedzierlski 2003:538-539). Most sociolinguistic analyses – in particular those discussed in Labov (2001:228) – assume face-to-face contact as the prototypical scenario in cases of diffusion and show that the influence of mass media is limited or even negligible:

> A uniform increase in contact with other dialects may also be an effect of the mass media. But all of the evidence generated in this volume and elsewhere points to the conclusion that language is not systematically affected by the

[4] There are two distinct definitions of diffusion in contemporary sociolinguistics. One refers to the process of a linguistic element entering previously constrained (syntactic) environments (cf. e.g. Bybee 2002). In this study, diffusion is considered synonymous with the concept of spread and refers to a process of spreading from one speech community to another.

> mass media, and is influenced primarily in face-to-face interaction with peers. (Labov 2001:228)

The hypothesis that mass media lack a substantial effect on language change rests on many studies conducted in American cities but also other locations, e.g. Britain, Canada, Australia and New Zealand (Tagliamonte 2001:41). Most of these studies have, however, focused exclusively on phonological change and dismiss spread of lexical innovations. With respect to the spread of lexical elements, Romaine (1994:34) shows that the lexical innovation "nerd" emerged in Scandiinavia through an American movie and concludes:

> … the possibilities for change of this type are indeed enormous nowadays, considering how much more mobile most people are, and how much exposure people get to speech norms outside their immediate community through the mass media. (Romaine 1994:34)

Although more recent accounts put stronger emphasis on the mass media as a method of diffusion - which has so far been disregarded (Hickey 2003:360) - the effect of mass media on language change remains controversial. Tagliamonte, for instance, summarizses the current state of affairs as follows:

> In some cases, media language appears to faithfully reflect ambient community norms. The forms and ranking of intensifiers *very*, *really*, and *so* in the television series *Friends* mirrored reported usage (Tagliamonte and Roberts 2005). However, a study of quotative *be like* in American film found neither sufficient tokens nor the patterns (i.e. constraints) that had been consistently reported in the literature (Dion and Poplack 2007). This suggests that the rapid spread of *be like* in North America was not the result of, nor influenced by, the media. (Tagliamonte 2011:41)

Meyerhoff and Niedzielski (2003) similarly assert that the recognition of constraints and the functionality of innovations relys upon quality face-to-face interaction. In their variationist study of *be like* in NZE, they show that cases of superficial contact, as in the case of acquisition via mass media, only transmits

superficial information about the innovation at hand (Meyerhoff & Niedzielski 2003:537-538). In cases of geographically non-continuous or even remote varieties, this implies that the variety-specific functionality, positioning and attitudes which undergo more severe re-negotiation presumably lead to rather distinct usage profiles and positional distributions. The crucial point here is to differentiate between *transmission*, which refers to native-language acquisition by children (cf. Labov 2010:307) and *diffusion*, which refers either more generally to the transfer of features from one speech community to another or, more specifically, from one adult speaker to another (cf. Labov 2010:308). In most cases of borrowing, the processes we observe are best described as diffusion rather than transmission, although transmission will take place after the innovation has become nativized. Another important difference between transmission and diffusion relates to the fact that transmission faithfully preserves the language variation patterns while diffusion does not (Hazen 2010:12). In cases of transmission, intra-linguistic constraints are preserved because children readily learn underlying grammatical usage constraints: in cases of diffusion, constraints are lost because adults are not apt to learn the underlying grammar of innovations. This difference is crucial, as Labov considered the preservation of intra-linguistic constraints during L1 acquisition as a key difference between diffusion and transmission (cf. Labov 2007).

The important issue at hand is, thus, not only to explore which processes are at work in cases of diffusion from one speech community to another but also to review which processes are at work when linguistic innovations diffuse throughout the respective speech communities.

LIKE is an ideal item for investigating such processes; this vernacular feature is common, widespread, and currently undergoing rapid change as it spreads throughout the English-speaking world (Tagliamonte 2005:1898). In a study pursuing similar ends, Buchstaller and D'Arcy (2009) hypothesize that "global innovations must be considered in light of local systems into which they are adopted […] [and that] the form and amount of contact must be correlated

with respect to the knowledge transfer they allow" (2009:291). In other words, in cases of global diffusion (in this study synonymous with global spread) moving outward from an original epicenter, the innovative features are not simply adopted, but their implementation is accompanied by re-contextualization, re-organization and re-negotiation of their meaning (cf. Kachru 1992; Buchstaller & D'Arcy 2009:292-294). These linguistic newcomers thereby undergo transformation and adaptation when interloping into pragmatic niches of the respective varieties as a result of the social and linguistic local underpinning. Consider Buchstaller and D'Arcy (2009:317-318):

> [I]nstead of simply accepting or rejecting an innovation, potential adopters are often active participants in the diffusion process, struggling to give meaning to the new idea as it is applied to their local context. […] In other words, global resources are negotiated in situ as they are integrated into pre-existing local norms and practices.

In Labovian terms, such processes constitute "contact across (national) communities in which speakers (often adults) acquire new variants from an originating community (i.e. diffusion; Labov 2007)" (Buchstaller & D'Arcy 2009:291-292). The theoretical implications of such scenarios are of particular relevance here: how stable are supra-local or universal trajectories concerning the diffusion of innovative forms through social strata in geographically distinct settings, and to what degree do such processes of adaptation and adjustment lead to locally distinct patterns of social stratification?

Following Britain (2002:618), there are at least three scenarios in cases of contact between a global or supra-local innovation and the norms and practices of local speech communities: (i) wholesale adoption; (ii) flat rejection; or (iii) interaction between the globally available form and its local implementation. Investigating the mechanisms at work in such scenarios, Buchstaller (2008) as well as Buchstaller and D'Arcy (2009) have indicated that linguistic elements are not simply borrowed or adopted wholesale (i.e.

Britain's (2002:618) first scenario) but they undergo re-interpretation as they are adapted to the local systems. This suggests that LIKE, when implemented in local systems, undergoes similar modifications leading to regionally distinct usage patterns. Indeed, this not only applies to the linguistic elements themselves, but also to the attitudes attached to them. An attitudinal study conducted on *be like* in EngE by Buchstaller (2006b), for instance, strongly suggests that "the adoption of global resources is a more agentive process, whereby attitudes are re-evaluated and re-created by speakers of the borrowing variety" (Buchstaller 2006b:362). Buchstaller's (2006b) analysis, therefore, calls into question the universality of the association of vernacular uses of LIKE with female adolescents leaving room for regionally distinct associations between LIKE and possible reference groups and social categories. In other words, the assumption that LIKE is generally associated with the California "Valley Girl" persona, as attested to by D'Arcy (2007) for North American English, may not hold for other regional varieties of English. This "reallocation of attitudes" (Buchstaller 2006b:363), may impact the gender- and age-specific usage patterns leading to a diversity of variety-specific distributions.

The increasing attention paid to the "sociolinguistic mechanisms of globalization" (Buchstaller 2008:15) poses an additional, albeit related, problem. For the past thirty or so years, sociolinguistic studies have focused on monolingual settings and do "not deal with influences that may stem from dialect contact or the substrate effects of other mother tongues" (Labov 2001:518). The more common case of linguistic and cultural diversity in multilingual settings has for the most part been neglected (Sture Ureland 1989:242-245):

> The overall perspective is excessively monolingual and only inner-linguistic evolutionary processes are stressed, assuming that contact is secondary and of minor importance. […] It is a monolingual world without bilingual or bilectal speakers, in which each language or dialect functions completely independently of all other languages and language varieties.

The underlying premise of a monolingual speech community is, however, more often than not fictitious, particularly in second language varieties such as IndE. The stability of recurring patterns in sociolinguistic diffusion may erode if applied to different local systems in which innovative forms meet linguistically diverse settings and quite different competing variants. A case in point is Sankoff et al.'s (1997) study of discourse marker use in a contact situation. Indeed, Sankoff et al. (1997:213) found that the use of French *comme* is very suggestive of interference effects from English vernacular uses of LIKE. This implies that various issues relating to possible factors influencing linguistic behaviour have, unfortunately, been neglected.

Before discussing possible shortcomings of contemporary sociolinguistic theory, let us survey its basic theoretical concepts and mechanisms.

2.4 Sociolinguistic variation and change

The aim of discovering general laws of linguistic change is by no means recent. Such general principles would allow one to predict that change, once initiated, "will move through the speech community in a uniform fashion" (Labov 2010:184).

On a general level, two types of change can be distinguished in sociolinguistics: (i) change from above and (ii) change from below. The distinction between *above* and *below* was introduced by Labov (1966) in his New York City study (Labov 1966) and refers "simultaneously to the level of social awareness and position in the socioeconomic hierarchy" (Labov 1994:78).

"*Change from above* is introduced by the dominant social class, often with full public awareness" (Labov 1994:78). Hence, this type of change is a conscious process involving the adoption of linguistic features due to their status as prestige variants. Borrowed linguistic elements are commonly lexical, but also extend to phonological or syntactical features perceived as carrying

prestige in the view of the socially dominant class. The occurrence of these prestige variants is, nevertheless, mostly restricted to careful speech "reflecting a superposed dialect learned after the vernacular is acquired" (Labov 1994:79). The connotations these words carry communicate membership in a higher social milieu, thus serving an accommodative function.

Changes from below are systematic changes operating mostly below the level of social awareness. It is only when such changes are nearing completion that the speech community becomes aware that they have been taking place. Changes from below are, predominantly, introduced by social groups which can be located in the middle of the socioeconomic continuum and involve the change of features due to language-internal causes.

Throughout the past three decades, various studies have examined these types of changes. The amount of research has resulted in a more detailed and enlarged database which has subsequently led to a much better and more accurate understanding of the general trends and principles underlying linguistic change. With respect to the rate of change, the picture which has emerged over the past years, resembles an S-shaped curve which appears to be underlying most changes studied so far. An idealized graphical representation of the S-shaped curve is provided in Figure 1.

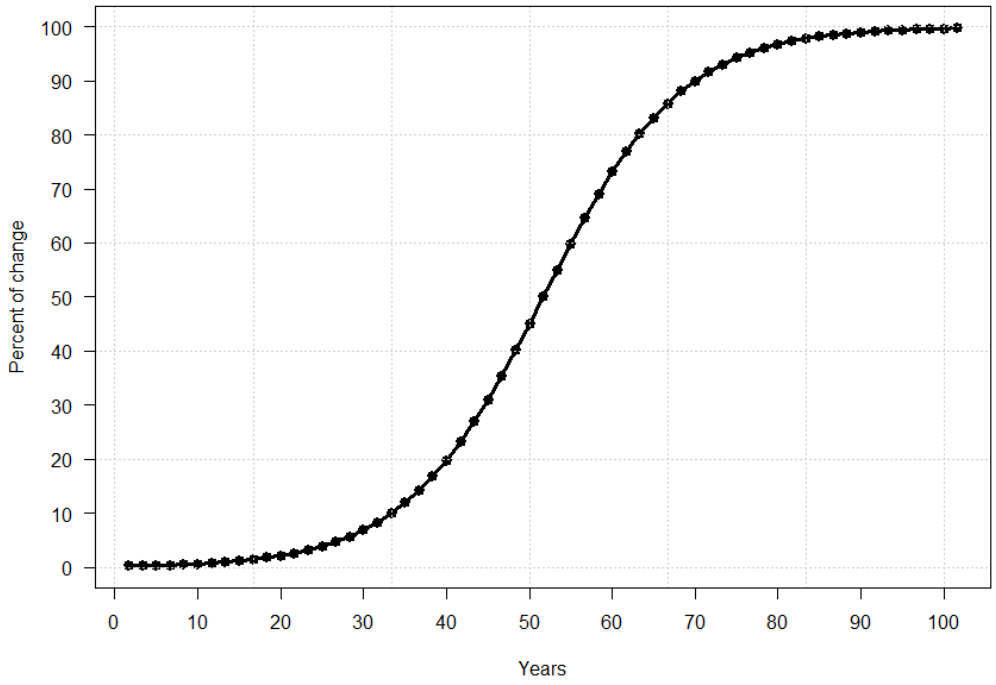

Figure 1: S-shaped curve representing the rate of change (cf. Labov 2001:450)

The S-shaped curve indicates that change does not follow a monotonic, linear rate: the rate of change is rather slow initially, then increases exponentially and finally slows down after the incoming form has spread and entered most environments. Indeed, the trajectory expressed by the S-shaped curve has given rise to the comparison between language change and the spread of infectious diseases. In this sense, the diffusion of linguistic forms is equivalent to an "epidemic spread throughout a population. At first, only a few persons are affected. Then, the disease or change picks up momentum and finally runs its course" (Romaine 2005:1698). Similar to the spread of infectious diseases, the explanation for the non-linearity of change is the amount of exposure of speakers to incoming forms (cf. Labov 1994:66): while speakers using an old form are only rarely exposed to the new form initially, the amount of exposure increases as soon as more speakers adopt the incoming form, leading to an increased amount of exposure of "old-form" users. At midpoint, the amount of exposure reaches a maximum and subsequently

decreases as the conservative form is only rarely used and exposure to this variant is reduced to a minimum.

According to Labov (1994), different intervals of the S-shaped curve represent five distinct phases of change which differ quantitatively with respect to the percentage with which an incoming variant is used (cf. Nevalainen & Raumolin-Brunberg 2003:55):

(i)	Incipient	below 15 percent
(ii)	New and vigorous	between 15 and 35 percent
(iii)	Mid-range	between 36 and 65 percent
(iv)	Nearing completion	between 66 and 85 percent
(v)	Completed	over 85 percent

Moreover, these phases can be characterized by different levels of social stratification and gender differentiation. For instance, the incipient phase and the incoming form are not yet associated with a certain social reference group. The distinctive features of incipient changes are that only a few speakers have adopted the incoming form and that its overall use is still infrequent.

On the other hand, new and vigorous changes are marked by a rapid increase in the respective variant and pronounced age-grading, as the innovative variant hardly exists in the speech of older. In addition, new and vigorous changes have commonly not breached the level of social awareness and are thus "never referred to in discussions of language with community members" (Labov 1994:82). During this phase of change, incoming forms are responsive to social situations and show consistent patterns of social evaluation, although they are not yet subject to correction in careful speech (Labov 1994:82). During midrange, the rate of change begins to decrease, resulting in a lower degree of age-grading. When changes are nearing completion, they are widely used, but show a notable degree of social stratification. The age coefficients continue to decrease as the incoming variant proceeds to intrude into the speech of more conservative speakers. Finally, once a change is complete, it "is completely below the level of social awareness" (Labov 1994:79) and age-grading is disappearing.

So far, the account of linguistic change and variation has focused on the general tendencies of the spread of forms itself which is accurately reflected in the S-shaped pattern displayed in Figure 1. But while this S-shaped curve is a neat display, the complexity of processes underlying change substantially exceeds this rather sketchy pattern. Following Labov (2001:308), stages of change are characterized by specific gender differences in the rate of the incoming form. Thus, if the gender of speakers and their linguistic performance are both considered, the five-stage model presented above can be refined to show six distinct stages:

In stage zero, no gender difference is to be expected as the incoming form is not yet associated with a reference group – for example, the vernacular of the female caretaker.

The first stage is identified as the phase in which an incoming form is associated with a specific reference group, while the incoming form begins to be associated with one or the other gender only during the second stage of the change.

In its third stage, gender differentiation takes place as "males in the lower social classes show a consistent pattern of retreating from or resisting female-dominated change" (Labov 2001:308). The reverse effect, on the other hand, seems to be negligible as females appear not to retreat or resist male-dominated change.

In its fourth stage, a first generation acceleration takes place as the children of young mothers who have already acquired the incoming form enter the speech community. Males learn to use the particular form and show a sharp increase in its use, while females show steady linear progression in frequency. In other words, "[M]en are at the level of linguistic change characteristic of their mothers" (Labov 2001:306-307).

The fifth stage is marked by a second generation acceleration, when young males – i.e. the children of stage two mothers – receive a second step-wise acceleration, while females continue to progress linearly.

In its sixth stage, the change approaches its completion and the gender differentiation weakens, but it will only disappear when the whole speech community adopts the now mature and integrated form. If the form is associated with a specific social reference category, however, it will typically develop a linear alignment with this group or social class. In addition, one commonly observes an interaction between social class and gender, meaning that the degree of gender differentiation differs from class to class. For example, women tend to use less overtly stigmatized non-standard features while males typically show higher rates of overtly stigmatized non-standard features.

To summarize, most processes of change follow a general scheme: while in its initial phase, change proceeds with a very low rate, increasing exponentially once the linguistic feature becomes associated with a specific social group. It is only then that gender differentiation sets in, with females taking the lead in the majority of cases while males commonly fail to adopt the new feature. After about one generation, when this linguistic feature has spread sufficiently and exhibits substantial gender differentiation, the youngest male group begins adopting the incoming feature as they acquire the incoming form from their female caretaker, i.e. in most cases, their mothers. At this point, the gender difference begins to level out, while the rate of change begins to decrease as in most environments the incoming feature has replaced its competitors. The adoption of the feature by male children is repeated after another generation, leading to another rapid decrease in gender differentiation which begins to diminish as the change nears completion. The feature loses its association with a certain gender and, as a consequence, the gender difference wanes and will subsequently disappear. Finally, the change is complete, and the incoming form has intruded into most environments. The remaining environments which the new form has failed to intrude into are often calques, i.e. environments which resist change and become idiomatic.

An idealized graphical representation of the resulting nearly linear progression of female speakers and the step-wise progression of males including the respective phases is provided in Figure 2.

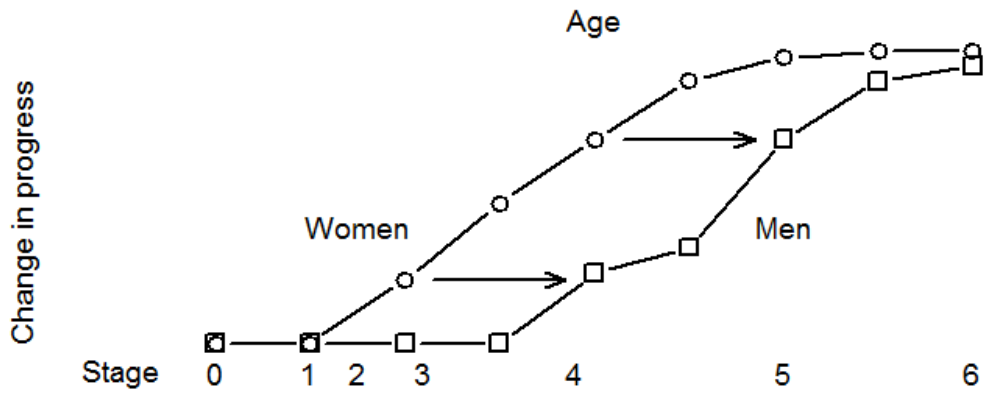

Figure 2: Six-stage model of gender relations in linguistic change from below (cf. Labov 2001:309)

The arrows in Figure 2 to indicate the direction of influence go from female to male (cf. Labov 2001:309).

2.4.1 Real-time and apparent-time

So far, we have discussed how change proceeds but not the data on which the analysis is based. There are two ways in which we can analyze language change and variation: from a real-time and from an apparent-time perspective. Both approaches form the fundamentals of sociolinguistic research that has been carried out over the past 30 years. Figure 3 presents the difference between the real-time perspective and the apparent-time construct.

As shown in Figure 3, real-time studies analyze certain linguistic variants by drawing samples at different points in time, while apparent-time studies analyze the distribution of a form on one particular date and try to recreate its spread by analyzing the frequencies of the respective form across age cohorts. Accordingly, real-time analyses study differences in the frequency or rates of

incoming variants by collecting data at two or more distinct points. If this variant is significantly more frequent in the data compiled at a later point than it is in the earlier data and all other variables such as age, gender and social class of speakers are held constant, the assumption that the form has increased in frequency, i.e. it has undergone change, is justified.

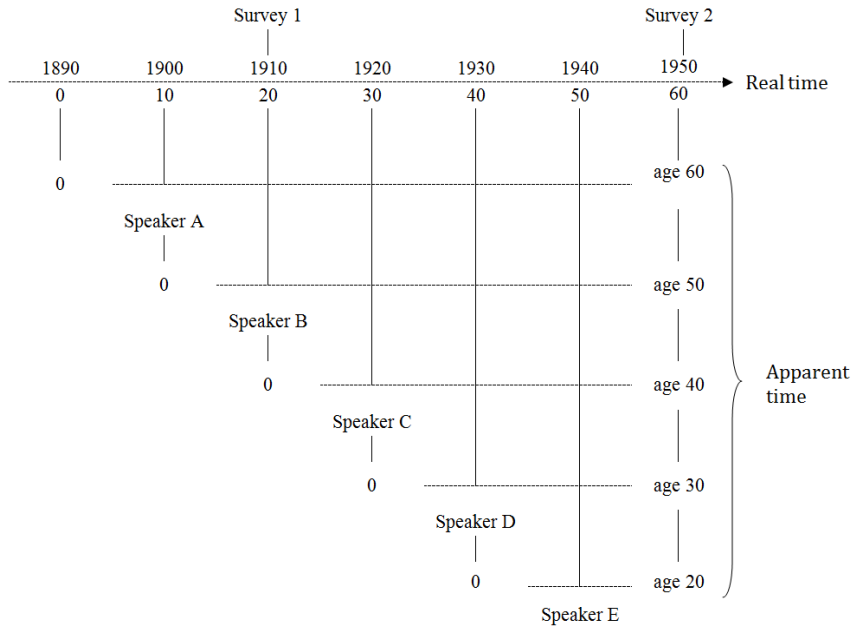

Figure 3: Real- and apparent-time in language change (cf. Downes 1998:238)

Real-time studies can be conducted in two distinct ways: as trend studies or panel studies. Trend studies replicate "an earlier study with the same population and the same methodology in sampling and analysis" (Nevalainen & Raumolin-Brunberg 2003:55). For this kind of approach to yield reliable results, it is necessary that other factors which can influence linguistic performance have remained constant. On the other hand, panel studies collect data, e.g. by means of interviews, from the same individuals at different points in time.

Real-time data provide the best material for tracking and analyzing the patterns and progression of ongoing change. However, in most cases data for real-time analyses is unavailable; the apparent-time construct is utilized because it provides the next best tool, and is most effective when the speakers

under consideration share the same background. The assumption underlying the apparent-time construct is the notion that differences in the frequency of forms by speakers belonging to distinct age groups reflect different historical stages of communal grammars. Thus, Chambers and Trudgill (1998:151) write:

> Studying the diffusion of innovations in apparent time involves surveying the differences between the speech of people of different ages in the same community, while controlling the other variables such as sex, social class and ethnicity.

Accordingly, apparent-time studies rely on data collected at only one point in time.

The assumption that apparent-time distributions reflect real-time change is, however, not unproblematic as the same patterns of variation which reflect ongoing linguistic change may also reflect stable age-grading, i.e. no change at all. This means that monotonic age-grading is not sufficient as evidence for ongoing change. Therefore, apparent-time analyses require additional information to discern stable age-grading from ongoing change. This is particularly relevant in cases where the analysis of a form is based solely on apparent-time data. Although it has become a common approach in sociolinguistics to limit analysis to apparent-time data, this is highly problematic as the same apparent-time distribution can be the result of age-grading, communal change, or generational change. The difference between these types of change is crucial as age-grading does not necessarily entail ongoing change; it can occur in stable conditions, while generational change is the prototypical variant of ongoing change. This distinction between types of variation is essential, because prematurely stipulating ongoing change purely based on monotonic age grading will more often than not lead to erroneous interpretations of the data. This erroneous attribution of the cause or pattern of the variation observed within communities leads almost of necessity to a flawed understanding of the individual patterns and mechanisms underlying the observed variation. This is essentially why the type of relation between

apparent-time and real-time is crucial, particularly with respect to age-grading versus generational change in the present analysis.

A schematized representation of these types of relations between change and variation is provided in Table 1.

Table 1: Patterns of change in the individual and the community (cf. Labov 1994:83)

		Individual	Community
1	Stability	Stable	Stable
2	Age-grading	Unstable	Stable
3	Generational change	Stable	Unstable
4	Communal change	Unstable	Unstable

As shown in Table 1, different types of relationships between variation and change express themselves as distinct combinations of individual and communal linguistic behavior, e.g., a form is stable, that is, not undergoing change, if neither the individual nor the community changes linguistic performance. To clarify these distinct types of variation and their relation to change, consider Labov's definitions of the respective types of variation and their relation to linguistic change in (2).

(2) Patterns of Linguistic Change (Labov 1994:84)

> Age-grading
>
> "If individuals change their linguistic behaviour throughout their lifetimes, but the community as a whole does not change, the pattern can be characterized as one of age-grading."
>
> Generational change
>
> "Individual speakers enter the community with a characteristic frequency for a particular variable, maintained throughout their lives; but regular increases in the values adopted by individuals, often incremented by generations, lead to linguistic change for the community."

Communal change

"In communal change all members of the community alter their frequencies together or acquire new forms simultaneously."

Age-grading refers to situations in which the pattern of variation remains stable on the communal level as younger speakers adapt their language to the norms of elders as they themselves grow older. This kind of change in linguistic performance is particularly difficult to detect in apparent-time designs. In contrast to stabilized age-grading, cases of generational change occur when speakers retain their grammar once it has stabilized at about the age of 17 (Labov 2001:451). This is shown in Figure 4 while the frequency of the linguistic phenomenon increases in the speech community as new speakers continually acquire higher frequencies of the respective form.

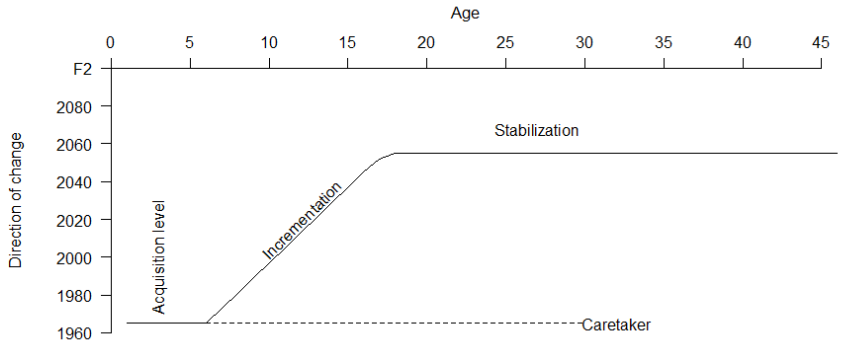

Figure 4: A linear model of incrementation for a single speaker from 1 to 45 years of age (cf. Labov 2001:448)

Figure 4 depicts the acquisition, incrementation and stabilization of an incoming form in the speech of one individual. Initially, the child acquires the incoming form from the primary caretaker and uses it with an almost identical rate as its linguistic role model. Starting at age 4 and continuing for the next 13 years, i.e. during the formative years, the incoming form will increase in frequency until incrementation stops at the age of 17, and the rate of use stabilizes.

Figure 5 illustrates the consequences of this process for a speech community which has been undergoing change since 1925. The progression of lines illustrates the synchronic distribution of an incoming form with respect to the age of speakers at four distinct points in time (1925, 1959, 1975, 2000). The rate of the incoming form is assumed to stabilize in individual speakers aged 17 to 25; that is, speakers retain the level of use.

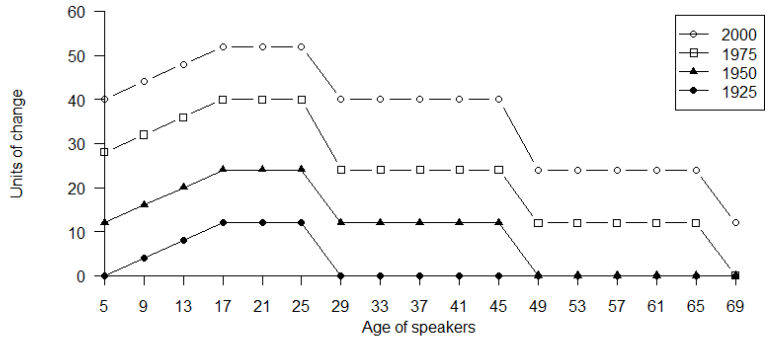

Figure 5: Age profiles of linguistic change in progress with uniform incrementation of the change (cf. Labov 2001:449)

According to this idealized display, an incoming form will increase uniformly over time as a result of the caretaker's effect on the child's adopting this new feature, and the increasing rate of use during incrementation. Caretakers transmit their level of use of the incoming form to their children, who therefore start off with a higher baseline rate. As speakers stabilize at age 17, in each generation their children will be exposed to higher frequencies of the new feature which will subsequently increase over time as the overall rate of use increases within the speech community.

In addition, Figure 5 displays another feature typical of ongoing change which is an apparent-time peak in the use of incoming variants near the age of stabilization, i.e. in adolescence. While this peak was not expected in earlier research on ongoing change, results of a peak in usage rates before age 20 accumulated over the next few years led to a revision of earlier sociolinguistic theories. In fact, the apparent-time peak among (female) adolescents is now

considered to be a general requirement of change in progress (Labov 2001:455). Furthermore, the existence of a peak depends on the leaders of change, as female adolescents will only exhibit a peak at about 13 to 17 years of age if the change is female dominated. Male adolescents, on the other hand, follow a near linear trajectory in female dominated changes and thus do not exhibit the peak in such female dominated changes (Labov 2001:456).

However, if an apparent-time analysis is based solely on the age and gender of speakers and is not accompanied by a complementary real-time analysis, issues relating to the exact type of change remain unresolved. To address this issue without requiring additional real-time data, it has become fashionable to look at other extra-linguistic variables, i.e. social class. The benefit of taking social class into account lies in the fact a curvilinear pattern of usage rates among speakers with differing socio-economic statuses may becoming visible. If this pattern emerges, it suggests ongoing change. In stable situations, on the other hand, "flat age distributions for adults combine with a monotonic social stratification" (Labov 2001:460).

2.4.2 Social class

Social class is considered by many sociolinguists, most prominently by Labov, to be the essential and primary extra-linguistic variable. This view is supported by the fact that "most of the linguistic changes in progress studied in the 2nd half of the 20th century show a high degree of social differentiation" (Labov 2002). In fact, social stratification is viewed as primary, while other extra-linguistic variables gain importance only during a subsequent phase of re-interpretation. Consider Labov (2001:320, 322):

> There can be no doubt that the social category of gender is part and parcel of the class system […], and that women's behavior is related to their social status. As Eckert's work shows, the "Burnout" and "Jock" categories are age-specific transformations of the *more general social class categories*. [emphasis MS]

Gender and age differences are thus regarded as secondary epiphenomena which depend on and are parasitic to pre-existing social stratification. From this perspective, social stratification is a prerequisite for linguistic change as change builds on the association of linguistic forms with certain social groups:

> The major acceleration of diffusion within a community is the result of the reinterpretation of social category differences as gender differences. Once a social category is reinterpreted as gender difference, it is subject to the basic mechanism of reinterpretation as the unmarked community level, leading to rising levels of linguistic change for the community as a whole. (Labov 2002)

Nevertheless, despite the common association of social stratification with linguistic variation the exact nature of the social model underlying linguistic variation usually remains underspecified (cf. Milroy & Gordon 2003:95; Cheshire 1991:3). Indeed, a unified theoretical framework which is able to account for the

> recurrent and robust correlation between language and class has not been clearly articulated. Furthermore, several principles based on undertheorized assumptions about stratificational social class have become almost axiomatic. (Milroy & Gordon 2003:95)

The lack of a coherent theoretical framework is particularly obvious in relation to the concept of social class. On a general level, a social class refers to a group of people sharing "similar occupations, incomes, life-styles and beliefs" (Milroy & Gordon 2003:95). Despite the lack of a coherent model of social stratification, most sociolinguistic studies quite pragmatically borrow their scaling of socioeconomic status from social sciences or pre-existing data sets. The implications of the underlying model are, however, only rarely addressed. In essence, there are two approaches to the social structure in which speech communities are embedded: the consensus view and the conflict view of class relations.

Consensus models assume the existence of shared norms and the common evaluation of linguistic features (cf. Milroy & Gordon 2003:96). Consensus

models are particularly suited for explaining higher rates of standard features in more careful speech implying that speakers with a lower socioeconomic status are aware of linguistic norms shared by the community as a whole. Conflict models, on the other hand, perform better when explaining stable socially stratified patterns, such as the persistence of non-standard vernaculars in certain lower-class varieties, or in certain ethnic groups e.g. African-American Vernacular English (Milroy & Gordon 2003:96). The difference between consensus and conflict models is particularly relevant with respect to the interpretation of distinct linguistic behaviors in a single geographic area. Interpretations of ethnically motivated language use in a framework based on a consensus model stipulate the existence of two or more separate speech communities, while conflict models draw attention to social inequality related to ethnicity, social identity as well as access to and participation in the linguistic market (cf. Milroy & Gordon 2003:96). Despite lacking a coherent theoretical framework, the results of sociolinguistic analyses are extremely consistent and provide detailed insights into the location of leaders of change as well as the interaction of social class with other extra-linguistic factors.

In contrast to traditional theories of language change, according to which the leaders of linguistic change are located either at the top or the bottom of the social hierarchy, modern quantitative approaches have indicated that more often than not it is speakers from the middle class who are most important for diffusing change throughout speech communities (Labov 2001:31). Research on the social stratification of linguistic forms and the social milieu of leaders in language change in the second half of the twentieth century (e.g. Haeri 1996; Labov 1990) have corroborated this so-called 'Curvilinear Principle' according to which "[l]inguistic change from below originates in a central social group located in the interior of the socioeconomic Hierarchy" (Labov 2001:188). The curvilinear hypothesis stands in stark contrast to traditional models of language change which equated language change with linguistic decay and characterized innovators as "uncultivated and careless" (Whitney 1904).

Traditional theories of this type posed that linguistic change was observable mostly among the lowest social classes, either gradually moving upwards or not affecting the speech of the "educated and lettered class" (Whitney 1904). The opposite view was held, for example, by Tarde (1873) who claimed that linguistic change sprang from the highest social class and "believed that the condensations of rapid speech represented an increase in efficiency of speech, and were characteristic of the most intelligent and educated speakers" (Labov 2001:30-31). Thus, the growing bulk of research strongly suggests that the social location of innovators and innovating groups are the upper working and lower middle classes which are almost identical in their "advancement of the change in progress in vernacular speech" (Labov 2001:32).

An alternative approach towards socially stratified behavior which elaborates and extends the curvilinear pattern is promoted by Kroch (1978), who views socially motivated linguistic variation in terms of residence and promotion of innovation. According to Kroch (1978), the socioeconomic continuum is also a continuum of linguistic practice in which greater socioeconomic status brings greater resistance to change. This models shares certain theoretical properties with Labov's (1966) interpretation of socially stratified language use which asserts the correlation between the degree of attention being paid to linguistic performance and the socio-economic status of speakers. The more speakers attend to linguistic performance, the less likely it is that this speaker employs non-standard forms. Hence, innovations, which are almost universally viewed as deviations from the pre-existing standard, are more likely to occur among speakers who do not closely monitor their linguistic performance (cf. Eckert 2000:16-17).

Eckert's (1989) analysis of adolescent speech at Belten High in Detroit, Michigan (USA), on the other hand, stresses a conflict model view of proactive identity marking. Adolescent speakers marked their social identity by active and more or less conscious adoption of linguistic features associated with one of the two distinct social identity groups, i.e. jocks and burnouts. The jocks and burnouts distinctions are reflected in more than mere linguistic performance,

extending to other lifestyle issues such as clothing. According to this view, linguistic means serve to mark the association with distinct communities of practice and, hence, perform a symbolic rather than a linguistic function. With respect to linguistic standards, jocks and burnouts differ in the sense that jocks prefer supra-regional, and more suburban styles, while burnouts mark their identity by using forms associated with the local community, similar to the use of linguistic features reflecting the local "insular culture" found in Labov's (1972) Martha's Vineyard study.

Just as Eckert (2000) found that the linguistic performance of jocks and burnouts at Belten High reflected the occupation of their parents to a very limited extent, the assignment of social class in the present study is restricted to speakers who have entered the work force and are no longer in or have never entered higher education. In other words, university students are not considered in analyses which focus on social class.

2.4.3 Gender

Although sociolinguistic studies use the binary distinction of biological sex as an independent variable, the concept they are interested in is not the biological distinction, but the socio-cultural construct of gender. Indeed, sex and gender are often considered synonymous in sociolinguistic studies, though "sex […] is a biological given; [while] gender is a social acquisition" (Miller & Swift 1976:51; cf. Chambers 1995:103). The reason for employing the binary biological variable instead of the socio-cultural concept of gender is simple enough: a speaker's sex is readily observable while "an individual's gender-related place in society is a multidimensional complex that can only be characterized through careful analysis (Eckert 1989:146-247; cf. Chambers 1995:104). Therefore, the sex of speakers is used as a proxy for gender, which refers to a social category in light of which the meaning of sex-differences is interpreted (Milroy & Gordon 2003:100). In other words, biological sex serves as an approximate measure for socio-cultural gender, and we therefore assume

that "differences in patterns of variation between men and women are a function of gender and only indirectly a function of sex" (Eckert 1989:246-247).

Taking sex as a measure for gender is, however, not unproblematic since gender is not a binary category like sex, but "rather a continuum where speakers situate themselves socially between two reference poles" (Milroy & Gordon 2003:100).

This implies that gender can only be interpreted with respect to its interaction with other social categories. The interconnectedness of gender and other social variables such as social class is particularly pronounced in a Labovian variationist view of gender differences. According to this view, social differences are the primary factor determining linguistic performance, while other social factors such as gender or age have often been explained "with reference to, and as dependent on, social class" (Milroy & Gordon 2003:1010). For example, it has been suggested that women exhibit a higher degree of social awareness and that they conform more to overtly prescribed norms than men of equal social status. Indeed, women show consistently lower levels of stigmatized forms than men, thus showing a higher degree of conservatism with respect to prestige variants (Labov 2001:265). This general tendency is expressed in the principle of linguistic conformity of women according to which they avoid stigmatized forms and exhibit a strong preference for prestige variants (Labov 2001:266). This behavior has given rise to what Labov (2001:292-293) termed the 'Gender Paradox' according to which "women conform more closely than men to sociolinguistic norms that are overtly prescribed, but conform less than men when they are not" (Labov 2001:293). As Holmes (1997) points out, there has not as yet been a satisfactory explanation for this gender difference.

An even more pressing issue of approaches which stipulate a primacy of class over gender is related to the interaction between social-economic status and gender-specific linguistic behavior. A growing number of studies indicate

that class is not necessarily the dominant or primary factor, but that gender – or age for that matter – is more relevant with respect to determining linguistic variation. Horvath (1985), for example, concluded that when reorganizing Labov's New York City data, some variables "are more clearly stratified by gender than by class" (Milroy & Gordon 2003:102). Dubois and Horvath (1999) who studied the interaction between gender and other social variables, furthermore found that "gender indexing affects language quite differently in each generation and each generation orients towards different norms" (Milroy & Gordon 2003:105). Their data highlight that social class is not universally the determining factor and that de-contextualized generalizations about the effect of gender are problematic because they neglect and, hence, underestimate, the interaction of linguistic behavior and gendered roles in specific social contexts (cf. Milroy & Gordon 2003:106).

The most salient feature among the general trends substantiated through numerous studies and described even in the earliest investigations is a stable sex differentiation, i.e. the tendency of women to take the leading role in language change. This major tendency of women to be advancing in linguistic change from below has been confirmed by quantitative studies of change in distinct social and geographical conditions (e.g. Cedergren 1973; Haeri 1987; Labov 1966; Trudgill 1974). The role of women is especially crucial at the onset of change as they serve as initiators, while males adopt incoming forms only later in the process. As a general tendency, female are approximately one generation ahead of males in their rates of incoming variants (Labov 2001:294).

The degree of sex differentiation indicates the current status or phase of ongoing language change, i.e. whether an innovation has only recently been introduced, is currently undergoing rapid change, or has stabilized:

> Depending on the stage of the change within the purview of the investigators, we see females diverging from males, as in Vancouver; females advancing ahead and in parallel with males, as in Toronto; or converging with advanced position of females, as in Buenos Aires and Hong Kong. (Labov 2001:283)

Although there is evidence for male-dominated changes, the leading role of women is the unmarked case, because male-dominated changes form a "small minority" (Labov 2001:284). Examples of male-dominated changes are, for instance, the centralization of (ay) and (aw) in Labov's 1963 Martha's Vineyard study and the backing and rounding of (a) in Belfast observed by Milroy and Milroy (1978; cf. also Labov 2001:284). Males have been assumed to dominate change as they were found to use higher rates of standard forms while females exhibited higher frequencies of local colloquial forms in Muslim-oriented societies (Bakir 1986; Modaressi 1978). Haeri (1987) accounts for this by arguing that in this case, the standard does not mirror the prestige variant: while Classical Arabic is comparable to a standard variety in Western societies, the prestige varieties in Muslim-oriented societies may well be modern urban forms which would indicate "that women in those varieties were behaving like women in other varieties" (Labov 2001:270).

Labov accounts for the sex differentiation in terms of what he calls the 'Eckert Effect', according to which the adoption of certain forms depends on their association with social categories or reference groups. The association of a particular linguistic form with a certain reference group is more or less arbitrary, but once a change has entered the initial stage, it is likely that the incoming form is associated with the first vernacular of a child, i.e. the mother's or caretaker's vernacular. On other occasions, the form may be associated with other social categories and is used to symbolize the speaker's identification with this particular group, as, for example, the association of a speaker with the social category "burnout" in the Detroit high school (Eckert 1989). Eckert (1998) thus promotes a constructivist approach towards gender, which assumes that gender is fundamentally dependent on culture-specific social practice. Eckert (1998) argues that because women are excluded from certain

social domains and thus from ways of externalizing their desired social identity, they employ symbolic, i.e. linguistic means, to locate themselves unambiguously in given social categories (cf. Milroy & Gordon 2003:208). Men, on the other hand, do not rely as heavily on the symbolic resources of language as they possess other means of constituting their social identity. In this view, the dominance of females to lead in language change is caused by female speakers' desire for female speakers to identify themselves with a social reference group. Their male counterparts do not exhibit the same level of identity marking and indicate their affinity with this group by other than linguistic means.

Another point in case highlighting the problems related to a binary, i.e. sex-based interpretation of linguistic variation and the primacy of social class is Schilling-Estes' (1999) study on patterns of variation on the island of Ocracoke off the south-eastern coast of the United States. Schilling-Estes' findings led her to doubt the adequacy of a sex based distinction of speakers compared to a model which locates speakers on a continuum between pronouncedly masculine and feminine. Schilling-Estes (1999) investigated the speech of distinct networks of speakers, which she ranked according to their masculinity and feminity respectively. Based on this criterion, Schilling-Estes (1999, cf. also Wolfram & Schilling-Estes 1995, Wolfram 2008) differentiated between three male and one female group. The first male group, the "Poker Game Network", espouse "values associated with highly masculine fishermen of the island and explicitly exclude women from their weekly poker games" (Milroy & Gordon 2003:107). The second male group consisted of heterosexual males not part of the Poker Game Network, while the third group consisted of gay males who also were not part of the Poker Game Network. The results of her study showed that women and the Poker Game males exhibited diametrically opposed patterns for the use of (aj) while the non-Poker game group males were somewhere in the middle, and the gay males somewhat closer to the typically female pattern. This orderly pattern strongly indicates that the linguistic performance was not based on sex differentiation, but served to mark

group membership and the construction of self-images of the respective group members.

With respect to the present study, the findings by Schilling-Estes (1999) imply the need to differentiate between gender and sex when interpreting the results of patterns: while the variable that is used during data compilation is biological sex, the interpretation of sex differences requires utilization of social gender roles.

2.4.4 Identity, prestige and style

In order to mark differences in formality, monolinguals or mono-dialectal speakers shift style as a reaction to the situation. This shift in style depends upon the diversity of available forms in their communal repertories. This context-sensitive variation is approached differently, depending on the sociolinguistic framework of the respective scholars. Labov (2001), for example, views style-shifting as primarily responsive, i.e. "determined by components of the communicative context. […] More recently, style has been treated less as a response to a set of contextual variables than as strategic, proactive use of available linguistic resources to construct social meaning" (Milroy & Gordon 2003:199). In other words, stylistic variants are used not only to index group membership or the lack thereof; they also serve to create social categories, such as jocks and burnouts (cf. Milroy & Gordon 2003:199).

The response hypothesis is based on recurring patterns of higher rates of stylistically marked variants (substandard forms) in casual speech compared to more careful speech. Labov (1972) interpreted this to reflect a lower degree of attention paid to speech and concluded that linguistic norms prevail across the entire speech community. Following this argument, speakers located higher in the social hierarchy use lower rates of substandard variants as a consequence of a higher level of linguistic awareness.

According to Le Page and Tabouret-Keller (1985) adopting innovative linguistic forms is an 'act of identity' which has the symbolic potential of

externalizing group membership or the self-image of speakers. As certain linguistic forms become associated with group membership, they allow speakers to identify themselves as members of the respective groups and thus to profit from the connotations associated with the group:

> The fact that speakers talk in a certain way is taken as evidence for their desire to be identified with or be differentiated form a particular group. (Labov 2001:191)

One aspect which has received increasing attention with respect to proactive utilization of stylistically marked variants relates to the audience, i.e. the interlocutor, as a determining factor of variation, as proposed by Bell (1984, 2001). But while Bell confutes the assertion that stylistic variation is best accounted for by varying degrees of attention paid to speech, he still regards stylistic variation to be primarily responsive to the addressee. In this view, stylistic variation is a form of accommodation. Accommodating one's style facilitates social evaluation in the sense that speakers adapt their style to meet their desire to express social proximity and distance similar to politeness phenomena.

A similar, yet different, understanding is put forward by Labov in his more recent publications. In his revised analysis of stylistic variation, Labov (2010) goes beyond his earlier take on this matter in acknowledging that stylistic variation is not merely responsive, but that it is functional in the sense that it indexes social relations:

> The uniform patterns of social and stylistic stratification suggest that community members can make use of such variation to place speakers on scales of social distance and social power, and many experiments confirm this view.

According to recent assertions, Labov (2010) views stylistic variation as a display of familiarity, power relations and social hierarchy. Although this

revised understanding is more elaborate than the attention based-model, it is not proactive in the sense promoted, for example, by Eckert (2001).

Eckert's (1989) study of adolescent speech at Belten High School, Detroit, Michigan (USA), carries particular importance. Eckert found that the linguistic behavior of adolescents differed markedly depending on their association with either of two social groups named "jocks" and "burnouts". The contrasting orientations of adolescents transcend linguistic performance and encompass social practices, such as clothing, adornment, demeanor, leisure time activities and hang-out places (cf. Milroy & Gordon 2003:98). This orientation affects linguistic performance with respect to the respective norms and standards jocks and burnouts adhere to. Jocks are characteristically middle-class as they orient to supra-local institutions such as school and college rather than informal local networks. In addition, they orient to suburban rather than urban and to supra-local rather than local places (Milroy & Gordon 2003:98). Eckert's analysis, hence, provided a more elaborate account of the interplay between identity marking and social class with relation to systematic differences in linguistic variation. Schilling-Estes (1998), but in particular Eckert (2001), elaborate on the proactive aspect of style shifting in that they view linguistic variation as a way to create socially distinctive categories. Indeed, Eckert's study shows that the prestige or standard is inseparable from its social embeddedness, as it depends on the individual's willingness to externalize group membership and self-imagery. Hence, neither prestige features nor standard varieties are objective entities, but rely on subjective evaluations of the associations and connotations of linguistic forms with social reference groups or personality traits.

> The chief constraint on this proactive linguistic behaviour […] is the individual's capacity to position herself in relation to a particular style in order to render her choices meaningful. (Milroy & Gordon 2003:209)

In an extension of this framework, stylistic variation is regarded as a reflection of the speaker's self image. Though the externalization of identity

marking in relation to social categories is primary in Eckert's (2001) analysis of stylistic variation, her analysis fails to specify the speaker's motivation with respect to psychological categories. In other words, by using certain stylistic features, which are associated not necessarily with certain social groups but with specific personality features, speakers construct the concept of their own self via linguistic means.

While the considerations so far have focused mainly on inner-varietal or monolingual variation, the following passage will concern itself with the role of prestige and attitudinal factors in contact scenarios. Thomason and Kaufman (1991), in particular, stress that stylistic marking and prestige are not only relevant with respect to inner-varietal identity marking but also with respect to lexical borrowing in contact situations. For instance, Thomason and Kaufman (1991:43-44) point out that a

> social factor that is frequently invoked in discussions of language contact is prestige. Moravcsik (1978:109) puts forward the hypothesis that 'nothing can be borrowed […] from a language which is not regarded [as] prestigious by speakers of the borrowing language'.

But while Thomason and Kaufman (1991:44-45) admit that prestige borrowing is a common phenomenon, they draw attention to the inadequacy of prestige alone and stress the function of socio-cultural and economic dominance as a second and plausibly more relevant factor. One case in point is the fact that lexical borrowing occurs predominantly from the dominant to the non-dominant language, while the reverse is substantially less frequent. Instances such as these challenge the prestige claim and show that even a stigmatized feature may be borrowed due to dominance relations.

> In such cases [true stubstrate rather than an adstratum or superstratum] the dominance relationship is clear, and the interference features are sure to be nonprestigious, if not definitely stigmatized. Moreover, prestige often seems to be irrelevant in cases of borrowing. (Thomason & Kaufman 1991:44)

In the present case, both mechanisms borrowing due to prestige and due to dominance reciprocally reinforce the global spread of vernacular LIKE. In light of these considerations, Andersen's (2001:221) hypothesis that the use of LIKE among London adolescents can be viewed as an American borrowing gains plausibility. Indeed, it is reasonable to assume that the global diffusion of LIKE represents a case of Americanization, or Anglicization of global culture. Consider, for example, Hjarvard (2004:75):

> [T]he media both are vehicles of Anglo-Saxon culture and contribute to the anglicization of global culture […]. The media are more than a neutral channel through which Anglo-American culture spreads; by virtue of their institutional structure and a strong dominance of English-speaking actors in the software industry in a broader sense (i.e., computers, television, music, etc.) they actively contribute to cementing the paramountcy of English over other languages.

2.4.5 Traditional dialectology and the modern variationist paradigm

So far, we have taken for granted that sociolinguistics is a rather new field of study initiated most prominently by Labov during the late 1960s. Nonetheless, modern sociolinguistics has precursors in traditional dialectology or sociology of language, which differ from modern sociolinguistics in several ways. In contrast to traditional dialectology, modern variationist work has mostly neglected regional variation, as it is concerned with universal mechanisms underlying processes of language change and correlations between social stratification and linguistic variation rather than determining isoglosses, i.e. dialect boundaries. The difference between traditional dialectology and modern sociolinguistics is also reflected in their different outlooks and aims, as well as in methodology and data sources. While traditional dialectology for the most part relied on non-mobile, older, rural males as informants (so-called NORMs), modern variationist studies commonly focus on urban speech communities or adolescents. The special attention paid to adolescents indexes

the difference in outlook mentioned earlier, as this subgroup is particularly relevant with respect to introducing and diffusing innovative forms throughout the speech community as a whole. Tagliamonte (2005:1897), for instance, writes that "[t]he innovative features are highly conspicuous and typically associated with the younger generation."

One particularly relevant approach investigating regional differences is the study of New Englishes beginning in the early 1980s (cf. Bauer 2002; Cheshire 1991; Kachru 1992; Kortmann et al. 2004; McArthur 1998, 2002; Pride 1982; Platt, Weber, Lian 1984; Schneider 2003, 2007). The spread of the English language across the world as a result of England's colonial history is unmatched in human history and has made English the lingua franca in various fields such as economics, politics, trade and science. After the decline of the British Empire, English has remained the language of choice in various former colonies which have nowadays become nations in their own right. While the former colonies gained political, economic and cultural independence, English has often remained the language of education, administration and everyday practice, leading to the nativization of local varieties, i.e. the development of regional standards and, hence, the emergence of New Englishes. Nonetheless, the study of regional varieties have mainly focused on structural diversity, e.g. the presence and absence of non-standard features, while for the most part, neglecting the sociolinguistic underpinning of these phenomena.

The diversity of socio-cultural settings is not unproblematic with respect to common sociolinguistic approaches to variation. Cheshire (1991), for example, asserts that concepts – such as the idea of the "speech community" or that of a "native speaker" – are seriously challenged when confronted not with monolingual, but multilingual speech communities. Buchstaller and D'Arcy (2009) raise a similar yet separate issue concerning the validity of cross-variety comparisons. In their cross-varietal analysis of *be like*, Buchstaller and D'Arcy (2009:298) point out that…

> the direct comparability [...] of previous analyses of verbs of quotation in general and of *be like* in particular remains questionable since they tend to be based on dissimilar methodological premises and applications in terms of the definition of the variable and constraints, the form selected as part of the envelope of variation, quantitative methods, and statistical models.

Moreover, the concept of class which has been utilized to explain correlations between stylistic variation and the socio-economic status of speakers (cf. Labov 1972) is not necessarily applicable to societies organized differently. A quite similar critique has been brought forward with respect to the cultural sensitivity of gender roles (cf. Eckert 1998). The fact that language change is inextricably interwoven into the specific socio-cultural settings in regionally distinct speech communities has only recently begun to attract attention among variationists. Based on Dubois and Horvath's (1999) study of a Cajun speech community, Milroy and Gordon (2003:106) direct attention to "difficulties with decontextualized generalizations about the role of gender in linguistic variation and change". Eckert (1998), for example, stresses the dependence of gender on the socio-cultural context in which it is constructed. Indeed Eckert (1998:66) asserts that gendered sensitivity of linguistic practices "differ considerably from culture to culture, from place to place, from group to group, living at the intersection of all the other aspects of social identity" (cf. Milroy & Gordon 2003:106). Furthermore, the linguistic behavior of men and women is based on a sizeable number of cases, "but [these are] scattered across the few communities that have been selected for a sociolinguistic study" (Labov 2001:284). Though more recent studies based on the *Atlas of North American English* (Labov, Ash & Boberg 2006) provide a more general and geographically wider perspective, ongoing change has so far only been analyzed marginally from a global perspective, notable exceptions being Buchstaller (2008), Buchstaller and D'Arcy (2009), and Meyerhoff and Niedzielski (2003). If analyses of ongoing change have investigated change in different speech communities, the aim was not to retrace linguistic change from a global perspective, but to establish to which degree the impact of certain

sociolinguistic variables are independent of the specific regions in which change occurred. In addition, if studies of linguistic changes have taken regionally distinct speech communities into account, they have taken a comparative perspective but have not aimed to retrace change over time from a global perspective.

2.5 Synopsis

This chapter has introduced the theoretical framework of this study and presented crucial concepts of modern sociolinguistic theory of language variation and change. The chapter began by showing that language variation is the default case and that the linguistic system is never homeogeneous, but can best be characterized as orderly heterogeiety. Next, the chapter has discussed how modern theorizing has approached the impact of globalization on language change and showed that classical mechanisms of ongoing change may need modification to adequately describe processes of ongoing change in times of globalized linguistic markerts.

With respect to prominent concepts within the framework of Labivian sociolinguistics, the previous chapter has specified apparent-time and real-time approaches to the study of ongoing change and has discussed various extra-linguistic factors that correlate with the linguistic heterogeneity we observe. In particular, the chapter has discussed different relations between age and linguistic variation, and other factors that affect ongoing change such as typical gender differences, or the impact of prestige and acts of identity on the diffusion of a linguistic innovation through the speech community. Finally, the chapter discussed differences between modern sociolinguistic theory and traditional dialectology, since the present study aims to synthezize these strains of linguistics.

3 Discourse markers: Definition, features, and origin

The following chapter introduces and discusses the concept of discourse markers. More specifically, the following chapter defines discourse markers and their features and argues that LIKE is a member of this group of linguistic elements.

3.1 Discourse markers

In spoken conversation, speakers accommodate their language to their interlocutors, the situation, the expected degree of formality and various other contexts, at the same time using linguistic means to satisfy their communicative needs. In other words, while speakers try to guide the listener's interpretation of what is being said to prevent misunderstanding, they are not free to say what they like and how they like, but have to act according to parameters which govern social interactions in general and, hence, human discourse. Among these social parameters is the type of relationship between speaker and addressee. A necessary condition for linguistic accommodation is, of course, that speakers possess a variety of linguistic options or means to mark, for example, intimacy or distance. One of these means is discourse markers which enable speakers to communicate some form of interpersonal information[5] while simultaneously guiding utterance interpretation by establishing coherence between different discourse elements.

Over the past 25 years, discourse markers have increasingly attracted the interest of linguists. Indeed, research on discourse markers has since "abounded in various fields of linguistics and allied fields" (Schiffrin 2001:54), and become "a growth industry in linguistics" (Fraser 1998:301). To survey the

[5] Be it intra-linguistic, for example, focusing, hedging and relating certain segments of speech or extra-linguistic like buying processing time or communicating that an utterance does not fully adequately express what the speaker had in mind.

ever-growing body of literature on discourse markers, this section starts off by providing an account of what is understood by the term 'discourse marker', introducing different approaches to their study.

Before discourse markers themselves became a major topic in linguistic research, linguists investigated linguistic phenomena such as text cohesion (cf. Quirk et al. 1985), which subsequently led to a more focused analysis of discourse markers and their function in discourse. As traditional accounts tended to focus on discourse functions, discourse markers were assumed to indicate how the speaker views the connection between two linguistic units (Quirk et al. 1985:631-633), thereby serving to create cohesion within and between utterances as well as informing the hearer's interpretation of how discourse units are linked.

Since the early 1980s, when discourse markers began to attract attention in various fields of modern linguistics (Schiffrin 2001:54), research on discourse markers has analyzed their occurrence and their functions in multiple languages.[6] Moreover, their functionality has been analyzed across various genres (Brizuela, Andersen & Stallings 1999; Iyeiri, Yaguchi & Okabe 2005), in language contact situations (Sankoff et al. 1997), from a typological perspective (Fleischman 1999), in synchrony and from a diachronic perspective (Brinton 1996), as well as in L1 acquisition (cf. Schiffrin 2001:54-55) and L2 acquisition (Müller 2005). Although the concept of discourse markers is well known in linguistic analysis, it remains difficult to provide a comprehensive definition that is able to encompass all discourse markers and their features; in fact, "there is no generally agreed upon definition of the term 'discourse marker'" (Jucker & Ziv 1988:1). One of the most widely accepted definitions is given by Aijmer (2002). According to her, "discourse particles seem to be dispensable elements functioning as sign-posts in the communication facilitating the hearer's interpretation of the utterance on the basis of various contextual clues" (Aijmer 2002:2). The label 'discourse

6 See Schiffrin (2001: 54) for a more extensive overview.

marker', thus, applies to such seemingly semantically empty linguistic units as *well*, *but*, *y'know*, and *oh*, which function in cognitive, expressive, social, and textual domains and on different planes of discourse, simultaneously fulfilling various discourse functions (Schiffrin 2001:54). Rather poetically, Crystal remarks that pragmatic expressions function "as the oil which helps us perform the complex task of spontaneous speech production and interaction smoothly and efficiently" (1988:48).

One reasonable way to cope with the multitude of often vague definitions of discourse markers has been introduced by Hölker (1991:78-79), who proposed a list of four basic features present in the vast majority of elements referred to as discourse markers: (1) they do not affect the truth conditions of utterances; (2) they do not add anything to the propositional content of an utterance; (3) they are related to the speech situation and not to the situation talked about; and (4) they have an emotive, expressive function rather than a referential, denotative, or cognitive function (cf. Jucker & Ziv 1998:3)[7]. Although such criteria appear explicit, to actually classify and determine which items are discourse markers and which are not is difficult at times. These difficulties arise from various sources, one of which is terminological confusion created by discourse markers being assigned numerous labels, such as cue phrase (Hovy 1994; Knott & Dale 1994); discourse connective (Blackmore 1987); discourse operators (Redeker 1991); discourse particle (Abraham 1991; Kroon 1995; Schourup 1985); pragmatic marker (Fraser 1987, 1990; Hölker 1991; Schiffrin 1987); pragmatic particle (Östman 1982, 1995),

[7] The definition of discourse markers used here is congruent with the depiction of discourse markers by Aijmer (2002), Brinton (1996), Hölker (1991), Schiffrin (1987, 2001), and to a certain degree Sankoff et al. (1997). The tripartite distinction of discourse markers introduced by Vincent and Sankoff (1992), states that there are "three major types: discourse coordinators, interaction markers, and punctors" (Sankoff et al. 1997: 195). It is disregarded, however, because the discourse marker LIKE would then be classified as a punctor rather than an interaction marker, which seems to be problematic in the light of the functions associated with LIKE in the literature.

semantic conjunct (Quirk et al. 1985); sentence connective (Halliday & Hasan 1976); stance adverb (Biber et al. 1999:560-563), etc.[8] Despite this diversity, a limited number of labels have established themselves as most widely used terms for units here referred to as 'discourse marker' – each highlighting specific characteristics of the respective items. While Schourup, himself a pioneer in the field, referred in his early work to discourse markers as 'discourse particles' (cf. Schourup 1982), in later publications he objected to this label because the term "'particle' has traditionally been a syntactic term" (Schourup 1999:229), reserved for so-called 'modal particles' of German and other languages (cf. Brinton 1996:30), and does not adequately cover the variety of syntactic classes of the linguistic items commonly referred to as 'discourse markers' (cf. Müller 2005:3). Another frequently used label is 'pragmatic marker' (Andersen 1998; Brinton 1996; Caron-Prague & Caron 1991; Redeker 1990) which, in contrast to 'discourse particle', suggests the "relatively low degree of lexical specificity and a high degree of context sensitivity" (Andersen 2001:40). Another reason for avoiding the term 'discourse marker' relates to Fraser's (1990) taxonomy of pragmatic markers, according to which discourse markers are not synonymous with, but are a subtype of pragmatic markers, signaling "a sequential relationship between the current basic message and the previous discourse" (Fraser 1990:383). This distinction is also drawn by Lenk (1997:2), who comments that the choice of terminology is commonly motivated by the focus of the research:

> Studies that investigate *pragmatic* markers often focus more on the interactional aspects between participants that are expressed through the use of particles. One of the most prominent functions of discourse markers, however, is to signal the kinds of relations a speaker perceives between different parts of the discourse.

Thus, the diversity of terminology reflects the diversity of subject areas in which discourse markers are studied, as well as the variety of functions

[8] See Fraser (1998: 301) for a more extensive list.

assigned to them. Despite – or possibly because of – the various theoretical approaches and frameworks which have been employed to account for discourse marker usage, a common agreement on a terminological cover term has not yet been established.

Both terminology and what is taken to be a discourse marker remain controversial. For example, the term 'discourse connective' commonly refers to lexical expressions such as *so*, *now*, *hence*, *thereby*, *therefore*, and *thus* which serve to connect units of discourse, while the term 'discourse particle' commonly refers to items such as *oh*, *uhm*, *um*, and *y'know* which serve quite discrete functions. The term 'pragmatic expression', different still, tends to denote linguistic units which consist of more than one word, e.g. *I mean*, *you know*, and *you see* (cf. Jucker & Ziv 1998:2). "Thus, the multiplicity in the terminology reflects on the one hand the broad range of elements in different languages with broadly comparable functions, and, on the other hand, it reflects distinct theoretical perspectives" (Jucker & Ziv 1998:2). Hence, it is not surprising that the variety of items subsumed under the diverse technical terms perform various, more or less distinct, functions. Among the functions reported to be performed by discourse markers are, for example, discourse connection, signaling intimacy and topic switches, turn-taking, confirmation-seeking, hesitation marking, boundary marking, filling, prompting, repair marking, hedging, and focusing (cf. Jucker & Ziv 1998:1).

The distinction between 'discourse connectives' and 'discourse particles' is particularly relevant with respect to LIKE and thus requires elaboration. D'Arcy (2005, 2007) strictly differentiates the discourse marker from the discourse particle LIKE. The former term is equivalent to a discourse connective and, therefore, applies to instances of clause-initial and non-clausal LIKE, which serve to link, specify or exemplify previous utterances. LIKE as a discourse particle, on the other hand, encompasses instances of clause-medial and clause-final LIKE, which modify elements mostly on a phrasal level. In the present study, discourse marker serves as a cover term comprising both the discourse marker and discourse particle.

3.1.1 Features of discourse markers

Due to the fuzziness and the internal diversity of the concept 'discourse marker', it is not possible to supply a conjunctive definition that would not allow exceptions or grey areas (cf. Jucker & Ziv 1998:2; Sankoff et al. 1997:195). Hence, descriptive definitions are difficult to implement. To cope with the heterogeneous nature of the concept, a range of characteristic features of 'discourse markers' have been proposed in the literature. Brinton (1996) provides an extensive feature list, enabling the grading of discourse markers according their prototypicality: as not all features apply to each discourse marker, the more prototypical a discourse marker is, the more features it possesses. Hence, the concept 'discourse marker' exhibits scalable prototypicality with some items being more central and others peripheral, according to the number of properties they possess.

The following passage serves to introduce Brinton's (1996:33-38) list of the characteristics common to 'discourse markers':

1. Brinton (1996:33) notes that regarding phonological and lexical characteristics, discourse markers…
   (i) are ‚short' items (Östman 1982:149, 1995) and often phonologically reduced (Schiffrin 1987:328);
   (ii) form a separate tone group (Quirk et al. 1985:1112) with falling rising or rising intonation (Crystal & Davy 1975:90; cf. Schiffrin 1987:328);
   (iii) are marginal forms and, hence, difficult to place within a traditional word class.
2. In terms of syntax, discourse markers:
   (i) commonly occur in sentence-initial position (Keller 1979:222), although a number of specific markers frequently occur in clause-medial position. Items with clear grammatical functions, e.g. conjunctions, may be included if they serve pragmatic functions;

(ii) occur outside the syntactic structure or are only loosely attached to it. Hence, they have no clear grammatical function and are sometimes referred to as 'agrammatical' (Goldberg 1980:7);

(iii) are optional rather than obligatory.

3. Semantically, discourse markers:

(i) have little or no propositional meaning or are at least difficult to specify lexically (Schiffrin 1986:42, 47, 67; Östman 1982:153) and are, therefore, not easily translated (Svarvik 1979:169; Stubbs 1983:69). A number of linguists, however, oppose the notion that discourse markers are meaningless, i.e. completely devoid of semantic content (Andersen 1997, 2001; Erman 1987:15; Watts 1988:248, 250; Wierzbicka 1986a, 1986b).

4. Sociolinguistic and stylistic features of discourse markers are that they:

(i) are a feature of oral rather than written discourse and are associated with informality and the grammatical 'fragmentation' caused by the lack of planning time. Nevertheless, they are not restricted to oral discourse, although the forms used in writing may differ from those used in speech (Östman 1982:170; Fraser 1990:389; Redeker 1990:379);

(ii) appear with high frequency in oral discourse, sometimes more than one occurring in a single sentence;

(iii) are stylistically stigmatized and negatively evaluated, especially in written and formal discourse, as they are considered a sign of dysfluency and carelessness;

(iv) are assumed to be gender-specific and more typical of women's speech.

5. Discourse markers are multifunctional and operate simultaneously on several linguistic levels. This multifunctionalty is regarded as one of the defining features of discourse markers (Jucker & Ziv 1998) and fosters ambiguity since one and the same element cannot be definitely assigned one specific function. This has led researchers to assign pragmatic functions to discourse markers which are contradictory; e.g., simultaneously hedging and focusing (cf. Siegel 2002; Underhill 1988). According to Hölker (1991:78-79), these difficulties in assigning definitive functions to discourse markers are a result of their semiotic status as symptoms, i.e. as signs which communicate inwardness of speakers and subjective evaluations (Bühler 1934). Hölker (1991:78-79) emphasizes in this context that discourse particles do not relate to the situation talked about but to the speaker's stance towards the truth of an utterance. This characteristic of discourse markers connects them to epidemic modality.

On another semiotic level, discourse markers and particles serve a socio-symbolic function as indexes of a speaker's association with a certain social group, or of relations between interlocutors. For example, discourse markers correlate with informality and, hence, serve to express or locate speakers in the perceived socioeconomic hierarchy. A related yet different function served by discourse markers concerns the externalization of the speaker's self-image. Attitudinal studies (Buchstaller 2006b; Dailey-O'Cain 2000) show that LIKE is associated not only with certain parts of the speech community, e.g. young females, but also with certain personality traits. Indeed, LIKE users are perceived to be "more attractive, more cheerful, and more successful" (Dailey-O'Cain 2000:73). Following Eckert (2001), the discourse marker LIKE is a stylistic device functioning on a socio-symbolic level by furnishing the construction of social categories. In this interpretation, LIKE appeals to the interpersonal function of language by allowing speakers to proactively deploy

linguistic means to externalize their self-image, as with other symbolic resources such as clothing.

Although Brinton's (1996) list of discourse marker characteristics covers the basic features quite accurately, the following section provides a more detailed account of features discussed by Brinton (1996) discusses in more general terms.

Besides providing some phonological characteristics of discourse markers, Quirk et al. (1985:631)[9] offer predominantly syntactic criteria, such as the inability of discourse markers to be the focus of cleft sentences; to be the basis of contrast in alternative interrogation or negation; to be the basis of focused subjuncts; and to come within the scope of predication pro-forms or ellipsis. For example:[10]

(3) a. …*It is nonetheless that you should send her the agenda.
b. …*Should you send her the agenda nonetheless or therefore?
c. …*You should only <nonetheLÈSS> send her the agenda.

Generally, discourse markers occur predominantly in spoken discourse, are very frequent, are often stigmatized or denigrated, and are difficult to translate. Besides being marginal with respect to word class, discourse markers never carry participant roles such as Agent, Location, Goal, etc., (Miller 2009:319) and do not affect the truth conditions of the propositional content of an utterance (Brinton 1996:1; Jucker & Ziv 1998:3). Hence, discourse particles do not add to the semantic meaning[11] of an utterance and, thus, appear to be dispensable

9 Quirk et al. (1985: 631-634) call these items adverbials, or, more precisely, conjuncts, rather than discourse markers.

10 The following examples were taken from Quirk et al. (1985: 631).

11 Schiffrin (2001: 58) as well as Siegel (2002) reject the claim that discourse markers do not add to the propositional content of utterances. While Schiffrin (2001) holds that discourse markers do add to the meaning of discourse to varying degrees, "i.e. as when *oh* displays information as 'new' or 'unexpected'" (2001: 58), Siegel (2002) argues that the discourse marker *like* changes the truth conditions before numerals, in existential THERE-constructions and in sluicing sentences.

elements which may be removed without altering the truth conditions of the proposition. Therefore, lexical forms are discourse markers only if they are "grammatically optional, that is, they do not change the semantic relationships between elements" (Fuller 2003; cf. also Schiffrin 1987; Schourup 1999). Sankoff et al. (1997:1995) add that discourse markers are subject to semantic bleaching as compared to their source forms and undergo greater phonological reduction than their source forms.

In contrast to other clause elements, discourse markers have "a relatively detached and 'superordinate' role'" (Quirk et al. 1985:631). Hence, discourse markers fulfill pragmatic functions, indicating a speaker's attitude about an utterance or the speaker's attitude towards the truth or the relevance of the proposition conveyed in the utterance (cf. Andersen 1998; Blakemore 2004:238-239). In this respect, discourse markers serve to inform the hearer's interpretation of what is being said as well as helping the hearer to evaluate the speaker's stance on the proposition made. Linguistic "[i]tems which have these properties are single words such as conjunctions (*but*), adverbs (*now*, *then*), interjections (*oh*), or lexicalized phrases (*y'know*, *I mean*, *after all*)." (Miller 2009:320).

Sankoff et al. suggest that discourse markers are "lexical items that relate to discourse rather than to syntax or semantics" (Sankoff et al. 1997:195). In accordance with this view, Schiffrin (2001:55-56), for example, elaborates on the formal criteria – such as syntactical detachability and its position within clausal structures – depicting discourse markers as non-obligatory, (predominantly) utterance-initial items that function in relation to ongoing talk and text (Schiffrin 2001:57).

Nevertheless, the classification of discourse markers remains difficult for two reasons: firstly, the demarcation between homophonous and historically related forms is by no means clear – in fact, their classification is a matter of degree, resembling a continuum rather than a dichotomy. Therefore, morphologically identical elements frequently belong to various parts of

speech, making definite assignment of function difficult. This is especially true because of the positional flexibility of discourse particles. Secondly, discourse particles and their more functional, lexical or grammatical counterparts, share not only morphological similarity, but also linguistic ancestry. This means that the same form co-occurs in different functions in real-time, representing different stages of grammticalization or 'pragmaticalization' (cf. Diewald 2010, 2011; Günthner & Mutz 2004).

3.2 Synopsis

This chapter has introduced and discussed what is generally and in this study understood by the term discourse marker. In addition, this chapter has provided a list of features typically associated with discourse markers and has shown that LIKE is a member of this group because it shares almost all the features of other discourse markers. Furthermore, this chapter served to discuss difficulties arising when classifying elements as discourse markers which result from the fuzziness of this concept.

4 Overview of previous research on LIKE

The following chapter introduces previous research on LIKE. The first part of the chapter is concerned with the diachronic development of the discourse marker LIKE. In this context the notion of pragmaticalization is defined and discussed with respect to its applicability to the present case. More specifically, the following chapter will present and discuss the most prominent theories of LIKE's grammaticalization and elaborate on the grammaticalization pathways of LIKE proposed by these theories. In the following chapter, the focus is placed on the the re-analysis of clause-medial LIKE and an alternative grammaticalization pathway for clause-final LIKE. The second part recapitulates LIKE's diachronic development and discusses whether it has grammaticalized in parallel in various locations simultaneously, or whether, when, and how it has spread from one or more source varieties to other target varieties. Subsequently, this chapter will present previous research on the attitudes associated with LIKE and on its syntactic positioning. Finally, the chapter will discuss in detail the discourse-pragmatic functions of LIKE and their relationship to its positioning within a clause. In addition, the final part of this chapter will highlight related uses of *like* and discuss why they cannot be regarded as realizations of LIKE, such as uses of *like* as a part of general extenders.

4.1 The history of LIKE

4.1.1 Grammaticalization

The first question that comes to mind is how and from which forms vernacular uses of LIKE developed. One possible explanation is related to the processes of grammaticalization, i.e. "[t]he process whereby lexical items take on certain grammatical functions in certain linguistic environments […]. Grammaticalization also refers to the process through which an already grammatical item takes over new or additional grammatical functions" (Golato

2000:33). In other words, "[g]rammaticalization is a process of gradual change, and its products may have different degrees of grammaticality" (Lehmann 1982:13). Hence, the process is unidirectional proceeding from lexical to grammatical. Nonetheless, there are a significant number of cases contradicting the unidirectionality hypothesis, but the number of exceptions is substantially lower than that which conforms to the common pattern (Heine 2002:97; Newmeyer 1998).

Although this concept is extremely powerful with respect to a number of systematic changes we observe when analyzing the history of languages, it is questionable whether this concept is adequate to account for the development not of grammatical features, but pragmatically functional forms (cf. Auer & Günthner 2005; Wischer & Diewald 2002). Indeed, the validity of theories of grammaticalization with respect to pragmatic markers has recently led to several publications addressing this very issue, i.e. differences and similarities between grammaticalization and pragmaticalization (cf. Brinton 1996; Günthner and Mutz 2004; Hopper & Traugott 1993; Traugott 1995, 1999; Traugott & Heine 1991). These studies focus specifically on whether the diachronic development of pragmatic markers is best ascribed to pathways similar to the traditional notion of grammaticalization, or whether their development constitutes a somewhat different process, i.e. pragmaticalization (Diewald 2011:374).

A first step towards accounting for pragmatic functions within a framework of grammaticalization is to expand the definition to encompass "the degree of grammatical function a linguistic item has on a scale between purely lexical and purely grammatical meaning. This entails both a diachronic and a synchronic perspective" (Wischer & Diewald 2002:ix). Taking Wischer and Diewald's position, grammaticalization is not limited to processes in diachrony as expressed in Givón's statement that (1971:413) "[t]oday's morphology is yesterday's syntax", but it also refers to a continuum and relation between lexical and grammatical function. Hence, we have two distinct and in part contradictory concepts of grammaticalization (Günthner & Mutz 2004:97-98).

> This type of change which leads to discourse and pragmatic markers, to elements which organize, structure, and contextualize discourse with respect to discourse-pragmatic concerns and not with respect to sentence-grammatical concerns (e.g., congruence, binding), contradicts classical grammaticalization. (Günthner and Mutz 2004:98)

This type of change, which leads to the development of discourse-pragmatic elements and not to grammatical elements, "contradicts classical grammaticalization" (Lehmann 1982, 1991; Günthner & Mutz 2004:97). To allow differentiation between these two more or less separate concepts, alternative terminologies for the latter type of change have been introduced, e.g. degrammaticalization, lexicalization, and pragmaticalization (Günthner & Mutz 2004:97). In contrast to Günthner & Mutz (2004), however, I will not differentiate between these two concepts, but treat pragmaticalization as a subtype of grammaticalization which thus serves as an umbrella concept. Indeed, Diewald (2011:384) asserts that

> pragmaticalization is a specific instance of grammaticalization which shows the crucial features of grammaticalization processes and is only distinguished from other grammaticalization processes by the functional domain it leads to and by some concomitant structural features (e.g., low degree of syntactic integration). As pragmaticalization (understood in these terms) is one instance of grammaticalization among many others, there is no reason to treat it on par with grammaticalization, i.e., on the same hierarchical level, in a classification of types of language change" (2011:384).

Concerning the grammaticalization of discourse markers, Traugott (1995; 1999) proposed that the pathways involved are not arbitrary, but share several distinctive features. For example, Traugott (1999) observes that the vast majority of discourse markers develop out of adverbs and conjunctions, thereby going through distinct phases. In a first step, the elements undergoing this process become increasingly decategorized into less central grammatical categories, i.e. discourse or cohesion markers (Günthner and Mutz 2004:84). The next stage is marked by an increase in subjectification and pragmatic

function: "meaning tends to become increasingly situated in the speaker's subjective belief-state/attitude toward the situation (Traugott & König 1991:205). The direction of change follows a trajectory from textual or referential to an interpersonal or metatextual meaning.

Another aspect characteristic of such processes is a broadening of scope from sentential to extra-sentential. On the one hand, this can encompasses an increase in scope from connecting clauses to indicating their relation with each other. On the other, it refers to an increase in scope from a single inner-clausal constituent such as a single phrase to larger constructions and subsequently clauses and sentences. Finally, while pragmatic or discourse markers profit from semantic bleaching as the resulting vagueness allows these elements to acquire new functions and fill formerly restricted positions, the initial meaning still resonates in these new uses. In the case of LIKE, this semantic bleaching is reminiscent of the similarity relation expressed by the comparative preposition which is still present – though to a lesser degree – in its so-called core meaning. According to Schourup (1982, 1985), the core meaning of vernacular uses of LIKE is to express a "possible unspecified minor nonequivalence of what is said and what is meant" (Schourup 1982:31). Thus, LIKE informs the listener about a similarity and non-identity between the actual and the intended utterance.

This schematic explanation does not provide an explanation of why discourse markers develop in the first place. A plausible hypothesis is given by Heine (2002:85-86), who redefines processes of grammaticalization in terms of semantic change (Heine 2002:84). To account for semantic changes within a theory of grammaticalization, Heine (2002) introduces a four-stage model of grammaticalization which describes the process by which a form-meaning pair acquires an additional new meaning. More specifically, Heine's (2002) model depicts the conditions for re-analysis by which the meaning and function of *like* changed into the present-day form LIKE. The basic idea is that the initial form or construction is used in a specific context, i.e. in bridging contexts, which allows for or "gives rise to an inference in favor of a new meaning" (Heine 2002.

86). This stage is crucial, as the new context triggers a reanalysis of the element in question. It is this re-analysis of the traditional form in this specific context which gives rise to the new semantic meanings or, as will be argued subsequently, pragmatic function. In a third stage, the form is used in a context which is compatible with the newly acquired, but not with the traditional meaning. The last stage of the process consists of conventionalization, i.e. the new "meaning no longer needs to be supported by the context that gave rise to it; it may be used in new contexts" (Heine 2002:86). Heine's model differs from a related precursor introduced by Diewald (1999), concentrating particularly on semantic change (Diewald 2002:117).

4.1.2 The grammaticalization of LIKE

Throughout the past two decades, the grammaticalization of LIKE has attracted the attention of researchers, leading to various publications on this topic. Most scholars agree that the discourse marker LIKE originated from the comparative preposition which in the course of time lost lexical meaning. The resulting lexically bleached variants exhibit an increase in syntactical flexibility; as Andersen puts it: "It is obvious that the word *like* is undergoing a process of grammaticalization from a lexical item to a particle which has greater syntactic freedom and is to some degree semantically reduced" (1997:40). The grammaticalization of LIKE is, nevertheless, presumed to be ongoing, as LIKE continues to intrude into formerly constrained syntactical environments (D'Arcy 2005). The assertion that this grammaticalization process is ongoing is supported by the multifunctionality of LIKE, which is commonly regarded as indicative of ongoing grammaticalization (cf. Romaine & Lange 1991; Buchstaller 2001).

One of the first studies on the grammaticalization of LIKE was published by Meehan (1991), who provided a historical account not only of the development of the discourse marker LIKE, but of the historical development of the general linguistic form *like*. Meehan's (1991) model of LIKE grammaticalization is

strictly linear and unidirectional in nature despite allowing for various forms of *like* occurring in synchrony.

Although not strictly limited to the discourse marker LIKE, Meehan's (1991) approach resulted in the introduction of the notion of a 'core-meaning', often paraphrased as 'similar to' meaning which is more or less salient in most forms of like including occurrences of LIKE as an adjective, adverb, conjunction, discourse marker, quotative, and noun. This core meaning dates back as early as the 14th century and functions to express a similarity relation by facilitating comparison. A closely related interpretation of LIKE's core meaning expresses approximation, which can be paraphrased as meaning 'similar to' and commonly occurs in the context of quantity phrases.

The more specific type, which developed from the core meaning of LIKE and which represents a later stage of the grammaticalization process, is its use as a conjunction. This form of *like* can be paraphrased as meaning 'as if'. It differs from the approximating function in that it has broader scope, and can take entire clauses, while comparative and approximating like predominately introduce post-verbal noun phrases (Meehan 1991:40).

A more recent type of *like*, which originated in the 19th century and which can be paraphrased with 'for example' and replaced by 'such as', is characterized by a further broadening of scope, as it can take either NPs, PPs, or entire clause as its complements "indicating that the meaning is becoming more generalized" (Meehan 1991:43). In addition, the information within the scope of this form tends to be new and focus the listener's attention. In contrast to discourse marker LIKE, this form cannot be removed without making the sentences ungrammatical, and thus, still has enough lexical content not be considered "simply a marker of new information" (Meehan 1991:43).

While Meehan (1991) proposed that LIKE originated from the comparative preposition and follows a sequential channel, Romaine & Lange (1991) proposed a somewhat similar grammaticalization path for LIKE based on Traugott's (1982) model. Although their model focuses on the syntactical

development, Romaine & Lange also include basic pragmatic and semantic aspects, such as a basic core meaning of LIKE equivalent to expressing approximation, comparison, or similarity. Romaine & Lange's (1991) consider LIKE to have developed diachronically from a comparative preposition, to a conjunction, to a discourse marker, and to a quotative complement. In contrast to Meehan's model, however, Romaine & Lange's recategorization model is not strictly sequential, but rather branching, as a "linear model of grammaticalization is inadequate to account for these developments [the multifunctionality of LIKE in synchrony]" (Romaine & Lange 1991:262). According to Romaine and Lange (1991), LIKE derives from the comparative preposition which takes a nominal complement. During an intermediate step, the scope of *like* widens from nominal to a sentential complement when *like* begins functioning as a conjunction. Only then is *like* reanalyzed as a "discourse marker with syntactic detachability and mobility" (Brinton & Traugott 2005:23).

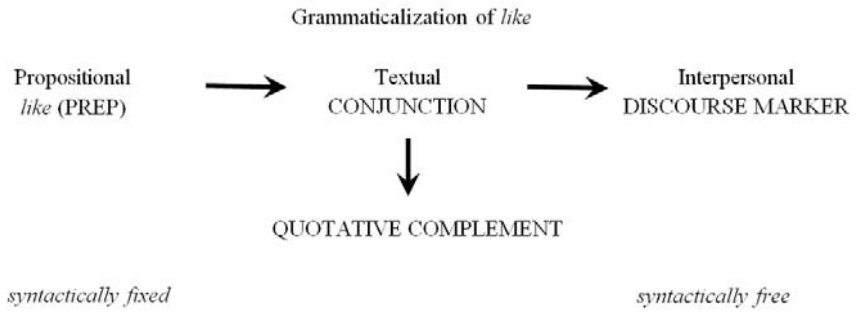

Figure 6: Grammaticalization path for like *(Romaine & Lange 1991:261)*

A more recent model introduced by Buchstaller (2001) and elaborated on in Fleischman and Yaguello (2004) is based on Lakoff's (1987) radial structure model. The need for modification derives from the fact that the grammaticalization channel as proposed by Traugott and Heine (1991) fails to show how "the synchronically co-existing meanings overlap and reinforce each other" (Buchstaller 2001:31). Buchstaller's (2001) model links interrelated semantic-pragmatic pathways, hence, enabling a more systematic account of

LIKE's multifunctionality which is portrayed as a network rather than a sequential or branching channel:

> The links between the synchronically co-occurring and often overlapping uses of *like* are metaphorical and metonymical extensions from one common comparative/approximative core and conversationalizations of conversational implicatures. They extended to a network of relations. The diverse functions that *like* has assumed synchronically are motivated by this model – they cannot be predicted but they are explained. (Buchstaller 2001:33)

Similar to those previously depicted, Buchstaller's model also includes and heavily depends on the notion of an underlying common core meaning which "is very closely related to its approximative semantics" (Buchstaller 2001:32) and is assumed to be comparative/approximative in nature. On theoretical grounds based on Haiman (1989), Buchstaller (2001:32) argues that, as comparisons highlight those elements which are compared, the link between focusing and comparison is salient and thereby links the comparative core meaning of LIKE to its focusing function.

This does not, however, answer the question of how focusing can give rise to either LIKE's occurrence as an epistemic hedge or to pragmatic hedging. To address this issue, Buchstaller (2001) argues that by signaling looseness of fit between two compared items, on a propositional level LIKE can be interpreted as an epistemic hedge. In support for her argument, Buchstaller notes that "the path from comparison to a hedge of epistemic uncertainty seems to be a well-trodden one cross-linguistically" (Buchstaller 2001:32). LIKE's pragmatic function is derived from the epistemic hedge as a face-saving device, which can be reinterpreted as a pragmatic hedge when it is "transferred to an interpersonal, affective level" (Buchstaller 2001:32).

LIKE is used as a filler most often when formulating problems arise (Buchstaller 2001:32; Siegel 2002) and can thus be accounted for in a psycholinguistic framework. Its filling function may have originated as an

extension of its pragmatic hedging function (Buchstaller 2001:32). Buchstaller's fundamental argument is that while the semantics of LIKE – the comparative 'similar to' sense – make it an ideal filling item, it also "works on the interpersonal level, as a floor-holding device, and as a signal of production problems" (Buchstaller 2001:32).

The link between the focusing and the quotative function of LIKE, on the other hand, is "explained […] by the fact that quotations are very often the most focused part of an utterance as they display immediacy and interpersonal involvement" (Buchstaller 2001:32). In addition, quotative LIKE commonly does not introduce actual utterances, but approximations, i.e. utterances which are not exact replications of utterances, but serve to illustrate what could have been uttered in the respective situation. The comparative core of the quotative LIKE allows for its classification "as a hedge both on the referential and on the interpersonal level, as the speaker retains a reduced responsibility with respect to what was said and how" (Buchstaller 2001:33).

Buchstaller's model explains LIKE's factual multifunctionality in synchrony as well as its diachronic development in terms of metonymical and metaphorical extensions of its comparative/approximating core meaning without limiting itself to a "clear-cut linear grammaticalisation path, such as the ones postulated by Traugott and Heine (1991)" (Buchstaller 2001:33). Further evidence for her proposed model is drawn from cross-linguistic data which indicate a general link between the core meaning of comparison and its derived functions such as focusing or hedging, especially as items from similar semantic source-domains follow parallel semantic-pragmatic paths of development in unrelated languages such as English and Thai (Buchstaller 2001:33).

The latest model accounting for the development of LIKE has been proposed by D'Arcy (2005:5), who offers a more refined version of the unidirectional model introduced by Romaine and Lange (1991) and

theoretically grounded in the framework offered by Traugott (1995) and Brinton (1996).

1. preposition/conjunction > sentence adverb > discourse marker
2. discourse marker > discourse particle > ?

Figure 7: Grammaticalization path of like as proposed by D'Arcy (2005:218-219)

D'Arcy (2005:51-73) certainly provides a compelling analysis of the discourse-pragmatic processes accompanying the grammaticalization of LIKE, such as subjectification and a shift from textual to interpersonal meaning; hence, D'Arcy (2005) – similar to Traugott & König (1991) and Traugott (1995) – consoders grammaticalization to be a process of unidirectional pragmatic strengthening. According to this view, the cline or grammaticalization channel of LIKE proceeds from "non-pragmatic (propositional) to pragmatic (i.e. textual/metalinguistic, subjective/interpersonal) meaning" (Brinton 2000:154).

D'Arcy (2005) reasonably assumes that the grammaticalization pathway of LIKE starts at the comparative preposition *like* or the conjunction *like*, which is best glossed as '*as if*'. Subsequently, the clause-marginal sentence adverb emerged which then developed into the discourse marker LIKE. Clause-medial LIKE is, according to D'Arcy (2005:218), a later development. While Traugott (1995) and Brinton (1996) offer analyses for the grammaticalization of discourse markers in general, their models are problematic when they applied to LIKE. For example, in applying Traugott's (1995) model, D'Arcy (2005) makes a very compelling argument for a grammaticalization path of LIKE, beginning with the comparative preposition, to sentence adverb, to discourse marker, to discourse particle. Nonetheless, D'Arcy (2005:69) admits that she has "no straightforward explanation for this positional shift", suggesting that equivalent trajectories are attested to for similar forms such as conjunctions (sentence final *but* and sentence final *though*) and discourse markers (*indeed*, *besides*) (cf. Traugott 1995).

With respect to semantics, several authors propose a direct link between the similarity relation expressed by the comparative preposition and the hedging function of clause-medial LIKE (Buchstaller 2001:23-24; Romaine & Lange 1991:260-261). Finally, the change from comparative forms, which establish similarity relations, to hedging devices is by no means unexpected – in fact, similar developments are attested to in several languages (Fleischman & Yaguello 2004; Meyerhoff & Niedzielski 1998; Sankoff et al. 1997).

4.1.3 Interim synopsis

This part of the chapter has presented and briefly discussed the concepts of grammaticalization and defined pragmaticalization as one of its subordinate processes. In addition, it has provided the most prominent theories of LIKE's grammaticalization. The most common theory with respect to the grammaticalization of LIKE proposes that LIKE developed from the comparative preposition (cf. D'Arcy 2005; Meehan 1991; Romaine & Lange 1991).

The following chapter will briefly recapitulate LIKE's diachronic development and discuss the issue of whether it has grammaticalized in parallel in various locations simultaneously or whether, when, and how it has spread from one or more source varieties to other target varieties. Subsequently, this chapter will present previous research on the attitudes associated with LIKE and previous claims about its syntactic positioning.

4.2 The development of LIKE

The previous account of LIKE's grammaticalization raises the question as to exactly when LIKE developed and how its development is linked to the sociolinguistic perspective adopted here.

At first sight, it may appear that the discourse marker LIKE is a single, homogenous form occurring in all possible utterance positions and syntactic

environments. A more fine-grained analysis of this form reveals, however, that LIKE is multifaceted. Until the latter half of the twentieth century, LIKE was commonly regarded as non-standard, colloquial, or even vulgar (Jespersen 1954:417) and assumed to be either a meaningless interjection akin to elements such as *uhm*, and *mh*, or an "expletive to provide emphasis or pause" (Morris 1969:757; Romaine & Lange 1991:245). Early accounts of LIKE in slang and dialect dictionaries depict it as a traditional, dialectal or non-standard feature of English (e.g. Wright 1857; Wright 1902; Grant & Dixon 1921; Partridge 1984). These early accounts focused mainly on clause-final LIKE as in (4), a variant which is commonly associated with Northern British varieties of English (Hedevind 1967:237) and which has more recently been described as "archaic or traditional (and obsolescing)" (D'Arcy 2005:5).

(4) a. Well I mean it's up to yourself <,> it's up to you *like*. (ICE-Ireland:S1B-016$A)
b. I feel we sh'd both be better for a change *like* (Jespersen 1954:418)
c. And she was good in the Who Dunnit *like*. (ICE-Ireland:S1A-018$A)

According to several authors (e.g. Andersen 1998, 2000, 2001; D'Arcy 2005; Siegel 2002; Tagliamonte 2005), a supposedly more recent and innovative American variant has quickly been gaining ground, especially in the language of teenagers. With respect to the origin of this innovative variant of LIKE, Andersen (2000) states that it is "said to have its roots in New York City counterculture groups (jazz, cool and beat) in the 1960s" (Andersen 2000:216, cf. also Wentworth & Flexner 1967; Chapman 1986; OED 1989 VIII:946). Although Croucher claims that LIKE, along with *you know*, "were implanted in the American culture by a popular creation in the 1980s, 'Valley Girls'" (2004b:38), this is probably incorrect as there are various attestations of vernacular uses of LIKE predating the mid- to late twentieth century. In fact, both Croucher's (2004a) and Andersen's (2000) hypotheses are qualified by D'Arcy (2005, 2007). Based on her analysis of LIKE use in Toronto English, she concludes that "the vernacular forms are not twentieth-century innovations that originate from the Valley Girls. Only the quotative may be sourced to this

group; the rest have extended histories in the English language" (D'Arcy 2007:411). Consider D'Arcy (2007:411) on this issue:

> The combination of empirical data from regional dialects of British English and the apparent-time results from Toronto suggest that the nonquotative vernacular functions of *like* have been increasing in frequency over the last 65 years or so, and the marker for seemingly longer still. In other words, they represent change in progress and cannot be isolated to the North American context.

In addition to D'Arcy (2005, 2007), Meehan (1991) and Schweinberger (2013) emphasize that vernacular uses of LIKE have a long-standing history dating back to the late eighteenth and early nineteenth centuries. The OED also provides various examples from this era as, for instance, in (5). The OED describes such instances of LIKE as dialectal and vulgar used "parenthetically to qualify a preceding statement: = 'as it were', 'so to speak'. Also, colloq. (orig. US), as a meaningless interjection or expletive." (OED, 2nd ed., 1989; online version March 2011).

(5) a. Father grew quite uneasy, *like*, for fear of his Lordship's taking offence. (1778:F. Burney Evelina II. xxiii. 222)

b. In an ordinary way *like*. (1826:J. Wilson Noctes Ambrosianae xxvii, in Blackwoods Edinb. Mag. July 91)

c. If your Honour were more amongst us, there might be more discipline *like*. (1838:E. Bulwer-Lytton Alice I. ii. iii. 157)

d. Why *like*, it's gaily nigh *like*, to four mile *like*. (1840:T. De Quincey Style in Blackwood's Edinb. Mag. Sept. 398/1)

The examples dating back to the eighteenth and nineteenth centuries are almost exclusively clause-final uses of LIKE. The earliest instances of clause-initial LIKE provided in the OED date back to the middle of the twentieth century as in (6). Clause-medial LIKE as in (7) seems to be an even later development, as the first attestations provided by the OED appear in data from the early 1970s.

(6) a. *Like* how much can you lay on [i.e. give] me? (1950:Neurotica Autumn 45)

b. What will be the contradictions that produce further change? *Like*, it seems to me that it would be virtually impossible to avoid some contradictions. (1973:Black Panther 17 Nov. 9/4)

(7) To concoct some fiendish scheme that might *like* give youse a fightin' chance. (1971:'H. Calvin' Poison Chasers xiii. 170)

However, while D'Arcy (2005) and Schweinberger (2012) concur that clause-final LIKE grammaticalized much earlier than clause-initial and clause-medial LIKE, the timeline suggested by the OED examples for clause-initial and clause-medial LIKE requires revisiting. To elaborate: Romaine and Lange (1991:270) hypothesize that vernacular uses of LIKE have probably been around for more than a century (cf. also D'Arcy 2007:401). Schweinberger (2013) offers further evidence for D'Arcy's (2007) interpretation and provides examples of clause-medial LIKE occurring in the speech of Northern Irish English speakers aged 65 and older as in (8).

(8) a. There not so much work, *like*, on farms around here now. (NITCS:L17.3# I LD37; Protestant female aged 65 to 75)

b. Well, I don't really know much, *like*, about making cheese. (NITCS:L12.3# I LM25; Catholic female aged 65 to 75)

c. It was off, *like*, the road this number of years, it is. (NITCS:L7.3# I WG23; Protestant male aged 65 to 75)

d. Oh, aye, aye. They were very awkward things to shoe. You know, you could hardly get the feet off the ground, they were, *like*, stupid, you know. (NITCS:L7.3# I TF80; Catholic male aged 65 to 75)

This finding is corroborated by apparent-time evidence provided by D'Arcy (2007:400), who concludes that

> [t]his suggests that the marker was already a feature of the vernacular before it was associated with the Beat and jazz groups of the 1950s and 1960s. In fact, working from the apparent-time hypothesis, in the 1930s, when these 80-year-olds were teenagers, *like* must have been relatively frequent in the ambient language as a discourse marker, a usage inherited by these speakers from the previous generations. The added perspective afforded by the British data further jeopardizes the plausibility of the counterculture genesis hypothesis. As both a marker and a particle, *like* is attested among the oldest

> speakers in the English, Scottish, and Northern Irish communities considered here, raising troubling questions about the American roots of these forms more generally.

Nonetheless, extrapolating from apparent-time data to real-time is not unproblematic (cf. section 2.4.1); hence, D'Arcy's (2007:400-401) assertions require complementary real-time evidence to be conclusive. Indeed, early instances of thediscourse marker LIKE can be observed in historical material included in the *Corpus of Irish English* (Hickey 2003). Consider (9):

(9) a. MIRABELL: By your leave, Witwoud, that were *like* enquiring after an old Fashion, to ask a Husband for his Wife. (Congreve, William 1700:The Way of the World; *Corpus of Irish English*)

b. 'Judy's out a luck,' said I, striving to laugh 'I'm out a luck,' said he, and I never saw a man look so cast down; he took up the halfpenny off the flag, and walked away quite sobered *like* by the shock. (Edgeworth, Maria 1801:Castle Rackrent, an Hibernian Tale; *Corpus of Irish English*)

c. She was dressed like a mad woman, moreover, more than *like* any one I ever saw afore or since, and I could not lake my eyes off her. (Edgeworth, Maria 1801:Castle Rackrent, an Hibernian Tale; *Corpus of Irish English*)

The instances in (9) seriously challenge the assumption that clause-medial LIKE is a twentieth-century development and substantiate D'Arcy's (2007) tentative assertions that LIKE neither originated in North American English, nor did it grammaticalize in the twentieth century.

4.3 Spread or parallel development?

The fact that vernacular uses of LIKE have a long history dating back about two centuries, poses a serious question: why did it take so long for discourse LIKE to start spreading and become the salient feature of contemporary vernacular English? In other words, why did LIKE remain in stasis for most of its lifetime, until the latter half of the twentieth century? The short answer is that – despite not having originated in AmE – it is quite plausible that the association of LIKE

with US American pop culture[12] triggered the spread of this non-standard form (Labov 2001:307). According to this view, the use of LIKE as a marker of identity or as a form expressing covert prestige within certain speech communities in the United States may have caused the dramatic rise in frequency. If we follow this account, the increase in the use of clause-medial LIKE in AmE may have reinforced its spread to or increase in varieties of English as an effect of globalization, specifically Americanization.

Meyerhoff and Niedzielski (2003) draw attention to a non-trivial issue concerning the analysis of global spread: before prematurely assuming that a feature is spreading out of an epicenter where the feature in question developed, parallel developments must be ruled out (2003:546). Indeed, Levey (2006) asserts that the development of elements functioning similar to English LIKE have been attested to in a variety of typologically distinct languages, leading him to hypothesize that the emergence of discourse markers has its origin "in underlying principles of conceptualising particular semantic domains which motivate trajectories of grammaticalisation" (Levey 2006:417). Moreover, Levey (2006:417) states that

> [t]here is accumulating evidence (cf. Andersen 1997; Meyerhoff and Niedzielski 1998; Fleischman 1999; Maschler 2001) of analogous discourse uses of *like*-lexemes in a range of languages. These cross-linguistic analogues exhibit striking parallels: they can be used in a variety of syntactic positions to sequence chunks of discourse, to hedge propositions, mark focus, and foreground reported speech/thought.

Indeed, the list of languages in which discourse markers have emerged possessing both functional as well as developmental analogies with LIKE is remarkable (Levey 2006:417-418): Canadian French *comme* (Sankoff et al. 1997:205); Hebrew *kaze* (Maschler 2001:296); Bislama *olsem* (Meyerhoff &

[12] In this context, the question of whether the reference category was New York counterculture groups or the Californian 'Valley Girls' is secondary and, hence, negligible.

Niedzieslki 1998:239); Finnish *niinku* (Fleischmann 1999, Heine & Kuteva 2002:274); Swedish *liksom* (Kotsinas 1994:87); and Italian *tipo* (che) (De Mauro et al. 1993, Fleischman 1999). According to Levey (2006:418), Fleischman (1999) even extends this list to encompass German, Japanese, Lahu, Portuguese and Russian.

Given that discourse markers such as LIKE appear to follow a consistent trajectory, Levey (2006) concludes that "these discourse markers cannot be easily ascribed to borrowing processes or common genetic inheritance" (Levey 2006:418). This commonality is suggestive and it is tempting to believe that the omnipresence of LIKE in geographically distinct varieties of English reflects parallel developments rather than diffusion from a common epicenter. In particular, Meyerhoff and Niedzielski (2003) draw attention to the fact that, when analyzing features in geographically discontinuous locales, scholars should confirm that the same linguistic constraints operate on identical forms (2003:544). According to this view, a similarity of constraints legitimizes the hypothesis that a form has spread from variety A to variety B, while diverging constraints signal parallel development (2003:545-546).

While establishing or reaffirming norms for good practice is undoubtedly a *sine qua non*, the present study argues against the primacy of language-internal constraints and stresses the importance of - language-external social factors. Hence, the justification for assuming the spread instead of parallel development relies not on the similarity of linguistic constraints but on the historical timeline. Indeed, the hypothesis that LIKE developed independently is inadequate to explain the discrepancy between the origin of this feature about two centuries ago and its recent global increase, i.e. over the past fifty or so years. Hence, it is implausible that LIKE use remained locally confined to EngE and IrE at a constant level for about two centuries until it developed independently in various regionally distinct varieties within a period of thirty or so years. The improbability of this scenario as well as the abundant exposure to AmE vernacular in the media, leads me to conclude that the global

availability of LIKE is a result of spread, but not of independent, parallel development, though parallel development cannot be ruled out entirely.

The nature of this spread is complex in that the most probable scenario consists of three distinct stages. During its first stage, clause-final LIKE spread from southern parts of England outward to Ireland and Scotland. This interregional diffusion probably took place during the late eigthteenth and ninetheenth centuries. This hypothesis is supported by two observations. First, clause-final LIKE is attested to in rural dialects across Britain – according to the global synopsis of Kortmann & Szmrecsanyi (2004:1162-1163) focussing LIKE one of the most widespread non-standard features across the British Isles and it is attested in seven out of eight British varieties surveyed in Kortmann et al. (2004) – which indicates that clause-final LIKE had diffused into almost all regions of mainland Britain. Secondly, the elevated frequencies of clause-final LIKE in Ireland (cf. section 7.4.1.3) and Scotland (cf. e.g. Romaine & Lange 1991) compared to its frequencies in southern parts of Britain (cf. section 7.3) suggest that clause-final LIKE represents a calque from earlier forms of English which have strongly been influenced and have borrowed thoroughly from earlier forms of standard English. The second wave of outward spread took place during the late ninetheenth and early twentieth centuries. During this phase, LIKE was introduced to colonial settlements in Australia and New Zealand as well as Northern American varieties of English. In addition, LIKE continued to affect the dialect formation of IrE and was adopted due to its functionality as a pragmatic element of the target variety. This scenario is supported by the use of clause-final LIKE among older speakers in Toronto and historical data of NZE (cf. D'Arcy 2007) and attestations of clause-final LIKE in the speech of 65- to 75-year old speakers in IrE data compiled during the early 1970s (Schweinberger 2013). The third and final phase of spread took place during the latter half of the twentieth century and had as its source not Britain, but the United States. The empirical support for this third phase lies in the omnipresence of LIKE in inner and outer circle varieties – according to the *Electronic World Altas of Varieties of English* it is used in at least 48 of the 74

varieties included in that data (Kortmann & Lunkenheimer 2011). The stark contrast between the fact that LIKE was until 50 years ago attested predominantly in the British Isles and some colonial varieties, on the one hand, and the current situation, on the other hand, where it is common in most varieties of English around the world strongly suggests that there must have been a third phase of spread. In the global synopsis (chapter 8), it will be argued that this third wave of spead relied, on the one hand, on the culturally dominant role and prestige of AmE and, on the other, it relied on dialect contact in the wake of WWII and on the omnipresence of audio-visual mass media, i.e. television and radio, as a means of diffusion.

4.4 LIKE across varieties of English

After taking a closer look at how LIKE developed, we are now in a position to survey its use across varieties of English today. This section will start by providing a general overview of where LIKE is attested based on Kortmann et al.'s (2004) survey and then proceed by focusing more specifically on distinct regions and summarize what individual studies have come to conclude about LIKE use in these regions.

Although, systematic surveys of the occurrence and frequency of the discourse marker LIKE are presently still wanting, LIKE as a focusing device is widely attested across the varieties of English surveyed in Kortmann et al. (2004).

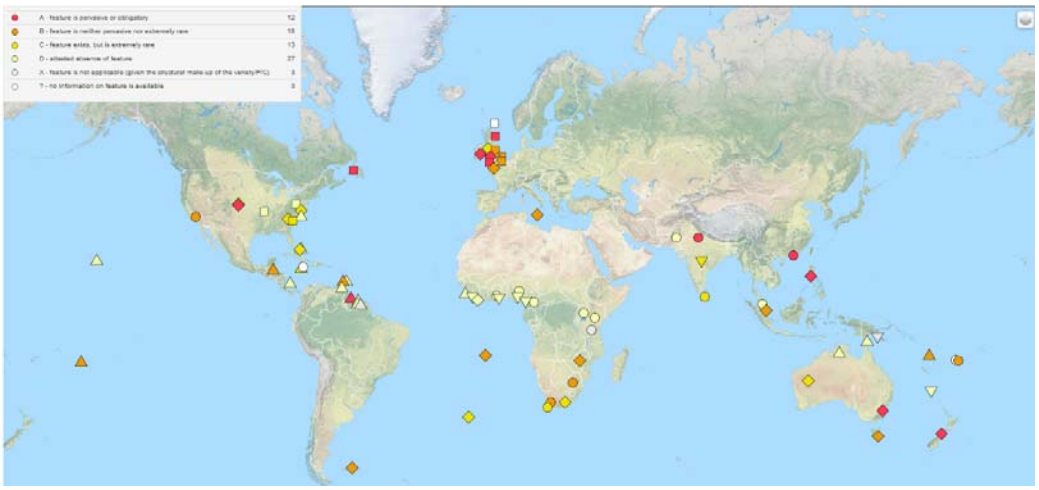

Figure 8: Like as a focussing device across varieties of English (Adopted from Kortmann & Lunkenheimer 2011)

In both Kortmann et al. (2004) and Kortmann and Lunkenheimer (2011), LIKE is unanimously referred to as focusing LIKE. I will presume here that this label encompasses all uses of LIKE, whether they are deployed as focus markers or not. According to Kortmann et al. (2004), LIKE is a widespread feature. In fact, it is listed as a pervasive or obligatory feature in eleven varieties, as a feature that is neither pervasive nor obligatory in 19 varieties and a feature that exists albeit extremely rarely in twelve of the 74 varieties of English included in Kortmann et al. (2004) (cf. Figure 8). However, its absence is attested in 26 varieties included in the survey.

LIKE appears to be particularly frequent in traditional L1 and high-contact L1 varieties: LIKE is attested in seven out of ten traditional L1 varieties – in four of which it is even listed as pervasive or obligatory – and in twenty out of twenty-one high-contact L1 varieties (cf. Kortmann et al. 2004, Kortmann & Lunkenheimer 2011). This means that it is attested in 87 percent of the L1 varieties surveyed in Kortmann et al. (2004). However, its absence is typically attested for English-based Creoles and Pidgins (cf. Figure 8 & Table 2).

Table 2: LIKE across varieties of English (Adopted from Kortmann & Lunkenheimer 2011)

	Traditional L1 varieties	**High-contact L1 varieties**	**Indiginized L2 varieties**	**English-based Creoles**	**Pidgins**
Pervasive or obligatory	4 (40.0%)	4 (19.0%)	2 (11.8%)	1 (5.3%)	---
Neither pervasive nor extremely rare	2 (20.0%)	8 (38.1%)	4 (23.5%)	4 (21.1%)	1 (14.3%)
exists, but is extremely rare	1 (10.0%)	8 (38.1%)	1 (5.9%)	1 (5.3%)	1 (14.3%)
Attested absence	2 (20.0%)	1 (4.8%)	7 (41.2%)	12 (63.2%)	4 (57.1%)
Not applicable	---	---	1 (5.9%)	1 (5.3%)	1 (14.3%)
Don't know	1 (10.0%)	---	2 (11.8%)	---	---
Total	10	21	17	19	7

What is striking about Table 2 is the decrease in use and attestations from left to right, i.e. from traditional L1 and high-contact L1 varieties on the one hand to English-based creoles and Pidgins on the other. This impression is substantiated by progression of the percentages which are added in brackets next to the numbers in Table 2: the higher values in L1 varieties are typically to the upper left while the highest values among English-based Creoles and Pidgins are to the lower right.

With respect to region, there appears to be a trend that LIKE use is particularly common on the British Isles, in North America and in Australia and New Zealand, while it is typically absent from varieties spoken in Africa and the Caribbean (cf. Figure 8). This regional distribution does, however, reflect the distribution of different types of varieties, i.e. traditional L1 varieties vs. high-contact L1 varieties vs. indiginized L2 varieties vs. English-based Creoles vs. Pidgins. Nevertheless, according to Kortmann et al. (2004) and other sources, LIKE is particularly frequent in AusE (Kortmann & Szmrecsanyi 2004:1163, 1174), NZE (Miller 2009:317-337), and in IrE (Amador-Moreno 2010a:36-41, Harries 1993:176; Hickey 2007:376; Siemund, Maier & Schweinberger 2009:21-30; Schweinberger 2012), as well as across the British Isles (Kortmann & Szmrecsanyi 2004:1162), especially across northern non-standard varieties (Kortmann 2004:1100; Beal 2004:136) which have been refered to as Celtic Englishes (Kortmann 2004:1100). It is also accounted for in

the Southeast of England (Anderwald 2004); in CanE[13] (D'Arcy 2005), North AmE[14] (Buchstaller 2001; Dailey-O'Cain 2000; Fuller 2003; Schourup 1982, 1985; Siegel 2002); in IndE (Siemund, Maier & Schweinberger 2009; Valentine 1991); and in dialects and Creole varieties in the Americas and the Caribbean, e.g. Appalachian English, Newfoundland English, Chicano English, Belizean Creole, South Eastern American English enclave dialects, Trinidadian and mesolectal Tobagonian Creoles (Bayley & Santa Ana 2004; Schneider 2004:1115); in Nigerian Pidgin and Cameroon English (Kortmann & Szmrecsanyi 2004:1182).

The following subsections will elaborate on this survey and provides a more fine-grained overview of LIKE use across varieties of English, combining the regional distribution of LIKE use with a description of its positioning, origin, and pragmatic functions.

4.4.1.1 LIKE on the British Isles

In the Northeastern varieties of English in the British Isles, "as in Scots, *like* is used as a focusing device, with different discourse functions according to its position in the sentence. The most traditional function is as an emphatic device in clause-final position" (Beal 2004:136), as in (10):

(10) I'm a Geordie, me, like (Beal 2004:135)

Beal (2004) furthermore states that in clause-final position LIKE "can also be used in interrogatives, where it often conveys a sense of interest or surprise" (Beal 2004:136) as in (11), but it can also occur in clause-initial position to introduce or focus on a new topic as in (12):

(11) How'd you get away with that *like*? (Beal 2004:136)

(12) *Like* for one round five quid, that was *like* three quid, *like* two-fifty each. (Beal 2004:136)

[13] D'Arcy (2005) used data compiled in the city of Toronto, exclusively.
[14] Siegel (2002) used data compiled in the city of Philadelphia, exclusively.

With respect to clause-medial LIKE, Beal (2004) notes that "in younger speakers, in the North-east as in many other parts of the English-speaking world, *like* is also used within clauses, often as an explanatory device [as in (13)]. This means that *like* can occur several times within one sentence in the speech of younger people North-east of England" (Beal 2004:136) as in (12) above.

(13) They were *like* lightning, as they say, …his legs. (Beal 2004:136)

Beal's (2004) analysis supports the view that clause-medial LIKE is a rather new phenomenon which has entered northern varieties of British English by being adopted primarily by younger age-cohorts especially. In addition, she notes that quotative *BE like,* as with clause-medial LIKE, has only recently entered northern varieties of English and that only speaakers born after 1974 have used it in the NECTE corpus.

A similar picture is described by Anderwald (2004) and Andersen (2000) with respect to the south-eastern dialects of England, where "the discourse marker *like* seems to be an imported feature from the U.S." (Andersen 2001:216). In addition, Anderwald (2004:192) notes that "[l]ike *innit*, the pragmatic marker *like* is used almost exclusively by adolescents and young adults" and according to Andersen, 83 percent of the tokens of the pragmatic marker like are uttered by speakers aged 41 or younger in his material (cf. Andersen 2001:225). In contrast to Beal, who promotes the view that LIKE is predominantly a focus marker, Anderwald notes that LIKE "has a wide range of functions: it is used in 'ad hoc constructions', i.e. for purposes of approximation and exemplification [as in (14a) and (14b)]; it is used to construct a metalingistic focus [as in (15)]; it is used as a quotative after BE [...], and, very frequently, it is a hesitational device or a discourse link (Andersen 2001:209-299)" (Anderwald 2004:192-193) as in (16) below:

(14) a. It's just *like* all sticking out all over the place. (Andersen 2001:237)
b. You know what I mean it's *like* all plotted (Andersen 2001:237)

(15) It's *like* one day developing, right (Andersen 2001:242)

(16) I know and *like* … on Friday yeah … (Andersen 2001:255)

The hypothesis that clause-medial LIKE has developed rather recently is supported by the fact that "pragmatic *like* is not found in FRED material, which dates from the 1970s and the 1980s and contains the speech of mostly older speakers" (Anderwald 2004:193). However, Anderwald also notes with respect to clause-final LIKE that "the use of a distinct, 'traditional' dialectal *like* [is attested] for the Southeast of England as well, supporting Andersen's hunch that this dialectal *like* is not exclusively a northern phenomenon" (Anderwald 2004:192-193) as in examples (17) and (18) from FRED.

(17) but they 're dead and gone now *like*. And eh, I went out with eh,… (FRED LND_003 from Anderwald 2004:193)

(18) Used to come down here *like* and have the day (FRED KEN_001 from Anderwald 2004:193)

Anderwald (2004) also remarks that clause-final LIKE is "quite distinct from the new [clause-medial] uses as recorded in COLT by Andersen" (2004:193). The absence of clause-medial LIKE in the FRED data lets Anderwald suspect that "[it] is not implausible however that London is the source for the outward spread of these new - imported – uses of *like*, especially - perhaps most notably – of quotative *like* which is currently being recorded all over Great Britain (cf. Macaulay 2001)" (Anderwald 2004:193).

In Scottish English, "[t]here are two constructions with *like*, both discussed in detail in Miller and Weinert (1998). The older construction has *like* in clause-final position and is used by speakers to provide explanations and forestall objections" (Miller 2004:69-70) as in (19) below. LIKE also occurs in interrogative clauses, in which case "*like* can be paraphrased as IT clefts. [...] *Like* does not occur at pauses or when the speaker has planning problems. It is regularly equivalent to WH or IT clefts – [...] [and] regularly highlights items constituting an explanation" (Miller 2004:69-70). In a second and more recent construction, LIKE occurs in any position except clause-final as in (20) – both forms are unconstrained with regard to clause-type, i.e. to their occurrence in interrogative clauses. With respect to the discourse functions, Miller (2004) as

well as Miller and Weinert (1995; 1998) reject the claim that *like* serves as a filler as it is commonly not accompanied by hesitation or pauses which, for their part, would indicate a need for further processing time to structure a subsequent utterance. Hence, Miller and Weinert (1995; 1998) argue that *like* is integrated into the structure of the clause syntax (cf. Miller 2009:324).

In addition to LIKE, Miller comments on the alternative construction *likesae* which stands in complementary distribution, e.g. Irvine Welch uses *likesae* instead of LIKE in his novel *Trainspotting* (cf. Miller 2004:70).

(19) You had a wooden spile – you bored on the top of the barrel…and then you had ready a spile, which was a wooden cone about that length…and a soft wood naturally was porous and it would help to get his froth to let it work down – you had to be very careful you didn't take it right down *like*/it went flat. (Miller 2004:69-70)

(20) a. I mean and *like* you've not got any obstacles here have you? (Miller 2004:70)

b. To the lefthand side of East Lake? *Like* the very far end of East Lake?

In IrE, "[f]ocuser *like* is found in all age groups and is particularly common in explanatory contexts" (Hickey 2007:376) as in (21) and serves to "focus on a constituent or command the listener's attention (Harris 1993:176). In general terms, i.e. without regard to its pragmatic function, LIKE most frequently occurs in clause-marginal position, vastly outnumbering clause-medial use (cf. Siemund, Maier & Schweinberger 2009:21-30) and is allowed for in interrogative clauses as in (21).

(21)a. They'd go into the houses, *like*, to play the cards. (Hickey 2007:376)

b. 'Tis quality now, *like*, and all this milk and everything. You're getting paid on the quality of milk, *like*, and you could lose, *like*, you know … (Hickey 2007:376)

c. I'm just telling you what I heard, *like*. (Hickey 2007:376)

(22) Did you get to see him, *like*? (Harris 1993:176)

Amador-Moreno (2010b:531-544), investigating differences between the discourse marker LIKE usage in non-fictional data (LCIE = Limerick Corpus of Irish English) and fictional texts (*The Curious Incident of the Dog in the Nightdress* by Ross O'Carroll) in IrE, finds that the functional behaviour of the

discourse marker LIKE is very similar in both contexts (cf. Amador-Moreno 2010b:539). In both types of data, LIKE "indicates fuzzy thought; it is used as a hedge, often with a degree of hesitation involved. […] [M]id-position *like* is used when a speaker is searching for an appropriate expression, or an alternative term. It is also very often used with numbers, as a synonym of 'roughly', 'approximately' and it tends to be employed when a speaker/character is trying to emphasize the expression of certain feelings, when exaggerating, or describing unusual actions/surprising events" (Amador-Moreno 2010b:539).

4.4.1.2 LIKE in Australia and New Zealand

In Australia and New Zealand, LIKE is a pervasive or obligatory feature (Kortmann & Lunkenheimer 2011) and according to Miller (2009:317-337), the discourse functions of LIKE in AusE and NZE vary with respect to the position with a clause. While clause-initial LIKE serves to highlight clauses and phrases which may be used to exemplify a previous statement, clause-medial LIKE highlights phrases and the information they carry (Miller 2009:317-318). Clause-final LIKE, on the other hand, is used to anticipate objections, to provide explanations and to ask for explanation (Miller 2009:318). Accordingly, "[c]lause-final and clause-medial *like* both have an interpersonal role, the former because of its use by speakers to persuade their hearers to go along with an explanation or assertion, the latter because it is closely implicated in the give-and-take of discourse" (Miller 2009:336).

As argued previously on grounds of SctE data (cf. Miller & Weinert 1995; 1998), Miller (2009) counters claims with regard to a filling function of *like* in NZE and AusE, as it is neither accompanied by hesitations, nor by false starts "which would indicate lexical indecision or problems in planning syntactic structure" (2009:323).

4.4.1.3 LIKE in Indian English

The use of LIKE in IndE exhibits surprising results with respect to the positional distribution of LIKE as the usage pattern shows striking similarities

with the Irish usage of this marker (Siemund, Maier & Schweinberger 2009). According to this study, IndE speakers strongly prefer LIKE in clause-marginal position and, interestingly, use LIKE very frequently in clause-final position. Another study which makes a brief mention of LIKE in IndE was conducted by Valentine.

Valentine's data appears to be atypical when compared to other data sets used in the investigation of LIKE. In fact, her data consists of a corpus of

> "the natural speech of English from educated Indian women in a range of formal and informal settings. All were bilingual users, native speakers of Hindi. These speakers were residents of cities in North India: Delhi, Allahabad, and Meerut. Their ages ranged from 21-year-old college students to 60-year-old working and non-working women" (Valentine 1991:325).

With respect to LIKE, Valentine (1991:332-333) notes that although used less frequently and not performing the whole range of functions, LIKE in IndE is similar to the colloquial American use of this marker as it expresses a sense of superficiality (Valentine 1991:332).[15]

4.4.1.4 LIKE in US American and Canadian English

According to Tagliamonte (2005), the discourse marker LIKE is highly prevalent in the speech of adolescents in Toronto English, especially among 15- to 16-year old female speakers. Its use is "not haphazard, random, or indiscriminate. Instead, [it is] quite circumscribed and linguistically defined" (2005:1896). D'Arcy (2005) provides one the most detailed current accounts of LIKE.

Numerous studies have investigated the use of LIKE in various regions across the USA (e.g. Buchstaller 2001; Dailey-O'Cain 2000; Schourup 1982,

[15] Nevertheless, Valentine was not aware of uses of LIKE in other varieties; she notes in her conclusion that the relation and emergence of IndE should be studied more extensively with respect to other varieties of English (cf. Valentine 1991: 333-334).

1985; Siegel 2002). Although these studies have not or have only marginally employed a comparative or variationist perspective, i.e. explicitly compared their findings with the results of studies conducted on other regional varieties of English, US American English is probably the variety in which LIKE use has been most common. Dailey-O'Cain (2000) found that younger speakers in US AmE exhibit ahigher frequency of LIKE use, while gender did not appear to have a significant impact. Croucher (2004a) did find a gender difference, with female speakers using LIKE significantly more than male speakers and attributes this difference primarily to cultural causes (2004a:43). He claims that LIKE was "implanted in the American culture by a popular cultural creation in the 1980s, 'Valley Girls'" (2004a:43). Croucher (2004a) also found that LIKE does not represent natural pauses in speech and thus cannot be considered a filler element such as *uh* or *um* (2004:43). In her study of discourse marker use by native and non-native speakers, Müller (2005) found in her analysis of Californian English that speakers older than 25 used LIKE significantly less often than speakers under 25 (Müller 2005:231). Müller also found that females used LIKE to introduce examples less frequently than males of the same age (2005:328). Fuller (2003) found that LIKE "is employed to mark both inexactness and focus" (2003:375). In addition, she did not find evidence that LIKE use is caused or affected by the other speakers using LIKE (2003:375). Underhill (1988), who investigated the syntactic environments and functioning of LIKE, found that LIKE is used to introduce new information (1988:236-237), or in some cases to express vagueness (in cases where it can essentially be replaced by approximately) while stressing that its main function is to focus constituents within its scope (1988:237-238), or entire sentences (1988:242). Siegel (2002) argues that LIKE is typically a hedge, while Underhill states that this is less often so than typically presumed (1988:241). However, the majority of researchers who have investigated LIKE use from a functional perspective agree that it serves to convey inexactness (e.g. Fuller 2003; Schourup 1982; Siegel 2002).

The next section addresses attitudes associated with LIKE and surveys the findings of studies which have addressed this issue.

4.5 Attitudes toward LIKE

Following Williams, Coupland, and Garret (1999:333), attitudinal studies are particularly relevant regrading insight into the effects of stereotypes, i.e. perceived personal and social attributes, on communication and language change.

LIKE has received scholarly attention in various fields of linguistics over the past two or three decades, as well as attracting the popular press (Diamond 2000; Gross 2009; Johnson 1998; Levey 1999; Petersen 2004). In particular by purists such as Newman (1974), White (1955:303) or Wentworth & Flexner (1967), LIKE has been viewed negatively and regarded as symptomatic of careless speech, of lacking linguistic or cognitive function, and dismissed as meaningless (Newman 1974:15). Although these prejudices still appear to be part of public opinion, linguistics has come to reject the claim that LIKE is a mere 'meaningless interjection'. It has been argued repeatedly that in most of its occurrences, LIKE cannot be considered a mere hesitation phenomenon (e.g. Andersen 1997, 1998, 2000, 2001; Miller & Weinert 1995, 1998; Miller 2009; Müller 2005; Schourup 1982, 1985; Siegel 2002; Underhill 1988). In fact, "to claim that the 'aberrant' uses of LIKE are all meaningless is simply wrong" (Schourup 1982:45).

Recent attitudinal studies on LIKE have shown that there are widely shared beliefs about this feature which only superficially reflect its use. Dailey-O'Cain (2000), for instance, found that LIKE is perceived to be more frequently used by females than males. Similarly, she found that LIKE use is thought to be a feature of young speakers, typically adolescents; indeed, "young women are perceived as using like most often" (Dailey-O'Cain 2000:60). Remarkably, speakers were also rated as more attractive, more successful, more cheerful,

and more friendly when they used LIKE compared to when they did not (Dailey-O'Cain 2000:73). However, when speakers used LIKE, they were also perceived to be less educated, less intelligent and less interesting to older speakers (Dailey-O'Cain 2000:73).

A similar study by Buchstaller (2006b) focuses on differences and similarities concerning the attitudes towards the quotative *be like* in AmE and EngE. Buchstaller (2006b) elaborates on Dailey-O'Cain's analysis by investigating how and which attitudes are transferred from the donor to the preceptor variety. In this study, Buchstaller (2006b:362) concludes that attitudes are not merely borrowed, but are modified by the speakers in the preceptor variety to match their local norms and practices:

> This effectively means that, in cases of borrowing, the stereotypes attached to linguistic items are not simply taken over along with the surface item. Rather, the adoption of global resources is a more agentive process, whereby attitudes are re-evaluated and re-created by speakers of the borrowing variety.

Buchstaller's (2006b) analysis is highly intriguing, as the perceived attitudes of both discourse and quotative uses of LIKE are very similar (cf. Dailey-O'Cain 2000). Indeed, Buchstaller found that *be like* is often perceived as a typically adolescent feature, while the perceptional associations with gender and class are insignificant (2006b:368). This is remarkable, as Dailey-O'Cain (2000) found that *be like* in AmE is perceived not only as typical of adolescents but also of females. This difference is indicative of regionally distinct associations of *be like* which corroborate the claim that perceived attributes are not simply borrowed, but re-evaluated in light of local practices: "During the adoption process, speakers in the U.K. are attaching new and potentially different local social meaning to them [the incoming linguistic forms]" (Buchstaller (2006b:370). With respect to personality traits associated with *be like*, Buchstaller's (2006b) analysis corroborates and elaborates on Dailey-O'Cain's (2000) findings: speakers using *be like* have been rated more

"lively, cool and carefree, and with few academic aspirations" (Buchstaller 2006b:372) compared to speakers who do not use this form. With respect to perceptions about its geo-spatial origin, Buchstaller's (2006b) results are very interesting, as they do not match the commonly held hypothesis among academics that *be like* is typically an American feature. In fact, "British speakers do not seem to overwhelmingly perceive it [be like] that way. *Be like*'s association with the U.S. and more specifically California – frequently reported from America – does not translate into a strong perceptual trend in the U.K." (Buchstaller 2006b:375).

Hence, Buchstaller's (2006b) findings call into question whether innovative forms are associated with the reference group that triggered its spread. In other words, innovative forms "may lose or gain associations during the process [of adoption], or, alternatively, already existing percepts may be re-analyzed and re-evaluated" (Buchstaller 2006b).

The insights offered by Dailey-O'Cain's (2000) and Buchstaller's (2006b) attitudinal study are highly intriguing when contrasted with common lore or the accounts given by prescriptivists. While common perception matches the findings of systematic quantitative studies with respect to age, other popular claims, e.g. the claim that LIKE is particularly favored by female speakers, are not unanimously endorsed by thorough scientific analysis.

The present cross-varietal survey will elaborate on these perceptions and evaluate the extent to which they agree with the settings in different localities. Finally, based on Buchstaller's (2006b) findings, it appears questionable that the supposed association between LIKE and American adolescents as a social reference category sufficiently explains the rise of LIKE in geographically distinct areas. In fact, it is very plausible that the association with personality traits such as "lively, cool and carefree" (Buchstaller 2006b:372) have higher explanatory power than do models merely emphasizing the importance of social class. Nonetheless, personality traits might in turn be perceived to be typical features of certain social cohorts and, therefore, serve to construct

social categories as proposed by, for instance, Eckert (2001). This point will be re-addressed later.

4.6 The syntax of LIKE

According to popular lore, LIKE "can be used anywhere in a sentence" (D'Arcy 2007:388). The belief that LIKE is syntactically unconstrained is, however, not limited to laymen, but is shared by scholars (Brinton 1996:34; Siegel 2002:38). Despite the popularity of claims to the contrary, even early scholarly accounts of LIKE (e.g. Schourup 1982; Underhill 1988) documented that this vernacular feature occurs predominantly in circumscribed and linguistically defined environments (Tagliamonte 2005:1986).

This misconception probably arises from the confusion of syntactic optionality and syntactic flexibility. In other words, scholars have mistakenly equated the fact that "the absence of discourse markers does not render a sentence ungrammatical/ or unintelligible" (Fraser 1988:22) with a lack of syntactic constraints. However, the fact that discourse markers are grammatical optionality does not necessarily imply that they are "'outside' the syntactic structure" (Erman 2001:1339, Brinton 1996:34; Jucker & Ziv 1998:3). If discourse markers in general and the discourse marker LIKE in particular were 'outside' the syntactic structure as Erman (2001:1339) states, then it should be expected that these features are random or indiscriminate. This view, which holds that LIKE use may most accurately be described as *random variation*, has repeatedly been disproved by systematic, quantitative analyses of its use (e.g. Andersen 1998, 2001; D'Arcy 2005, 2007, 2008; Tagliamonte 2005). Accordingly, it is now widely accepted that the syntactic constraints and positioning of LIKE are neither haphazard, indiscriminate, nor arbitrary (cf. D'Arcy 2005, 2007; Tagliamonte 2005).

In fact, there is ample evidence for an underlying systematicity of LIKE use, as it is restricted to definable syntactic contexts and in contrast to other

discourse markers such as *well*, *oh*, *I mean* and *you know*, it appears to be "more deeply integrated as a clausal constituent of the syntactic context in which it occurs" (Andersen 1997:37). Underhill notes in this context that "it is not the case that *like* can be placed anywhere in a sentence" (1988:243), as it is 'closely rule-governed', nearly always introducing a constituent, and frequently marks new, and significant information (1988:236). The claim that LIKE has "developed gradually and systematically" (D'Arcy 2005:ii) is expressed even more emphatically by D'Arcy, who states that "LIKE is not random, but interacts with syntactic structure in regular and predictable ways" (D'Arcy 2005:ii).[16] Thus, it is, in fact, more accurate to describe its occurrence as an "*orderly heterogeneity*" (cf. Weinreich, Labov & Herzog 1968:100) rather than the indiscriminate, random variation described above. Hence, under closer inspection, one finds that LIKE comprises a heterogeneity of quite distinct uses which occur under specific conditions and in rather well circumscribed contexts (e.g. D'Arcy 2005:ii; Tagliamonte 2005:1897). For example, Andersen (2001) observes that LIKE is less likely to occur within phrases particularly when they are idiomatic or have a high degree of syntactic rigidity. Furthermore, he suggests that LIKE is more likely to immediately precede the lexical material of a phrase, as opposed to grammatical words (cf. also Tagliamonte 2005:1901). A case in point is the positioning of LIKE with respect to the verb. Consider D'Arcy (2007:408):

> In the current data set, the syntagmatic order of *like* and verbs is highly fixed: the particle categorically occurs to the immediate left of the lexical verb. Thus, when functional morphemes such as modal verbs, auxiliary verbs, and infinitival to are present, *like* appears between these and the main verb.

[16] Although many scholars agree that LIKE is at least to some degree governed by syntactic constraints which restrict its occurrence to well-defined slots (Andersen 1997, 1998, 2000, 2001; Dailey O'Cain 2000; D'Arcy 2005; Romaine & Lange 1995; Tagliamonte 2005; Underhill 1988), the rigidity and nature of these constraints as well as their area of application with respect to geographical variation require further research.

This observation has been confirmed for a number of L1 varieties, as Underhill (1988) has shown for AmE, D'Arcy (2005) and Tagliamonte (2005) for Toronto English, and Andersen (1998) for London teenage speak, i.e. vernacular EngE. Indeed, this constraint appears to be consistent across other varieties (cf. (23)).

(23) a. But as I said all of the banks are *like* closing…. (ICE Jamaica:S1A-063$B)
b. No the one where they were uhm they were *like* worshipping that golden cow or something that they have made. (ICE Philippines:S1A-007$B)
c. [S]he was *like* standing sort of waiflike in the background. (ICE New Zealand:S1A-069$W)
d. And to purchase a ticket you have to *like* fight with persons just to get a ticket. (ICE-India:S1A-061$B)
e. [H]e was *like* going nuts. (ICE Ireland:S1A-048$A)
f. [H]e could *like* stretch. (ICE Canada:S1A-075$B)

Nonetheless, there are rare exceptions which occur predominantly, though not exclusively, in outer circle or second language varieties (cf. (24)).

(24) a. Oh no yeah you *like* have two more weeks before that. (ICE Philippines:S1A-013$A)
b. [S]he used characters of the Corinthians who *like* were great uhm great uh great soldiers. (ICE Jamaica:S1A-065$NA)
c. She's a lot more tired and she sort of like has to take time off work over it. (ICE GB:S1A-079$C)

In summary, the assertion that speakers use LIKE "almost anywhere in a sentence'' (Randall 1988:206) is incorrect. Despite allowing for a certain degree of syntagmatic flexibility, LIKE is not random, "although to some hearers it may seem so'' (Meehan 1991:40).

4.7 The discourse-pragmatic functions of LIKE

The vast majority of research dedicated to the analysis of discourse markers in general and LIKE in particular has mainly focused on pragmatic functions, for the most part following discourse analytic approaches. In recent years, "systematic corpus-based examinations of language variation and change produced a gently expanding body of evidence [which empirically validated] that they are neither superfluous nor random insertions in discourse" (Pichler & Levey 2010:17). Pichler and Levey (2010:17) point out that

> discourse-pragmatic features are strategically used by interactants to signal speaker attitudes and structure discourse […], and exhibit structured heterogeneity in vernacular usage. (Pichler & Levey 2010:17)

Concerning methodology, discourse analytic studies of *like* have predominantly employed fine-grained, qualitative approaches to describe the contexts in which LIKE occurs as well as the communicative purposes it serves. Schourup (1982, 1985) has provided an extensive account of uses of LIKE in AmE which serve quite distinct functions. With respect to function, Schourup (1982, 1985) views LIKE as an evincive, i.e. a linguistic item which

> indicates that at the moment at which it is said the speaker is engaged in, or has just been engaged in, thinking; the evincive item indicates that this thinking is now occurring or has just now occurred but does not completely specify its content. (Schourup 1982:14)

In this sense, evincives are lexically empty and, although they have a core use, they require contextual interpretation and, thus, have procedural rather than lexical meaning – a claim which is argued for not only by Schourup (1982, 1985), but more recently especially by Andersen (1997, 1998, 2000, 2001).

Andersen pursued a similar, yet more elaborate approach to discourse markers in general and LIKE in particular. At its core, this approach rests on the Gricean principle of conversational relevance as re-formulated by Sperber

and Wilson (1986) (Brinton1996:7). According to Andersen (1997, 1998, 2001) and Blakemore (1987, 1988a, 1988b, 1990), similar to discourse markers such as *so*, *after all*, *therefore*, *you see*, *afterwards*, *but*, *however*, and *moreover*, LIKE does not so much carry propositional content, but "minimizes the hearer's processing costs by limiting the context, or set of assumptions (old information), used in interpreting the proposition (new information)" (Brinton 1996:7). From this perspective, the main function of discourse markers is thus to indicate how the relevance of one discourse segment or utterance is dependent on another, i.e. its immediate context (Redeker 1990:372; Blakemore 1987:125; Brinton 1996:30).

As LIKE has been studied extensively and from various perspectives, a wide variety of supposed functions and meanings have been proposed. Among the meanings and functions discussed within the linguistic literature, we find (i) hedging (Siegel 2002); (ii) non-contrastive focusing (Romaine & Lange 1991; Underhill 1988; Miller & Weinert 1995, 1998; Miller 2009); (iii) buying processing time (Siegel 2002); (iv) indicating that the passage to follow is difficult to formulate/ holding the floor (Schourup 1982; Siegel 2002); (vi) signaling a minor non-equivalence between what is said and what is in mind (Schourop 1982); (vii) signaling 'loose talk', i.e. marking non-literalness (Andersen 1997, 2000); (viii) signaling approximation (Andersen 1997, 1998, 2000, Buchstaller 2001; D'Arcy 2005; Schourup 1982, Siegel 2002); (ix) introducing exemplifications (Meehan 1991; Miller & Weinert 1995, 1998; Miller 2009; Schourup 1982); (x) signaling similarity (Meehan 1991).

Considering the flexibility of LIKE as well as its high frequency in spoken discourse, it is not surprising to find such a variety of functions being proposed and discussed within the respective literature.[17] Recently, a number of studies have emerged (e.g. Andersen 1998, 2000; Miller & Weinert 1995; Miller 2009)

[17] What is, nevertheless, quite remarkable, is the fact that LIKE is supposed to fulfil apparently converse functions simultaneously such as hedging (e.g. Siegel 2002) and focusing (e.g. Underhill 1988).

which link certain discourse-pragmatic functions to the clausal positions. Andersen (1997, 1998, 2000), for example, suggests that all uses of LIKE fall into either of two categories, which he terms clause-internal and clause-external. While clause-internal uses are "syntactically bound to and dependent on a linguistic structure as a pragmatic qualifier of the following expression" (Andersen 2000:273) as in (25), clause-external LIKE is "syntactically unbound (parenthetical) when it is external to and independent of syntactic structure" (Andersen 2000:273) as in (26).

(25) Clause-internal, syntactically bound LIKE

a. I thought Jews'd always been very *like* <,> stringently against divorce. (ICE Ireland:S1B-005$D)

b. I got quite good at *like* heating up thermometers and stuff and give myself a temperature and things. (ICE GB:S1A-076$B)

c. It's *like* right in the middle of nowhere. (ICE Canada:S2A-038$A)

d. But then all the more you need *like* a program for orienting. (ICE-Philippines:S1A-085$A>

e. Well they/we said that pidgins are *like* the beginning of Creole. (ICE-Jamaica:S1B-001$B)

(26) Clause-external, syntactically unbound LIKE

a. I mean that <,> *like* uh <,> there is no such <,> parking problem or some. (ICE India:S1A-077$A)

b. *Like* <,> I don't know probably ten <,> nine inches <,> nine or ten inches. (ICE Canada:S1B-063$B)

c. Uhm <,> I don't know <#> UCD *like* first of all well well UCC supposedly it's meant to be easier to get into second year Psychology. (ICE Ireland:S1A-048$B)

This distinction mirrors Brinton's (1996:1) classification into 'textual' and 'interpersonal', i.e., subjective uses, which were based on Halliday. This general differentiation is an adequate heuristic, as it can be applied to almost all instances occurring in natural-language data. Nevertheless, it does not suffice for more fine-grained analyses of highly frequent and extraordinarily flexible discourse features.

Andersen's (1997, 1998, 2000, 2001) analysis of LIKE deserves additional attention, as his study is perhaps the most extensive and fine-grained analysis of LIKE's pragmatic functions todate. To elaborate, he provides extensive

discussion of a wide variety of instances of LIKE and his analysis encompasses a wide variety of functions. In the following, I will provide an overview similar to Andersen's, which will provide and discuss examples of instances of LIKE, which occur in the present data. This discussion consists of three parts: the first will deal with clause-internal uses; and the second part will discuss clause-external uses. Finally, the third part will deal with instances of LIKE which occur in multiword units, such as *something like that* or *it's like*. Such multiword complexes are different from, but still relevant to, an analysis of instances of LIKE as a single discourse marker and, hence, require special attention.

4.7.1 Clause-internal LIKE

The following section will discuss instances of clause-internal LIKE. The structure of the following subsection tries to account for functional similarities between distinct occurrences of LIKE, in the sense that uses of LIKE, which share common features and have probably evolved from a common ancestor, are subsumed under the same headings.

4.7.1.1 LIKE as a hedge

One of the most widely attested functions of LIKE is its use as a pragmatic hedge. In these instances, LIKE serves to down tone a statement or assertion and conveys a sense of vagueness, which communicates the speaker's stance towards the (literal) truth of what was said. Hence, LIKE enables speakers to evaluate and to provide metalinguistic comment on propositions, thereby allowing speakers to pull back from assertions "in a rather non-committal fashion" (Meehan 1991:50). Accordingly, LIKE serves to down tone a statement by communicating to the hearer that the speaker is providing a more or less rough approximation of the actual state of affairs (Hedevind 1967:237). Using LIKE, speakers signal that the constituent which is modified by LIKE is not intended to be taken as precise or literally true, but that it is to be interpreted with a certain degree of vagueness, either relating to the amount referred to or

the literal truth of the proposition (cf. Andersen 1998, 2000). Indeed, Andersen (1997, 1998, 2000, 1001) describes the ability of LIKE to communicate metalinguistic comment, i.e. that LIKE in general expresses a "non-literal resemblance between utterance and thought" (2000:219), as an overarching feature of all uses of LIKE.

In a similar vein, Schourup (1982, 1985) asserts that in instances of hedging, LIKE "is used to express a possible unspecified minor nonequivalence of what is said and what is meant" (Schourup 1982:31). Hence, hedging LIKE is an equivalent of other hedges, i.e., segments such as *sort of* or *kind of*, which express that the modified segment, be it a single word, phrase, complement clause, or utterance, should be interpreted as less than literal (Andersen 1997, 1998, 2000, 2001). Although it is at times difficult to determine whether an instance of LIKE highlights or focuses a certain segment, the co-occurrence of other hedges provides ample reason to classify the respective LIKE as a hedge, as in (27).

(27) a. I thought that was pretty weird sort of *like* expecting you to do that. (ICE Canada:S1A-084$A)

b. [T]hat's who I sort of *like* hang about with and stuff and all. (ICE Ireland:S1A-014$B)

c. [W]e need a large kind of *like* psychological finale. (ICE Great Britain:S1B-079$A)

Instances such as those in (27) allow a speaker to distance him- or herself from the utterance, either relating to exactness of measurable amounts or with respect to how convinced a speaker is of something that is said about something or, more importantly, about someone. This is in line with Grant and Main Dixon's (1921:142) account of LIKE, which states that it is added adverbially to soften an expression. In fact, hedging is the earliest of the functions attested to for LIKE. In early accounts (cf. Wright 1902), LIKE is said to modify a statement and can be glossed as 'rather', or 'as it were' (Wilson 1915:98).

Within the group of instances in which LIKE is used as a hedge, three distinct types can be differentiated: (i) LIKE before (exact) numeric

expressions or other measurable constituents as in (28); (ii) LIKE before phrases which do not denote exact quantities, but more abstract notions, such as depth, size etc. as in (29); and (iii) instances where LIKE serves as a prototypically pragmatic hedge with little or no lexically definable content, in which case it precedes phrases commonly not denoting quantities, as in (30). In the following, a closer look will be taken at each of these types of LIKE as a hedge:

(28) a. [S]o we got *like* three points right here. (Santa Barbara Corpus SBC024:DAN)
b. They put it in a machine and it's in there for *like* three days<,> but only thing it comes out yellow. (ICE Jamaica:S1A-025$B)
c. I said yeah we have it it's it's *like* eighty-nine point five. (ICE Philippines:S1A-026$A)

(29) a. And it was *like* maybe a foot wide. (ICE- Philippines:S1A-007$B)
b. I don't really want do that for *like* too long. (ICE Jamaica:S1A-037$A)
c. It's only about *like* that size and you look into it and you can read but except I couldn't read anything. (ICE Ireland:S1A-035$B)

(30) a. It's *like* a kind of a turquoise background <,,> with <,> little flowers on it. (ICE Canada:S1A-098$A)
b. [I] mean it's not *like* directly obvious. (ICE New Zealand:S1A-042$J)
c. [T]hat's how it happens but <,,> slowly <,,> you became *like* used to <,,> all those things <,,>. (ICE India:S1A-090$B)

In the first two contexts (28-29), LIKE is used to express a notion of approximation, similar to but not exactly equivalent to other traditional adverbs of approximation. In contrast, the third instance is not so much expressing approximation or that the listener is not supposed to take the semantic content of the modified constituent as being precise, but it serves rather to express a subjective evaluation of the utterance and thereby allows the speaker to distance herself from what she has just said. Hence, this third variant links LIKE to epistemic modality, evidenciality, and issues concerning politeness.

When LIKE is used as a pragmatic hedge as in (31), it serves to tone down a statement, i.e., it expresses the speaker's attitude towards the utterance by indicating that the speaker is not fully convinced either of the truth expressed

in the statement, or that the utterance may not be framed or phrased optimally. In other words, LIKE – when used as a hedge – serves an expressive function which is to create a sense of vagueness and, thus, allows the speaker to distance him- or herself from the proposition. This element of subjective evaluation of the truth of a proposition links LIKE with the domain of epistemic modality (Buchstaller 2001:32). This is especially true when LIKE is used for politeness (Schourup 1982:47), in which case it falls into the category of 'modality markers' (House & Kasper 1981) and serves to avoid "a precise propositional specification thus circumventing the potential provocation such a specification might entail" (House & Kasper 1981:167).

(31) Uhm <,> I yet have to wait for their call on Monday before I finalize everything but so far uh as far as I'm concerned I'm going there on the fourth then we have *like* a reunion on the fifth then I'll be back on the sixth. (ICE-Philippines:S1A-087$A)

Buchstaller (2002) makes a strong case for the quotative complement LIKE having evolved from the pragmatic hedge variant of LIKE which she separates from the epistemic hedge variant (LIKE preceding numeric expressions and other measurable constituents in the present classification). According to Buchstaller (2002), epistemic hedge and pragmatic hedge differ with respect to the communicative level on which they operate. While the epistemic hedge operates on the referential-epistemic level, the pragmatic hedge is located on the interpersonal-pragmatic level of communication (2002:4). Both variants, nevertheless, share a "basic core meaning of like [which is] the notion of similarity which is the basic underlying notion of both comparison and approximation" (Buchstaller 2002:3). Siegel (2002) agrees with the concept of *like* carrying a core-meaning which expresses similarity, but, in addition, treats *like* as a marker of intimacy or solidarity, especially between adolescent interlocutors. With regard to semantics, Siegel (2002:66) claims that LIKE affects core semantics in the sense that it has an effect on the interpretation of quantifiers and truth conditions of sentences. Andersen (2000:264) argues in a similar vein and concludes that LIKE puts subsequent material in a metalinguistic focus and contributes to utterance interpretation.

On an interactional level, however, LIKE serves metalinguistic purposes or strategies, such as avoiding face-threatening and marking a speaker's identity. These metalinguistic functions refer simultaneously to the basic core meaning, but additionally, they require semantic bleaching as a prerequisite for indicating a less than literal interpretation. This is true, especially, when *like* occurs in questions where it enables speakers to distance themselves, soften the request and shield themselves in case of refusal as in (32) (Underhill 1988:241).

(32)[18] One sister asking another: Could I *like* borrow your sweater?

(33) (A guy walked up to a girl sitting in a bar.)
Guy: Can I *like* nibble on your neck? [It turns out he was her boyfriend.]

(34) A very wired teacher: I'm so tired. I'm really going to rest this weekend. I mean *like* stay in bed all day Saturday and Sunday.[19]

On the level of speaker interaction, *like* is supposed to mark 'loose talk' by indicating "simply that the speaker's words are an inexact formulation and should not be understood as a complete or accurate portrayal of what the speaker has in mind" (1982:31). In this sense, like expresses "a possible discrepancy between what the speaker is about to say and what the speaker feels ideally might or should be said" (Schourup 1982:31). Relating to this interpretation, Andersen (1998, 2000) suggests that LIKE serves to 'impoverish' the proposition in order to give hearers the maximum of useful information with a minimum of processing effort, thereby creating optimal relevance (Miller 2009:324).

[18] Examples (32) to (34) are adopted from Underhill (1988: 241).

[19] Instances such as (32) to (34) display the close link between hedging uses of LIKE and LIKE with a focusing function. In fact, in many instances it is very difficult to distinguish between these two functions, especially when no other contextual information is present. As there were no audio data available for the current analysis, the evaluation was limited to the written transcription of the spoken data provided by the ICE corpora. The transcriptions were nevertheless very precise, and included metalinguistic information such as pauses and turn delimiters.

4.7.1.2 LIKE preceding numerical and measurable expressions

Discourse markers are defined by various authors (e.g. Brinton 1996; Hölker 1991; Jucker & Ziv 1998; Schiffrin 1987, 2001) as items which do not interfere with the truth conditions of the propositions in which they are embedded. Nevertheless, LIKE infringes upon this property whenever it precedes numeric expressions (Andersen 2000; Siegel 2002). LIKE does, however, quite often precede numeric expressions, i.e. exact numbers (Schourup 1982:29), or other measurable constituents as in (35). In such cases, LIKE communicates a notion of approximation and behaves similarly to other traditional approximating adverbs, such as *roughly*, *about* or *approximately* (cf. D'Arcy 2005, 2008).

(35) LIKE before numeric expressions and other measurable constituents

a. They're trying to get a guy to drive down but you know it's *like* nine hours up and down drive. (ICE Ireland:S1A-027$A)

b. Souhm <,> well the persons who who do it for *like* eighteen hundred is it good. (ICE Jamaica:S1A-083$B)

c. It's been like <,> we've been on for *like* a month and a half and <,> we're doing great. (ICE Philippines:S1A-048$A)

The status of LIKE preceding numerical expressions or other measurable constituents, is, however, controversial as in such contexts LIKE is "a borderline case between adverbial and pragmatic marker" (Andersen 2000:260). D'Arcy (2005:33) argues that LIKE preceding numeric expressions cannot be considered a discourse marker, as it carries propositional content and thus defies the criterion of semantic emptiness. In other words, "there is a specifiable semantic difference between descriptions preceded by <u>like</u> and identical descriptions without <u>like</u>" (Schourup 1982:30).

D'Arcy's (2005) view is supported by the fact that *like* in measurable, numeric contexts interferes with the truth conditions of propositions (cf. Andersen 2000; Siegel 2002) as do "other truth-conditional adverbs such as *roughly*, *nearly* and *about*. Such uses of *like*, therefore, interfere with truth-conditions, which has been repeatedly stated as a defining feature of discourse or pragmatic markers (cf. Brinton 1996). An additional argument for this view is that *like* is not only in complementary distribution with other traditional

adverbs such as *roughly*, *about*, or *approximately* before measurable constituents, but that it does, in fact, even replace them. This means that it interferes significantly with other more syntactically entrenched parts of speech and is, therefore, "functioning primarily as an adverb in this context" (D'Arcy 2005:429). For these reasons, various researchers (Biber et al. 1999; D'Arcy 2005; Underhill 1988) consider it an adverbial rather than a discourse or pragmatic marker.

Despite acknowledging these arguments, Andersen (2000:260) still regards instances of LIKE which precede numeric expressions as pragmatic markers rather than adverbials. He argues that in these contexts:

> [LIKE] has the function of signalling that the utterance contains a loose interpretation of the speakers thought, and that the speaker does not commit herself to the literal truth of the utterance, in a way which the adverbials roughly and approximately could not do. (2000:260).

Schoroup's (1982) analysis supports Andersen's interpretation. According to Schoroup (1982:30), some instances of LIKE preceding numeric expressions have a different reading than other traditional adverbs of approximation as in (36). In cases such as those provided by Schoroup (1982), LIKE is more appropriately glossed as "as it were" or "so to say" (1982:30).

(36) SUE:
You know that um – they've been livin(g) in this big three-story house with basements- like four floors y'know- gigantic house on Summit… (LAB-B, 22)[20]

Schourup (1982:31) provides further support for the non-equivalence of LIKE and traditional adverbials of approximation in that it co-occurs with such traditional adverbials. This co-occurrence with traditional adverbials of approximation is, however, only very rarely the case (cf. Schweinberger 2010).

[20] This example was originally used in Schoroup (1982: 30).

The question which remains is whether LIKE preceding numeric expressions should be treated as an adverb of approximation or as a discourse marker.

A recent analysis of *like* in the Irish and Canadian components of the ICE corpora (Schweinberger 2010) supports D'Arcy's (2005) claim that *like* in the contexts described may be classified instead as an adverb of approximation. The results of that analysis show that although *like* co-occurs with traditional approximating adverbs with a significantly higher frequency than would be expected if *like* did not differ from such adverbs, the low rate of co-occurrences as well as the small effect size indicate that its syntactic behavior does not differ substantially from traditional approximating adverbs. A linear regression provided additional evidence that *like* replaces traditional approximating adverbs and should, therefore, be viewed as an adverb rather than a discourse marker. Accordingly, such instances of *like* will be excluded from the present analysis.

4.7.1.3 LIKE before inexact quantities

LIKE preceding inexact or more abstract segments differs from cases in which *like* precedes exact numeric expressions, as theses occurrences do indeed represent a true borderline case between adverbial and discourse marker use. Schourup (1982:31) – in an analogy to the instances dealt with above – argues that LIKE expresses not a mere approximation, but informs the listener about a minor nonequivalence between what the speaker had in mind and what the speaker has actually expressed:

> *Like* [italics M.S.] in the above cases [i.e., LIKE preceding inexact or more abstract segments] can be described as indicating a possible discrepancy between what the speaker is about to say and what the speaker feels ideally might or should be said. Like in this use can be seen as a device available to speakers to provide for a loose fit between their chosen words and the conceptual material their words are meant to reflect. (Schourup 1982:31)

On the one hand, such instances of LIKE are, according to Schourup (1982), very similar to the pragmatic or epistemic hedge as discussed in section (4.7.1.1). On the other hand, they differ from instances in which LIKE functions as a pragmatic hedge in so far as it occurs in a different conceptual environment. The sense of approximation which is more or less the sole function of LIKE preceding exact numeric expressions is not quite as appropriate when the constituent within the scope of LIKE is inexact, or a rather abstract entity such as depth, width, or size. The reason for this is that the sense of approximation is already conveyed in the abstractness in the constituent itself which, although denoting a quantity of some sort, remains vague and thus does not require additional approximation. Although uses of LIKE in this described context have features in common with both LIKE before exact numeric expressions and hedging LIKE, it appears to be more fitting to subsume instances of LIKE which precede inexact or abstract quantities under discourse markers which serve a hedging function.

4.7.1.4 Focus LIKE

In contrast to instances of LIKE which tone down statements and thus function as pragmatic hedges, various sources (Meehan 1991; Miller & Weinert 1995, 1998; Streeck 2002; Underhill 1988) argue that LIKE has taken on a role as textual focus marker, i.e. as "a type of unit that marks subsequent talk as salient, for example, as new information" (Streek 2002:583). Meehan even "hypothesizes that the main function of *like* in these constructions [before adjectives in particular] is to focus the listener's attention on specific information" (1991:44). Although predominantly found in the Celtic Englishes, such as IrE and SctE, this function is attested in various varieties of English as, for instance, in AmE as well as NZE and AusE (Beal 2004; Hickey 2007; Kortmann & Smreczanyi 2004; Miller & Weinert 1995, 1998; Miller 2009; Underhill 1988).

The paper which popularized the notion that LIKE is best interpreted as a focusing device was Underhill's (1988) study, in which he suggests that various

instances of clause-medial *like* serve to focus on or highlight constituents to its immediate right. The constituents LIKE highlights can be of varying length, i.e., it may modify single words (37), phrases (38), complement clauses (39) or matrix clauses and utterances (40) (Fleischman & Yaguello 2004:131).

(37) a. Gosh that was *like* brutal. (ICE Canada:S1A-083#A)
b. It's *like* burning. (ICE India:S1A-001#A)

(38) a. I'd go home *like* every weekend. (ICE Jamaica:S1A-067$B)
b. It's quite good I mean it's interesting 'cause it's *like* a new building and then everyone who works there uhm is given an ID with those bar I mean with the bar code in it. (ICE Philippines:S1A-012$A)

(39) a. She has episodes of *like* wanting to kill herself or to hurt herself. (ICE-Canada:S1A-028$C)
b. I got quite good at *like* heating up thermometers and stuff and give myself a temperature and things. (ICE-Great Britain:S1A-076$B)

(40) a. *Like* I saw Pikachu and Raichu fighting. (ICE Phippines:S1A-001$B)
b. That indeed there is no longer you see <,> either religious or devotional <,> *like* it seems to have lost its meaning altogether. (ICE India:S1A-093$B)

Focusing LIKE frequently marks new and, specifically, significant information (Underhill 1988:236) as in (37) to (43). These instances of LIKE are similar to uses as in (41), where LIKE not only marks unusual notions – particularly ideas which are not intended to be taken literally – but also indexes the unreality of a statement (Underhill 1988:241) as in (42).

(41) (A guy walked up to a girl sitting in a bar.)
Guy: Can I *like* nibble on your neck? [It turns out be was her boyfriend.]

(42) A very wired teacher: I'm so tired. I'm really going to rest this weekend. I mean *like* stay in bed all day Saturday and Sunday.

Miller and Weinert (1995) elaborate on Underhill's (1988) notion that LIKE marks focal information. They argue that in contrast to other focusing devices in English, such as IT-clefts, IT-BE sentences or WH-clefts, LIKE is used as a non-contrastive, non-presentational focus marker which is not restricted to new information, but serves to guide the listener's attention to parts of utterances which the speaker evaluates as especially important.

In contrast to Underhill's (1998) assumption, LIKE need not necessarily modify items to its immediate right, as clause-final LIKE can also fulfill this function although its scope points forward to its preceding constituent, as in (43) and not to the following segment, as in (37) to (40).

(43) a. Ah it's it's alright *like* <ICE-IRE:S1A-027$B> In small doses. (ICE Ireland:S1A-027$C)
b. That becomes dangerous *like* <ICE-IND:S1A-041$B> Yeah. (ICE India:S1A-041$A)
c. People were silly *like* you know. (ICE Ireland:S1A-023$B)

The observation that utterance-final LIKE frequently marks focus is corroborated by Columbus (2009), who views LIKE as an invariant tag. However, Columbus (2009) rejects the notion that this form is functionally restricted to focusing and draws attention to the multifunctionality of this element.

With respect to the origin of focusing LIKE, Buchstaller (2001) argues that it originated from and is an extension of the 'similar to' meaning of LIKE. Building on Haiman (1989:310), who argues that "the comparative construction is one which contrasts, and, hence, focuses the elements which are compared […] the element compared being highlighted", Buchstaller states that "the semantic link between comparing and focusing seems to be a fairly salient one" (2001:32).

While Underhill (1988) allows for alternative discourse functions of clause-medial *like* such as hedging, Miller and Weinert (1995, 1998), and also Miller (2009), claim that clause-medial *like* is almost exclusively employed for focusing the subsequent constituent. The main argument put forward on their behalf is the absence of pauses and hesitations accompanying *like* use: "[c]lause-medial like highlights phrases and the information they carry" (Miller 2009:318). Siegel (2002) as well as Fuller (2003:368) object to this claim, as the argument that like is restricted to indicating focus is not supported by data; in fact, Siegel claims that LIKE does not uniformly mark focus in all contexts, as, for example, when it precedes numeric expressions (Siegel 2002:41).

Although the focus function of LIKE is not uncontroversial, most researchers (e.g. Andersen 1998, 2000; D'Arcy 2005; Fuller 2003, Meehan 1991, Miller & Weinert 1998) agree that it does serve this function in certain contexts. Nevertheless, the strong hypothesis that it is restricted to this function in clause-medial environments (e.g. Miller 2009) remains contested. The controversial status of LIKE as being a maker of focus relates to the discrepancy between its morpho-syntactic similarity and its discourse-pragmatic multifunctionality. In other words, assigning unambiguous pragmatic functions to certain instances of LIKE is problematic if not impossible. In other words, based merely on morpho-syntactic criteria, differentiating between instances which serve to hedge or to focus a certain constituent, is almost impossible given that phonological cues are missing. Indeed, most instances cannot be confidently assigned to either of these functions. One case in point illustrating this conundrum is the fact that Siegel (2002:40-41) more or less dismisses the notion that LIKE marks focus. In her view, LIKE only superficially seems to mark focus, while it actually expresses what Schourup (1982:31) calls a minor non-equivalence between what is said and what is meant. Consider Siegel (2002:41): "[I]f *like* seems to mark new or focused material, it is because that is the material that speakers are most likely to be insecure about describing accurately". According to Siegel (2002), LIKE as a focus marker is, hence, a secondary phenomenon, based on the re-analysis of its use as an index of inexactness before new or focused material. In support of her argument, Siegel proceeds by asserting that "*like*'s being a focus marker cannot explain its ability to interfere with the Definiteness Effect only in sluicing and existential *there* constructions or to change truth conditions" (Siegel 2002:41).

The almost diametrically opposed interpretations of very similar instances of LIKE by Siegel (2002) and Underhill (1988) point to the difficulty of unambiguously assigning the speaker's intended discourse pragmatic function to certain individual instance. A promising approach for operationalizing this issue was suggested by D'Arcy (2005), who argued that focusing should co-

occur significantly more often with definite phrases, thus, testing for the co-occurrence of LIKE with definite articles. But although this approach seems quite appropriate at first, it does not hold for instances such as (44) and (45) below:

(44) a. So that's where you get these women or mo mm these parents who are definitely grandparent age you know if not great-grandparent age who have these *like* really young children. (ICE Canada:S2A-039$A)

b. And Nigel was there and Jim wasn't really drinking and Elaine was on a high cos she got this *like* letter about her getting the <,> the <,> whatever the the <,> the job <,> or not the job even the interview and that so she was on a high. (ICE Ireland:S1A-049$A)

(45) a. That is *like* really bad slang Spanish. (ICE Canada:S1A-051$B)

b. [I]t seems to be a lot more productive when you have *like* a male dominated field team right. (ICE Jamaica:S1A-007$A)

c. If the roof isn't good <,> that's *like* a very important part of any house you buy. (ICE Ireland:S1A-035$B)

Examples such as those in (45) show that although LIKE clearly highlights the following NP, the NP is not introduced by a definite but by an indefinite article, or even none at all. Hence, it is inappropriate to restrict focusing LIKE to NPs which are introduced by definite articles.

In order to avoid misclassification of cases which could reasonably be focusing although the modified NP is introduced by indefinite articles, the classification of focusing LIKE is based on directly linguistic contextual cues. Such clues may be additional intensifiers such as *very*, *really*, *totally*, *extremely*, etc., as in (46), or other linguistic elements which clearly indicate that the instance of LIKE in question served a highlighting function.

(46) a. Ngi <,> it's *like* very old. (ICE Philippines:S1A-078$B)

b. If the roof isn't good <,> that's *like* a very important part of any house you buy. (ICE Ireland:S1A-035$B)

As a final remark on focusing LIKE, as argued by Fleischman and Yagello (2004:132), it should be mentioned that there appears to be a division of labor between LIKE and intensifying adverbs: while LIKE marks focus, the intensifying adverbs serve to convey emphasis. Despite this, the co-occurrence

of LIKE with intensifying adverbs strongly indicates that LIKE in such contexts marks focus and is not used as a hedging device.

4.7.1.5 Clause-final LIKE

The 'traditional' clause-final use of LIKE (Andersen 2000:216) as in (47) appears to be simply another variant of the discourse marker LIKE, on par with the clause-initial or clause-medial instances discussed above. This is, however, not the case. In fact, clause-final LIKE exhibits distinct characteristics which set it apart from the instances of LIKE discussed so far. With respect to its regional or geographic distribution, clause-final LIKE has been associated mainly with the British Isles and is referred to as a typically "northern" phenomenon (Hedevind 1967:237; cf. also Miller & Weinert 1995:368).

(47) a. No how do you explain it *like*. (ICE Ireland:S1A-063$C)
b. But you can't come to the table <,> as in speaking to people at the table and meanwhile somebody's out putting a car bomb under their motor *like*. (ICE-Ireland:S2B-025$I)
c. My dad in chinos *like*. (ICE-Ireland:S1A-051$E)

Despite being superficially similar to the so-called 'more innovative American form' (Andersen 2000:216), it has distinctly different ancestral roots. The difference in origin surfaces in its 'reversed' direction of scope. Contrary to instances of LIKE which have forward scope, i.e. scope over the following constituent, clause-final LIKE has backward scope, hence, modifying the preceding word, phrase, or sentence. Jespersen (1954) accounted for this by suggesting that clause-final LIKE is a variant or later development of the suffix *-like* which was originally attached to adjectives or adverbs as a second component. Gradually the suffix became independent and "may now be added to any sb [substantive] and is frequently added to adj [adjectives]" (Jespersen 1954:417) with a meaning of 'similar to' or 'characteristic of'. Only later did this form evolve into the present day clause-final discourse marker which is associated with colloquial, dialectal and vulgar language. In this understanding, clause-final LIKE evolved from the suffix *-like* and not from the comparative preposition as its American cousin did. In this sense, it is not sufficient to speak

of "two distinct traditions concerning the use of *like* as a pragmatic marker, one of which stems from (rural) dialects of Britain. The other […] appears to be a relatively recent borrowing from American English" (Andersen 2000:216). It appears more appropriate to speak of two superficially similar forms, a comparative preposition, on the one hand, and a suffix, on the other, which perform equivalent functions, despite having quite distinct developmental origins.

On a functional level, Jespersen remarks that clause-final LIKE is used parenthetically by inferiors addressing superiors "to modify the whole of one's statement, a word or phrase, modestly indicating that one's choice of words was not, perhaps, quite felicitious" (1954:417). Jespersen's (1954) description ties in with other comments on LIKE use which predominantly depict LIKE as a marker of vagueness. Partridge (1984), for example, asserts that clause-final LIKE is "expressive of vagueness or after-thought modification" (1984:264), thus functioning as a hedging device which may be glossed as 'as it were', 'not altogether' or 'in a way'. According to this view, LIKE is used to qualify a preceding statement or to express afterthought, and can be glossed 'as it were' or 'so to speak' (OED 1989 VIII:946).

Later accounts more often than not reject the claim that clause-final LIKE serves a predominantly hedging function. Miller and Weinert (1995:388-390), for instance, argue that clause-final LIKE "does have some retroactive focusing power, but more importantly, it […] is 'clearing up misunderstanding'" (Miller & Weinert 1995:388-389), and can be interpreted as countering potential inferences, objections or doubts. Miller (2009) elaborates on this interpretation, stating that clause-final uses of LIKE "serve to mark argument and counterargument in formal discussion as well as private conversation" (2009:335), asserting further that such occurrences index particularly relevant and 'clinching' arguments.

Fleischman and Yaguello (2004:132) offer a different interpretation of clause-final LIKE. While Fleischman and Yaguello (2004) agree with Miller and

Weinert (1995) that clause-final LIKE has scope over the preceding segment, they propose that it is not used to counter objections; it disambiguates the scope of the highlighted segment so that the listener is guided to the interpretation that the adjectives *sad* and *quiet* in (48a) as well as *staunch* and *bad* in (48b) and not the intensifier *really* are the highlighted segments.

(48) a. [S]he's really sad and quiet LIKE. (example from Fleischman and Yaguello 2004:132)

b. Aye but her Granny's really really staunch *like* <#> And uh she wouldn't she found out what religion the nurses were and all and wouldn't let the Catholic ones come near her and all this sort of stuff. (ICE Ireland:S1A-005$C)

c. [B]ecause um i mean yeah people leave these places *like*. (ICE New Zealand:S1A-031$A)

d. Well like that would be really bad *like* <ICE-IRE:S1A-039$A> Mm. (ICE Ireland:S1A-039$B)

While it is plausible that clause-final LIKE is modifying and in the majority of cases probably highlighting the preceding segment, Fleischman and Yaguello's argument that it is used to disambiguate scope is problematic for three reasons: firstly, it is not convincing that the speaker in (48a) should highlight *really*. Secondly, the vast majority of instances of clause-final LIKE do not co-occur with intensifying adverbs such as *really*, *very*, etc. Thirdly, an alternative and probably simpler explanation is that whenever phrase-initial LIKE and phrase or clause-final LIKE co-occur as in (48a), they serve distinct functions. In fact, the first instance of LIKE in (48a) occurs in a lexicalized construction and would – at least in the present analysis – not be considered an instance of the discourse marker LIKE proper.

The most elaborate analysis of clause-final LIKE todate is provided by Columbus (2009). The results of her analysis of IrE data lead her to conclude that utterance-final LIKE is best described as being an invariant tag on par with tag questions such as *eh* or *right*. According to Columbus (2009), such instances of LIKE clearly mark focus, but they are performing functions similar to other invariant tags such as serving to mark emphasis. Hence, this discourse-analytic approach highlights the multifunctionality of this form and challenges accounts

which subsume all instances of clause-final LIKE into a single functional category.

4.7.1.6 Quotative *be like*

Of all variants of LIKE, quotative LIKE has attracted the most attention (e.g. Blyth, Recktenwald & Wang 1990; Buchstaller 2001, 2002; Ferrara & Bell 1995; Meehan 1991; Romaine & Lange 1991; Schourup 1982, 1985; Tagliamonte & Hudson 1999). In contrast to other uses of *like*, quotative *be like* has a fixed syntactic slot as in (49) and, thus, is part of an already grammaticalized quotative construction.

(49) a. They were just you know there dancing and I was *like* "what the hell is going on with these people". (ICE Phippines:S1A-007$B)
b. [I] was just *like* "oh cool". (ICE New Zealand:S1A-096$A)
c. [W]e were there and I'm *like* "Hiti you really need to work on your jokes you know". (ICE Jamaica:S1A-031$A)
d. And I was *like* "oh my God". (Santa Barbara Corpus:SBC045$CORINNA)

This means that quotative LIKE lacks the feature of syntactic optionality and is, therefore, not regarded as a variant of the discourse marker LIKE.[21] Nevertheless, it is briefly touched upon, as it shares common ancestry with discourse marker LIKE and is similarly associated with the speech of adolescents (cf. Dailey-O'Cain 2000). It is claimed that, in contrast to other verbs of saying, quotative *be like* "allows the speaker to retain the vividness of direct speech and thought while preserving the pragmatic force of indirect speech" (Romaine & Lange 1991:228).

With respect to the origin, Romaine and Lange (1991) proposed a step-wise grammaticalization starting with comparative preposition, which in turn gave rise to the conjunction variant of *like* which is best paraphrased as *as if.* Meehan (1991) describes the same grammaticalization path, arguing that it originated

[21] This is probably too simplistic as instances of the quotative complementizer may co-occur with other quotatives such as GO (+ LIKE) + quote, in which case the instance of LIKE fulfils the optionality criterion.

from the conjunction, which provided this variant of *like* with its fixed syntactic slot; and then became semantically bleached; it now serves "to focus on the highlighted information expressed in the quote. In addition, *like* is still reflecting the old 'similar to' meaning since the information contained in the quote is not exact" (Meehan 1991:48). Buchstaller (2001) proposes an approach towards the grammaticalization of quotative LIKE which builds on Lakoff's (1987) radial structure model, replacing the channel or pathway model introduced by Traugott and Heine (1991).

4.7.2 Clause-external LIKE

In contrast to clause-internal instances of LIKE which modify lexical elements within its scope, clause-external LIKE has no modifying function, but either links sequences or indicates planning or processing difficulties (Andersen 2000:254). Such non-modifying instances of LIKE frequently co-occur with pauses, false starts, self repairs, incomplete or terminated utterances (Andersen 2000:254). As not all clause-external instances of LIKE are identical, the following section provides a typology of different variants of LIKE which can be classified as either linking devices or elements which indicate processing difficulty.

4.7.2.1 Linking and cohesion

Until discourse markers became in themselves a topic in linguistics, they were mostly studied as devices which function to create discourse cohesion, as in Halliday and Hasan (1976). In this context, like is regarded as a linking device, marking "the relationship between chunks of text, highlighting certain chunks/putting them into focus or signaling that the current chunk of text serves as exemplification or explanation of a previous chunk" (Miller 2009:320).

Instances of LIKE which establish cohesion or link utterances commonly do so by introducing specifications of what was said before and can best be

glossed as *that is* as in (50). This type of LIKE is frequently, though not necessarily, preceded by pauses which indicate that the speaker has completed the previous sequence, but then returns to the topic and elaborates on or specifies certain aspects of the utterance to disambiguate components which may be misinterpreted or seem too vague. The specifications vary in length, as they can basically consist of single words (50a), phrases (50b), complex constructions (50c) and whole sentences (50d), or even multiple utterances (50e).

(50) a Specification (single word)

He's improved a wee bit you know <#> *Like* physically (ICE Ireland:S1A-002$?)

b. Specification (phrase)

So we thought well we don't really want that covered up <,> but we need a bit of privacy so we thought we'd actually get a screen you know *like* a standing screen (ICE-Ireland:S1A-013$A)

c. Specification (complex construction)

High dependency unit <#> *Like* a step forward from ICU. (ICE Ireland:S1A-078$B)

d. Specification (sentence)

And uh it never even occurred to me that she was or wasn't was or wasn't making or doing the same as us because <,> there was no way she was a weirdo <#> *Like* she never felt left out and we never felt like the oh weirdo. (ICE-Ireland:S1A-011$B)

e. Specification (utterances)

Uhm you see this is why I don't want a night job in a pub <S1A-048$B> Yeah yeah <S1A-048$A> *Like* I could go home and study say from seven to nine every night and go out then at half nine <#> But then I'd be in late and that'd mean I'd be tired be wrecked the next day. (ICE Ireland:S1A-048$A)

These uses of LIKE are well explained by Andersen (1998, 2000, 2001) who concludes that LIKE is employed to guide utterance interpretation by providing cues as to which components contain especially relevant information. In the present case, this means that LIKE has a twofold function. On the one hand, it serves as a marker of focal information (i.e. the string which is about to follow) while on the other hand, it functions as a linking device which informs the speaker that the segment that follows relates to what was said immediately

before. The linking of conceptually related utterances establishes its status as an extra-clausal instance of LIKE, while its focusing potential reflects a similarity with clause-internal uses. But despite sharing features with clause-internal uses of LIKE, the fact that it marks the beginning of a new sequence or chunk clearly sets it apart from purely clause-internal uses.

4.7.2.2 LIKE as a hesitation marker and repair indicator

Some instances are, nevertheless, most accurately described as mere hesitation phenomena – for instance, when a speaker is trying to find the most appropriate expression and after succeeding in finding it proceeds with a syntactically related sequence, as in (51). Nonetheless, it would be naïve to assume that LIKE is not performing any communicative function when used in this fashion. Indeed, instances which seem to be mere pause fillers are more plausibly functioning as floor-holding devices, a way of indexing that the speaker is about to proceed and wants to hold the turn. Hence, in particular instances, such as (51), appear to be way to buy processing time which allow speakers to briefly interrupt the utterance without losing the turn, as the utterance proceeds without syntactic loss. In contrast to instances in which LIKE indicates the beginning of a new, syntactically unrelated sequence or chunk, such hesitation indexing uses of LIKE are characterized by a perfectly complete and unimpaired syntactic string as in (51).

(51) But you know if I uh having a sore eyes is really uh bad for me you know because uh just *like* uh I was experiencing life of Randy Santiago remember? (ICE Philippines:S1A-069$D)

Schourup (1982), and more elaborately Siegel (2002:41), point out striking parallels between such instances of LIKE and interjections such as *oh* and *um*. Like other interjections, instances of LIKE, which function as fillers, occur "with the greatest frequency in positions of great lexical indecision such as a) preclausally, but after prefatory material; b) before filled or unfilled pauses; and c) before restarts" (Siegel 2002:41). In contrast to clause-initial LIKE, these

occurrences do not introduce explanatory passages and, hence, do not establish coherence relations between larger chunks of discourse.

In addition to the positions described by Siegel (2002:41), LIKE also occurs adjacent to aborts as in (52). On a metatexual level, such instances function as cues to inform the listener that either the speaker is searching for the appropriate expression, or that the speaker is about to begin a somehow separate, new communicative act, as in (52), and that the preceding sequence is not relevant for the interpretation of what is about to follow. In this sense, LIKE indicates to the listener that he or she is supposed to focus on the following string, as the previous utterance will not be taken up again, and whatever may be 'stored in the working memory' may be deleted.

(52) a. So would you recommend that people who *like* <,> it's that uh they're taking the pill because they want to stop painful periods. (ICE Jamaica:S1A-054$A)

b. [W]ell you can but not *like* […] i suppose it's just an exception at the moment… (ICE New Zealand:S1A-038$P)

Along these lines, one can discriminate between various types of clause-external variants of LIKE according to their linguistic context and their communicative function. Parameters of this sort are, for example, the syntactic relatedness of the sequences which are separated by LIKE (compare (51) to (52)); the co-occurrence of LIKE with pauses as in (53a) and (53b); or repetitions of identical strings before and after LIKE without pausing, as in (53c).

(53) a. No we I don't say they are <,> *like* uh <,> they are for what you call for the tourist purpose only and we will go and watch. (ICE India:S1A-008$A)

b. [B]ut the point is i mean *like* er <,> […] yeah so i mean he hasn't got he hasn't got he's strapped for cash. (ICE New Zealand:S1A-016$R)

c. You *like* you wanna come back to summer school (ICE Jamaica:S1A-063$A)

Instances of LIKE displaying typical signs of lexical indecision have been linked to signaling intimacy. The association between intimacy and lexical indecision may be considered dialectic in the sense that intimacy signaling may also be construed as a by-product of lexical indecision. In this view, speakers

who feel comfortable in the communicative situation begin statements before having fully structured them and use LIKE as a means to buy further processing time for completing utterance planning. The interpretation of LIKE as a signal of intimacy or informality is challenged by Miller (2009). While Miller (2009) admits that LIKE can reflect solidarity among speakers, and may also serve as a processing time buyer, he stresses that this explanation fails to account for the vast majority of uses in which LIKE does not co-occur with hesitations and pauses, and also does not apply to clause-final LIKE (2009:324). According to Miller (2009), LIKE functions predominately on the textual level, e.g. as a non-contrastive focusing device which anticipates possible follow-up questions.

4.7.2.3 LIKE as part of general extenders and complex lexicalizations

Variationist research only recently has begun paying attention to *like* as part of general extenders (GE), such as (54), or lexicalized expressions (LE), as in (55).

(54) a. Not paradise or *something like that* (ICE Jamaica:S1A-001$A)
b. [P]eople with drawing boards and *stuff like that*. (ICE GB:S1A-034$B)
c. Do you remember that's where Gerry went and got altitude sickness or *something like that.* (ICE Ireland:S1A-003$E)

(55) a. *It's like* people who have a visa. (ICE Jamaica:S1A-012$A)
b. *[I]t's like* <„> even though I would they would have felt <,> uh mummy is not letting us play and all so long. (ICE India:S1A-030$B)
c. I think um *it's like* all the the ideas that we have throughout the ages and the various they're just sort of echos of uh the past. (Santa Barbara Corpus:sbc 017$Micha)

GEs commonly occur in utterance-final position and, similar to other discourse markers, are associated with vagueness as they direct "the hearer to access a set of which the given item is a member, whose characteristics will enable the hearer to identify the set" (Channel 1994:122). Hence, they "represent a distinct set of linguistic elements which have received little attention from linguists" (Overstreet 1999:3) as vague language has often been met with derision, especially among purists (Jucker, Smith & Lüdge 2003:1737). Recent studies in particular have employed elaborate quantitative

methods in the analysis of this phenomenon (cf. Pichler & Levey 2010; Tagliamonte & Denis 2010). With respect to their sociolinguistic distribution in Toronto English, Tagliamonte and Denis (2010:28) find that GEs are more frequent among adolescents as compared to adults. In addition, Tagliamonte and Denis (2010:28) draw attention to the variety-specificity of both their grammaticalization and their discourse-pragmatic functions.

Complex lexicalizations, on the other hand, are a subgroup of phrasal discourse markers which occur clause-initially. CLs like "I think" or "I guess" function as epistemic parentheticals and have grammaticalized from matrix clauses such as "I think that…" or "I guess that…" (Brinton & Traugott 2005:137-138). CLs in which *like* is a non-optional element, such as "It's like", differ from a construction which is entirely optional. Cases which have been excluded from the data share at least two features which set them apart from structurally identical true matrix clauses. Firstly, they are not subsided by a pronominal from so that they precede a fully independent matrix clause. And secondly, they carry an epistemic function, as they express the speaker's certainty of the truth of the subsequent utterance and, thereby, serve to qualify how sure the speaker is of what is to follow.

The use of *like* in GEs or CLs differs from its use as a discourse marker in so far as in the latter case *like* is not part of fixed constructions which were deemed ungrammatical when *like* was removed. In this sense, *like* is not syntactically optional in these environments and, hence, does not fall within the definition of discourse markers employed here.

4.8 Synopsis

The previous chapter has introduced the reader to the current state of research on LIKE. The chapter has recapitulated LIKE's diachronic development and dismissed the idea that LIKE has grammaticalized in parallel in various locations simultaneously and argued for an interpretation which assumes that

it has undergone three major waves of spread. During the first period of spread, LIKE diffused through various regional varieties on the British mainland. The second wave is characterized by the introduction of LIKE into various colonial varieties, most notably to AusE and NZW as well as North American varieties of English. In addition, LIKE became a functional element of IrE, where it had been introduced before, during its dialect formation. The last wave of spread took place during the last fifty or so years, when the spread of LIKE was promoted through AmE and its covert prestige and dialect contact in the wake of World War II. With respect to the attitudes associated with LIKE, this chapter showed that LIKE is typically overtly stigmatized and perceived as being an American element. Speakers using LIKE are perceived as cheerful, social, young, and female. The chapter also showed that the positioning of LIKE is not random, but that its positioning is syntactically constrained so that there are specific slots where LIKE can occur. The final section of the chapter described and discussed LIKE's discourse-pragmatic functions and showed that the positioning and the functioning of LIKE interact. In other words, LIKE fulfills different discourse-pragmatic functions when it is used in different clausal positions. For instance, in clause-initial position, LIKE serves as a linking device used to establish coherence relations; it is used to modify elements to its right, i.e. to focus or hedge lower level constructions such as phrases, when used in clause-medial position.

5 Data and methodology

The following section describes the data, its structure, the processing and editing stages, as well as the methodology applied in evaluation of the analysis of the discourse marker LIKE across and within varieties of English. The section consists of three parts: the first concerns itself with general issues, such as the outline of and remarks on the methodology, as well as stating the basic research questions. The second part elaborates more specifically on issues relating to the data sources and data processing, such as the inclusion or exclusion of data and their analyses. The third part discusses the statistical models used in the quantitative analyses.

5.1 Introduction

The aim of this chapter is to justify the present study and to provide a comprehensive and systematic account of issues relating to methodology and data editing.

The present study combines sociolinguistic methodology with dialectometric analysis, offering both a new approach towards the analysis of non-standard LIKE and displaying the need for combining inter- and intra-varietal perspectives on usage patterns. By approaching the discourse marker LIKE from a cross-varietal perspective, it serves to illustrate regionally bound characteristics of LIKE use, allowing re-evaluating claims and hypotheses of previous research which presumed that LIKE usage is rather similar across varieties of English. Although often not explicitly, this is problematic, particularly since the bulk of research on LIKE has focused either on North American English (Dailey-O'Cain 2000; D'Arcy 2005, 2007; Schourup 1982; Siegel 2002; Tagliamonte 2005; Underhill 1988), or varieties of English spoken across the British Isles (Andersen 1998, 2000, 2001; Miller & Weinert 1995). The underlying premise that findings in one variety can be generalized to other, geographically distinct varieties , is either premature or overly simplistic.

Unfortunately, it is problematic to generalize from one variety over a culturally and linguistically distinct variety of English because this bears the risk of dismissing variety-specific usage patterns. Avoiding such unwarranted over-generalizations, the present analysis touches on wider theoretical implications for the local sociolinguistic underpinning of linguistic supra-localization, i.e., re-occurring patterns in the diffusion and transmission of forms (cf. Buchstaller & D'Arcy 2009:293).

Quantitative approaches to sociolinguistic phenomena have been employed ever since Labov's pioneering work on Martha's Vineyard (Labov 1972) and the New York City study (Labov 1966), but have mainly been used to investigate phonological variation and to a lesser extent morphological and syntactic variation (Macaulay 2005:12). Only recently have researchers begun to apply statistics-based quantitative methodology to discourse features[22]; mostly, however, by focusing on gender or age differences (Macaulay 2005:13). Such large-scale quantitative methods enable meaningful comparison and repeatability as well as increased reliability, while fine-grained, qualitative analyses allow for a minute and detailed analysis of discourse features on a micro-level. Ideally, both methods are combined, thereby consolidating the productive aspects of both approaches. Furthermore, the combination of cross-varietal and within-variety analyses enables the retracing of the spread of LIKE throughout the English-speaking world, but also the investigation of which subpopulation has been responsible for its spread, i.e. who were the leaders of that change.

The classification of LIKE in the present study mainly builds on D'Arcy's (2005, 2007), Miller's (2009) and Andersen's (1998, 2000) previous analyses of this vernacular feature. The resulting synthetic categorization is then applied to occurrences of LIKE in a sample of geographically distinct varieties

[22] One notable exception is D'Arcy (2005), who employs a large-scale, statistics-based approach towards the use discourse marker LIKE in Toronto English. Indeed, this study provides a fine-grained analysis of its apparent-time distribution in this regional variety.

of English based on eight regional components of the *International Corpus of English* to enable a cross-varietal comparison.

5.2 What does local implementation entail for LIKE?

Before specifying the research questions and hypotheses, it is necessary to elaborate on what it means to say that an innovation has undergone modification during local implementation.

As stated previously, the main objective of this study is to analyse processes that accompany the local implementation of globally available linguistic innovations. Therefore, a closer look at what local implementation means seem to be in order. The following section will thus discuss this concept and what it entails for the present study.

In the present context, local implementation refers first of all to the frequency of vernacular LIKE within a given speech community. Following Nevalainen & Raumolin-Brunberg (2003:55), the percentage of speakers within a speech community that use an innovation allows inferences to be made about the degree to which an innovation has diffused through a speech community and, hence, how accepted and established it is. Given the limited size of data sets used here, this only applies to high frequency phenomena, since infrequent constructions would be too rare in small or even mid-sized datasets to allow conclusions to be drawn about how established such constructions are. However, LIKE is a high-frequency phenomenon, and the percentage of speakers who use it does allows conclusions to be drawn about its acceptance and use in a given speech community.

Second, local implementation refers to differences between the use in the donor or source variety and the host or target community. These differences can be realized on any mirco-linguistic level, be it phonetic, phonological, morphological, syntactical, semantic, or pragmatic. In the present case (cf. also section 5.4.5), effects of localization are considered to be positional deviations

in the target variety. The underlying assumption is that the positioning of LIKE affects the pragmatics and thus correlates with the function of LIKE. For instance, clause-medial LIKE serves to modify lower level constructions such as phrases to its right by hedging or focusing them (cf. section 4.7.1), while clause-initial LIKE serves to establish coherence relations between higher level constructions such as clauses and sentences typically by indicating that what is to follow is an exemplification of the content of the previous proposition. Therefore, if the positioning of LIKE in a given target variety differs from the positional distribution of LIKE in probable source varieties, this strongly indicated a difference in use and, more specifically, in functional employment of this vernacular form. The difference in positional distribution then indicates that LIKE has been adopted by the local speech community and modified to match their communicative needs.

Thirdly, LIKE serves as a social index by marking in-group membership or as a display of certain extra-linguistic attitudes. For instance, Dailey-O'Cain (2000) found that "young women are perceived as using *like* most often" (Dailey-O'Cain 2000:60). Remarkably, speakers were also rated more attractive, more successful, more cheerful, and more friendly when they used LIKE as compared to when they did not (Dailey-O'Cain 2000:73). However, if speakers used LIKE, they were also perceived to be less educated, less intelligent and with respect to older speakers less interesting (Dailey-O'Cain 2000:73). The findings of Dailey-O'Cain (2000) validate the hypothesis that LIKE serves as a social index and is associated not only with certain reference groups such as young females, but also with certain personality traits. It is likely that the association of LIKE with certain personality traits is secondary and builds on its role as a marker of group membership, because the personality traits appear to be related to subjective generalizations about features of the social reference group. To elaborate, young females are more generally considered attractive, cheerful, and friendly, and LIKE's association with these personality traits is thus based on its association with this subsection of society rather than the other way around. In other words, it less

likely that it is primarily its association with the favourable personality traits that make LIKE an attractive element for young females. The reason for why the social index is primary is the favourable argument lies in the fact that the opposed view fails to answer why LIKE is associated with the respective personality traits in the first place. However, this is not to say that young females use it because it is associated with these personality traits – but the association with these traits is not primary but secondary. Differences in the social profile of LIKE use would thus strongly suggest that LIKE use in the target variety indexes different group memberships and entails different personality traits than it does in the source variety. Again, this would strongly support the assumption that the extra-linguistic functions of LIKE have undergone modifications and have been adopted during local implementation.

The preceeding subsection has elaborated on the notion of local implementation and defined the ways in which differences in LIKE's frequency, its positional distribution, and its social profile would indicate modifications during local implementation and why these differences between probable source varieties, specifically AmE but also to a lesser extent EngE, suggest that LIKE has been modified by the local speech community.

5.3 Research questions

In essence, the present investigation of the discourse marker LIKE focuses on linguistic supra-localization, i.e. the local implementation of globally available innovative forms. The overarching objective is to evaluate the stability of the Labovian model of the social motivation of linguistic change in culturally diverse and multilingual settings. Therefore, the present study aims at testing the consistency of correlations between extra-linguistic variables and LIKE use in diverse contact scenarios. Hence, this analysis elaborates specifically on previous research, such as Buchstaller (2008), Buchstaller and D'Arcy (2009), as well as Meyerhoff and Niedzielski (2003).

The research questions addressed here are in line with the research questions stated in Buchstaller and D'Arcy (2009:293):

- Who uses these innovations in different localities?
- Are the distributional and perceptual loads of travelling resources constant across geographic space?
- What social and linguistic functions are constant across geographic space?

The latter question is particularly relevant, as the vast majority of sociolinguistic studies have dealt with monolingual settings and thus neglect dialect contact or substrate/superstrate interference as relevant factors (Labov 2001:518). In contrast, the present analysis attempts to shed light on the implementation of supra-local innovations in diverse, often multilingual and geographically discontinuous speech communities (cf. Buchstaller 2008; Buchstaller & D'Arcy 2009; Meyerhoff & Niedzielski 2003). Thus, this study offers a more profound understanding of the adaptation of features in case quality face-to-face contact, which is the prototypical scenario in most sociolinguistic analyses, is missing. Thereby, it elaborates on alternative methods of transmission, such as the role of mass media, which has for the most part been disregarded by sociolinguists (Hickey 2003:341).

Hence, this investigation concerns itself with the issue of whether the diffusion of LIKE follows similar trajectories in geographically distinct locales. If this is true, it would suggest that LIKE is merely borrowed from a parent to a receptor variety. If, on the other hand, LIKE shows variety-specific adaptation during its transmission, this would suggest that the associations attached to LIKE are modified to local norms and practices. This would corroborate recent claims that the adoption of globally available variants involves proactive behavior on the side of the local speech community (cf. Kachru 1992; Buchstaller 2008; Buchstaller & D'Arcy 2009, Meyerhoff & Niedzielski 2003). Accordingly, the present investigation addresses issues such as the status of LIKE as an American borrowing (Andersen 2001:25); whether we observe

systematic variation in LIKE use among the regional varieties of English; and how we can account for such patterns in case they exist? In addition, the analysis aims to answer questions relating to the variety-specificity of LIKE use, for example, whether certain varieties favor particular variants of this marker, or whether all variants of LIKE are subject to similar trajectories of change. These questions ultimately aim at evaluating the degree to which extra-linguistic factors have influenced LIKE use in quite distinct and diverse varieties of English.

5.4 Central hypotheses

In view of the above considerations, and based on both contemporary sociolinguistic theory, and previous research on LIKE, the present investigation aims to test five central hypotheses. The first two address claims stated in previous studies on LIKE use and thus relate to its sociolinguistic profile. The latter three hypotheses relate to the theoretical framework more generally and aim to evaluate widespread sociolinguistic assumptions and mechanisms of ongoing linguistic change. The following section will present, describe and discuss each hypothesis in detail.

5.4.1 Hypothesis 1: LIKE is a marker of teenage speech

The following subsection states the first hypothesis and will describe and discuss this hypothesis with respect to previous claims about LIKE use.

Hypothesis I:

LIKE is most common among adolescents and young adults.

The speech of teenagers and young adolescents in contemporary, urban speech communities represents a "tremendous breeding ground for linguistic innovation. […] The language of these speakers is a gold mine for innovative linguistic features, revealing evidence for both grammatical, as well as

sociolinguistic change" Tagliamonte 2005:1913). The relevant literature on LIKE strongly supports the proposal that this also holds for LIKE (e.g. Andersen 1998, 2000; D'Arcy 2005, 2007, 2008; Siegel 2002; Tagliamonte 2005; Underhill 1988). Accordingly, this hypothesis aims to evaluate to what extent LIKE is a marker of teenage speech, not only in standard varieties of English but across varieties of English more generally. In fact, this hypothsis tests the claim that LIKE is an angloversal marker of teenage speech and evaluates the factors that contribute to mark it either as typically teenage-speech-related, or as being a pragmatic marker employed by other social strata or even the population at large. If LIKE is indeed confined to the speech of adolescents and younger adults, then this would confirm previous analysis and strongly suggest that LIKE is not only functional from a language-internal perspective but that it also possesses an extra-linguistic function as a social index. If LIKE is not confined to the speech of adolescents and younger adults, however, then this would suggest stabilization, i.e. that LIKE has already diffused through the entire speech community, and would indicate that LIKE's functionality is predominantly language-internal – in this case carrying various discourse-pragmatic functions.

5.4.2 Hypothesis 2: LIKE as a marker of female speech

The following subsection states the second hypothesis and justifies this hypothesis with respect to previous claims about gendered LIKE use.

Hypothesis II

LIKE is typically a female feature and thus gendered.

Although the relevant literature on LIKE is inconsistent with respect to the existence and degree of gender differences of LIKE use (cf. section 5.8.3), the female lead in cases of vigorous and incipient change is one of the most robust aspects of ongoing change. Labov (2002), for instance, asserts that "in the great majority of cases, it is women who are ahead – usually by a full generation" (2002). Accordingly, the third hypothesis aims to evaluate the extent to which

the mechanisms that have been observed for ongoing change apply to the spread of LIKE. In addition, this hypothesis serves to shed light on whether and where LIKE is more common amoung female speakers. Similar to hypothesis 1, this second hypothesis targets the functionality of LIKE – if LIKE is used by speakers with a certain social profile, this would support the view that LIKE is a social index and not confined to its intra-linguistic, pragmatic functionality.

5.4.3 Hypothesis 3: The universality of the Labovian model

The following subsection states the third hypothesis and describes this hypothesis in light of current sociolinguistic theories and the broader outlook of the present study.

The third hypothesis aims to evaluate the universality of the Labovian model of change from below on all mircolinguistic levels. In other words, it tests the extent to which mechanisms which have been observed in cases of phonological change from below are also at work in cases of pragmatic change.

Hypothesis III: The Labovian model of social motivation for linguistic change is valid with regard to pragmatic change.

This hypothesis implies that recurring patterns attested during the incrementation and diffusion of phonological innovations such as a female lead, social stratification, and age-grading in apparent-time are not confined to phonological change, but also features of pragmatic change and thus the spread of pragmatic innovations. In essence, the main objective of this hypothesis is to test how fitting the Labovian model of change from below is in cases pragmatic rather than phonological innovations such as the spread of LIKE. The third hypothesis relates to the broader theoretical framework of this study and serves to empirically determine to what extent the Labovian model requires modification when applied to the diffusion of pragmatic innovations. Although this objective seems straightforward, there are so far no systematic studies that evaluate the adequacy of the Labovian model of change from below, although it is frequently taken as a reference point in cases of ongoing change (cf. e.g.

D'Arcy 2005, 2007; Tagliamonte 2005). Therefore, testing this hypothesis has repercussions not only on LIKE use, but on sociolinguistic theorizing more generally – at least when analyses go beyond the study of phonological phenomena.

5.4.4 Hypothesis 4: Diffusion and stratification

The following subsection states and discribes the fourth hypothesis. Similar to the third, this fourth hypothesis also addresses a theoretical issue and aims to evaluate basic assumptions underlying modern sociolinguistic theories.

> Hypothesis IV: LIKE use is socially stratified, particularly in locales where it has only recently been introduced. In other words, the higher the frequency of LIKE, the lower the degree of social stratification.

According to Labov (2002), "most of the linguistic changes in progress studied in the 2nd half of the 20th century show a high degree of social stratification." Considering that social stratification wanes once a change is nearing completion, then this implies not only that LIKE use is socially stratified, but also that the degree of social stratification depends on the stage of change. The closer a change is to completion, the less social stratification is expected. The complementary implication is, of course, that the more vigorous a change is, the more social stratification is expected.

If LIKE indeed showed less social stratification in varieties where it is well attested and into which LIKE thus cannot have entered recently, this would support the view that social stratification is a valid index for ongoing change. If LIKE showed significant stratification in locales where it is well attested and which LIKE thus cannot have entered recently, then this could mean two things. First, that LIKE is undergoing another form of change in which a previously language-internal functional element has acquired an extra-linguistic function as a social index. This interpretation would, however, be somewhat unusual, as social marking is typically a feature of innovations or elements which have not

fully diffused through the speech community in the first palce. Second, it could mean that the diffusion of morpho-syntactic and pragmatic elements is generally accompanied by less social stratification and the existence of social stratification would therefore be of less value for detecting ongoing change for such elements. This latter point would have notable implications for sociolinguistic theory and the practice of sociolinguistic research, which builds heavily on the apparent-time construct and the generalizing from phonological processes of language change to other elements on other linguistic levels.

5.4.5 Hypothesis 5: LIKE use is modified during local implementation

The last hypothesis is perhaps most crucial with respect to this study. The following subsection states and justifies this hypothesis in light of the broader outlook of this study.

Hypothesis V: The regional locale and the cultural norms and practices shape the local implementation of LIKE.

This hypothesis, in essence, tests the assertion put forth in previous studies on the local implementation of globally available variants that (linguistic) innovations are not adopted wholesale, but that they undergo modification and re-interpretation during their implementation (cf. Kachru 1992; Buchstaller 2008; Buchstaller & D'Arcy 2009, Meyerhoff & Niedzielski 2003). The use of LIKE in geographically distinct locales should, therefore, exhibit variety-specific idiosyncrasies. More specifically, this hypothesis evaluates the degree to which the social and positonal profile of LIKE in recipient varieties mirrors the social and positonal profile in the target or donor variety. With respect to positioning, chapter 4.7 has shown that the position of LIKE with respect to the clause strongly correlates with distinct discourse-pragmatic functions. If a form undergoes functional modification during its implementation in a given speech community, this would be expressed in a difference between the positional profile of the donor and the recipient variety. If the social meaning

of LIKE is modified and LIKE becomes associated, for instance, with a reference group distinct from its reference group in the donor variety, then this would cause a difference between the social profiles of the donor and the recipient variety. Any difference between the use of LIKE in, for instance, AmE and PhiE or JamE would therefore indicate local modification and support the view that innovations are not adopted wholesale but that they are shaped during implementation to match local norms and needs. If the positioning of LIKE in the recipient variety is very similar to the positioning in the donor variety but the social profile is distinctly different, this would indicate that the social meaning of LIKE is (cognitively) more salient than its linguistic constraints and its language-internal functionality. If, however, the social profile of LIKE in the recipient variety is very similar to the social profile in the donor variety but the positioning is distinctly different, this would indicate that the language-internal functionality of LIKE is (cognitively) more salient than its social function. In any case, the degree of variability between varieties in terms of both positioning and social meaning will further our understanding of processes that accompany the global diffusion of pragmatic innovations.

5.5 Data sources

The following section will discuss the data on which the following analyses build. It will specifically introduce the design of the ICE corpora and discuss why they are considered an adequate data source for this study. The subsequent sections will then continue to discuss matters of data processing and data editing, such as which instances of LIKE were excluded from further analysis and why they had to be excluded.

5.5.1 Introduction

Although previous research has broadened our understanding of various issues relating to the discourse marker LIKE, neither large-scale cross-varietal

surveys of this marker nor variety-specific constraints on LIKE use have been established. As systematic accounts of the variation of LIKE usage across varieties of English are few, the present study dedicates itself to filling this. In particular, this study addresses basic questions concerning variety-specific usage patterns of discourse marker LIKE. This is relevant, as few sociolinguistic studies have focused on the transmission and diffusion of features across wider geographical areas. Buchstaller and D'Arcy (2009:298), for instance, emphasize in their cross-varietal investigation of *be like*, which is very similar to the present study in its theoretical outlook and methodology, that…

> the direct comparability […] of previous analyses of verbs of quotation in general and of *be like* in particular remains questionable since they tend to be based on dissimilar methodological premises and applications in terms of the definition of the variable and constraints, the form selected as part of the envelope of variation, quantitative methods, and statistical models.

To overcome the limitations of these previous studies, the present investigation uses a computationally edited version of the ICE and employs coherent methodology to guarantee cross-varietal comparability of the data. Hence, the present quantitative analysis of LIKE across varieties of English draws its data from the *International Corpus of English* (ICE). This family of corpora represents distinct regional varieties of English and, therefore, forms an appropriate starting point for a cross-varietal study.

The issue of data selection is not trivial, as Buchstaller and D'Arcy (2009) emphasize that "[w]hat is needed, therefore, are reliable and comparable methods applied rigorously and uniformly across datasets to uncover which constraints hold both across and within varieties of English worldwide" (2009:298). Therefore, the present investigation uses a computationally edited version of the ICE and employs coherent methodology to guarantee cross-varietal comparability of the data. This family of corpora represents distinct regional varieties of English and, therefore, forms an appropriate starting point for a cross-varietal study.

5.5.2 The ICE family of corpora

The ICE family of corpora meet the criteria mentioned above, thus allowing a balanced and extensive overview of LIKE use across regional varieties of English to be attained. Seven of the available components of the ICE are considered: *ICE Canada* (Canadian English); *ICE GB* (British English; R1); *ICE Ireland* (Hiberno-English); *ICE India* (Indian English); *ICE Jamaica* (Jamaican English); *ICE Philippines* (Philippine English); and *ICE New Zealand* (New Zealand English)[23]. In addition, the *Santa Barbara Corpus of Spoken AmE* was added to the analysis for three reasons. Firstly, the *Santa Barbara Corpus* contains AmE data and, hence, broadens the range of regional varieties. Secondly, the *Santa Barbara Corpus* matches the other ICE components, as it "forms part of the International Corpus of English (ICE). The Santa Barbara Corpus represents the main data of the American component of ICE[24] (http: //www.linguistics.ucsb.edu/research/sbcorpus. html; April 4th 2010). Thirdly, the data included in the *Santa Barbara Corpus* consists mainly of private dialogue, i.e. face-to-face conversation, and thus contains the data particularly relevant for the present purpose.

[23] ICE Hong Kong and ICE Singapore could not be analysed, as the respective ICE teams did not grant access to the speaker information which is crucial. ICE East Africa, on the other hand, was not included as it deviates significantly from the other components in terms of both size (the spoken part of ICE East Africa consists of 714,916 words compared to an average of the other ICE corpora of 651,822 words) and annotation.

[24] The comparability of the Santa Barbara Corpus also stems from the fact that "[i]n order to meet the specific design specifications of the International Corpus of English (allowing comparison between American and other national varieties of English), the Santa Barbara Corpus data have been supplemented by additional materials in certain genres (e.g. read speech), filling out the American component of ICE" (http: //www.linguistics.ucsb.edu/research/sbcorpus.html; April 4th 2010).

The most prominent feature rendering the ICE family of corpora relevant for this investigation is that they were designed for comparability[25] and, hence, enable contrastive analyses of geographic varieties.[26]

All regional components of the ICE corpora contain 500 files, 300 of which represent spoken language of various text types and 200 of which represent written language of various text types. The transcribed dialogues are classified according to the type of spoken discourse (private or public dialogue, scripted or unscripted dialogue).[27] Therefore, the ICE corpora allow for a detailed analysis of the occurrence of LIKE in specific communicative situations. Beyond enabling large-scale, quantitative analyses, the ICE family of corpora also enables specialized and fine-grained exploration, as the linguistic context is available and data are not confined to specific communicative situations. Since discourse markers are most prevalent in informal spoken text types, the study will use data from files with the header S1A exclusively which indicates that the data represents private dialogues and more specifically, either face-to-face conversations or transcripts of telephone calls.

25 The ICE family of corpora was compiled during the 1990s or later, and the authors and speakers of the texts are aged 18 or above; they were educated through the medium of English, and were either born in the country in whose corpus they are included, or moved there at an early age; and they received their education through the medium of English in the country concerned (http: //ice-corpora.net/ice/design.htm, 29.4.2010).

26 "To ensure compatibility across the individual corpora in ICE, each team is following a common corpus design, as well as common schemes for textual and grammatical annotation. Each component corpus contains 500 texts of approximately 2,000 words each - a total of approximately one million words. Some of the texts are composite, made up of two or more samples of the same type" (http: //ice-corpora.net/ice/design.htm, 29.4.2010).

27 Each regional ICE component contains 120 monologues, of which 70 are unscripted (20 spontaneous commentaries, 30 unscripted Speeches, 10 demonstrations and 10 legal presentations) and 50 are scripted (20 Broadcast News, 20 Broadcast Talks and 10 non-broadcast Talks).

Table 3: Common design of the ICE components

Mode	Conversation type	Register	Text type	Number of text files
SPOKEN (300)	Dialogues (180)	Private (100) header=S1A	Face-to-face conversations	90
			Phonecalls	10
		Public (80) header=S1B	Classroom Lessons	20
			Broadcast Discussions	20
			Broadcast Interviews	10
			Parliamentary Debates	10
			Legal cross-examinations	10
			Business Transactions	10
	Monologues (120)	Unscripted (70) header=S2A	Spontaneous commentaries	20
			Unscripted Speeches	30
			Demonstrations	10
			Legal Presentations	10
		Scripted (50) header=S2B	Broadcast News	20
			Broadcast Talks	20
			Non-broadcast Talks	10

Beyond enabling large-scale, quantitative analyses, the ICE family of corpora also enables specialized and fine-grained exploration, as the linguistic context is available and data are not confined to specific communicative situations.

Adding to their advantageous design, the ICE corpora offer extensive information about the speakers: the age and gender of speakers to their level of education, occupation, 1st and 2nd languages, and ethnicity. Based on this additional speaker information, the ICE family offers the opportunity for fine-grained sociolinguistic analyses.

Despite their high degree of comparability, their wide range of registers and the detailed speaker information contained in them, the ICE have some shortcomings to be addressed in the present context.

With the exception of the *Santa Barbara Corpus*, neither of the ICE components contain detailed information about the phonological features of the original spoken data, nor do they offer the original audio files as complementary sources for data coding. This additional material would have been enormously helpful during the coding process: for example, coding of intonation would have permitted a more precise analysis of LIKE usage with regard to its position relative to the embedding intonation unit.

Furthermore, the ICE components are not representatively balanced with respect to the age, gender, and educational background of the speakers:

> The corpus contains samples of speech and writing by both males and females, and it includes a wide range of age groups. The proportions, however, are not representative of the proportions in the population as a whole: women are not equally represented in professions such as politics and law, and so do not produce equal amounts of discourse in these fields. Similarly, various age groups are not equally represented among students or academic authors. (http://ice-corpora.net/ice/design.htm, 29.4.2010).

This imbalance turned out to be a major obstacle for this study and ultimately led to the decision to focus exclusively on data containing the most informal spoken dialogue, i.e. private face-to-face conversations and private phone calls (files headed S1A) while disregarding more formal spoken data (files headed S1B, S2A and S2B).

Another disadvantage of the ICE family of corpora relates to its aim to represent the "national or regional variety of English", which is to say that the respective ICE components represent the national or regional standard varieties rather than representing a random sample of entire scope of language use in the respective region. Since the national standard is predominantly spoken by the educated elite of a country, the ICE corpora fail to reflect the language use of the population of English speaking individuals at large, but have a rather substantial bias towards educated speakers, i.e. they reflect the language of the elite rather than the speech community. Although this is a

drawback, the sample of speakers included in the respective ICE components was varied enough to allow testing for social stratification, because the speakers included in the data were biased towards higher education but not limited to speakers of such a profile.

5.6 Data processing

The frequency-based approach taken in the present analysis deploys the normalized frequency of LIKE instances per 1,000 words per speaker and, hence, requires the extraction of the precise word counts for each speaker present in the data. To perform this task, I devised R scripts (http://www.martinschweinberger .de/blog/codes-tutorials/) which serve both to extract the precise word counts for each speaker and combine these word counts with the biographical and sociological background information of the speakers. The resulting tables hold both word counts and biodata and constitute the basis for further speaker- and frequency-based analysis of LIKE use within and across varieties of English.

Subsequently, speakers who uttered fewer than 100 words had to be removed, as they would have skewed the data as even single uses of LIKE would have resulted in disproportionately high frequencies per 1,000 words. A related reason for excluding speakers who uttered fewer than 100 words is that their inclusion would, in fact, have led to very imprecise estimates of their linguistic behaviour, resulting in far less reliable statistical results. For an overview of the final data see Table 4.

Table 4: Overview oft he final data set for this study

	Speakers (N)	Speakers (words)	LIKE (N)	LIKE (mean ptw)
Canada	244	194,,574	900	4.625
GB	320	201,372	127	0.631
India	236	211,646	331	1.564
Ireland	309	189,787	936	4.932
Jamaica	228	207,807	526	2.531
New Zealand	227	229,193	529	2.308
Philippines	198	193,077	452	2.341
US	163	246,258	860	3.492
Total	1,925	1,673,714	4,661	2.785
	Female (N)	**Female (words)**	**LIKE (N females)**	**LIKE (mean ptw females)**
Canada	155	125,120	616	4.923
GB	156	101,620	61	0.600
India	115	98,563	214	2.171
Ireland	236	150,684	748	4.964
Jamaica	171	149,056	409	2.744
New Zealand	140	142,771	338	2.367
Philippines	141	132,441	352	2.658
US	91	138,888	539	3.881
Total	1205	1,039,143	3277	3.154
	Male (N)	**Male (words)**	**LIKE (N males)**	**LIKE (mean ptw males)**
Canada	89	69,454	284	4.089
GB	150	92,352	53	0.574
India	114	109,137	115	1.054
Ireland	73	39,103	188	4.808
Jamaica	56	58,650	117	1.995
New Zealand	87	86,422	191	2.210
Philippines	57	60,636	100	1.649
US	71	106,227	312	2.937
Total	697	621,981	1,360	2.187

Computing the token counts of all individual speakers represents a valuable extension of the ICE components, as it facilitates the use of the ICE for fine-grained sociolinguistic analyses across an extensive sample of geographically distinct varieties of English.

Hence, in contrast to most sociolinguistic studies on linguistic change and variation, the present study focuses exclusively on the use of LIKE. This may bring into question whether the present study can be called a study of linguistic variation proper, as alternative variants in the respective syntactic environment are not considered. This is crucial with respect to the discourse-pragmatic adaptation of LIKE during its implementation in local speech systems. Without taking competing or rival variants into account, the specific functional niches filled by LIKE remain unidentified. Indeed, the variety specificity stemming from interaction and competition with other linguistic means is mostly disregarded. In contrast to fine-grained discourse-pragmatic analyses which are better suited to addressing such variety-specific adaptations, this analysis assumes an underlying functional similarity allowing cross-varietal comparison. The purpose of the present investigation is not to carry out a comparative analysis of the specific discourse-pragmatic functions of LIKE, but to investigate the stability of re-occurring patterns in sociolinguistic diffusion and the degree to which these erode when applied to diverse local settings.

5.7 Data editing

For the empirical and systematic study of linguistic phenomena, relying on corpora has become vital. Nonetheless, corpora have deficiencies, as they cannot display all information contained in the original data – i.e. the factual discourse itself, or fine-grained phonetic and phonological properties. To compensate for the inherent limitations of data contained in corpora, it is necessary to consider all levels of discourse "such as phonetics, prosody, context and topic […] where the grammatical analysis arising from a mere

browsing of computer lists of examples will not suffice" (Andersen 1997:39). 'Browsing computer lists' is certainly valid for providing an informative approximation, further multilevel analysis is required to capture a more detailed picture. Unfortunately, the ICE data are not accompanied by audio files of the original communicative situations. Therefore, the classification of various instances of LIKE in the present case relies heavily on morpho-syntactic features, and has also been informed to a large extent by meta-linguistic annotation and commentaries included in the transcriptions, i.e. the presence of pauses or utterance boundaries.

Accordingly, but also for reasons of efficiency, concordancing software (MonoConc Pro 2.2) was used in this study to automatize the search for the relevant target forms, i.e. the orthographic sequence *like*. Unfortunatelty, the orthographic sequence *like* is by no means restricted to discourse marker uses because *like* also functions as a verb, a comparative preposition, a noun, an adverb, among others.

Exemplifications of discourse marker uses of the word *like* occurring in natural language data are given in (56).

(56) a. *Like* every time we spend a decent amount of time together i think i'm so happy. (ICE New Zealand:S1A-055$A)

b. No the one where they were uhm they were *like* worshipping that golden cow or something that they have made. (ICE Philippines:S1A-007$B)

c. That's amazing *like*. (ICE Ireland:S1A-036$A)

d. I mean I love American crap especially comedies *like* crap comedies that everybody thinks are crap. (ICE Great Britain:S1A-041$A)

The discourse marker LIKE differs from instances of orthographically and phonologically equivalent standard uses of *like*. Such standard uses comprise uses of *like* as a verb, as a noun, as a comparative preposition, as an adverb, and as an element of general extenders and lexicalized forms as in (57f). The difference between standard uses and uses of *like* as a discourse marker is that the latter is (i) grammatically optional and thus does not change the semantic relationships between elements (Fuller 2003; Schiffrin 1987; Schourup 1999); (ii) semantically bleached compared to their more lexical source forms

(Sankoff et al. 1997:195); and (iii) do not interfere with the truth conditions of the propositions in which LIKE occurs.

(57) a. *Like* as a verb
I still *like* to go to parties. (Santa Barbara Corpus SBC006:ALINA)

b. *Like* as a comparative preposition
[H]e's exactly *like* the bloke I fell in love with (ICE Great Britain:S1A-006:B)

c. *Like* as a comparative preposition which is best glossed as 'as for example'
Okay for instance uh a lot of nice people work here man *like* John Andrea<,> uh Raymond Charles (ICE Jamaica:S1A-008$A)

d. *Like* as a comparative preposition which is best glossed as 'as if'
It wasn't it didn't look *like* it was gonna fall (ICE Philippines:S1A-007$A)

e. *Like* as a noun
[T]here's about four wards in Lisburn that are sort of Twinbrook Poleglass and the *like*. (ICE Ireland:S1A-034$A)

f. *Like* as a suffix
He walked in a bum-*like* manner.

g. *Like* as a part of general extenders
You had a buy a two-piece suit or something *like* that. (ICE Ireland:S1A-029$B)

e. *Like* as part of lexicalizations
It's *like* I really feel upset. (ICE Canada:S1A-051$A)

As optionality is a defining feature of discourse markers and is a key criterion for the current purpose, it was decided to exclude quotative *like*, as it cannot be removed without affecting the acceptability of the utterance (compare (58) to (59)). This procedure is non-trivial, as previous research, e.g. Schourup (1982) and Underhill (1988), include quotatives into their analyses of LIKE use.

(58) Quotative *like* (BE+*like* constructions)

a. And then he walked up to the car door. I was *like* Hi. (ICE Jamaica:S1A-034$B)
b. So I'm *like* okay so do you leave or what do you do. (Santa Barbara Corpus 044:LAJUA)
c. And so I'm standing there in this florist's and I'm *like* what do I do (ICE Canada:S2A-037$A)

(59) Quotatives without *like* (BE constructions)

?a. And then he walked up to the car door. I was Hi. (ICE Jamaica:S1A-034$B)

*b. So I'm okay so do you leave or what do you do. (Santa Barbara Corpus SBC044:LAJUA)

*c. And so I'm standing there in this florist's and I'm what do I do (ICE Canada:S2A-037$A)

Although some studies (cf. Schourup 1982, 1985; Andersen 1997, 1998) have classified LIKE before numbers and other quantitative expressions as a discourse marker when it is used adverbially to signal approximation, such instances are not considered discourse marker uses of LIKE in this study. The reason for this is that in its adverbial function, *like* replaces alternative adverbs such as *about*, *around*, and *approximately* (D'Arcy 2005, 2008; Schweinberger 2010) and can be substituted with various adverbs without noticeably altering the meaning or acceptability of the utterance in which it occurs (cf. Andersen 1997; D'Arcy 2005). In addition, this use of *like* interferes with the truth conditions of the underlying proposition (cf. Siegel 2002), leading to "a specifiable semantic difference between descriptions preceded by *like* [italics M.S.] and identical descriptions without *like* [italics M.S.]" (Schourup's 1982:30). Furthermore, Meehan (1991:40) supports this analysis and suggests that this 'approximately' reading of *like* "can be thought of as a specific interpretation of 'similar to' [i.e. the adverbial extension of the adjectival use of *like*]" (1991:40) and, thus, behaves rather atypically compared to other discourse markers and constitutes a borderline case between discourse marker and adverbial (Andersen 1997:37).

Nevertheless, there are valid reasons for including such instances in the data analysis. For example, Andersen (1997, 1998) mentions that some cases of LIKE in this context do not belong in the adverb category as "in a number of cases *like* precedes a measurable unit without expressing inexactness" (1997:40). In such cases, the instance cannot be equated with other approximating instances, as it apparently serves a different pragmatic function, i.e. it serves a focusing function or, as Schourup (1982, 1985) suggests, it seems

adequate to gloss it as 'for instance'. Although such cases admittedly remain problematic, they were excluded for the reasons given above (cf. (60)).

(60) *Like* preceding numeric expressions

a. [I]t costs me *like* a a fiver more to come in for nine o'clock. (ICE Great Britain:S1A-008$A)
b. Yeah cos it took *like* five ten seconds all. (ICE Canada:S1B-062$B)
c. When you would say that it's night market it would be cheap it would be really *like* fifty percent off the original price. (ICE Philippines:S1A-080$A)

Deciding to exclude occurrences of LIKE was, however, not unproblematic when the status of the respective form was not clear-cut, as instances such as in (61) where *like* can be interpreted as both a discourse marker and a comparative preposition.

(61) Unclassifiable instances of *like*

a. What happened to you I mean not *like* dating *like* a formal date. (ICE Philippines:S1A-038$B)
b. It's a reflection of the brain and it's communication *like* books but it's much quicker. (Santa Barbara Corpus SBC017:MICHA)

Excluding significant proportions of occurrences of *like* is quite delicate, but absolutely necessary to guarantee a high quality of the data. It is, nevertheless, unfortunate that neither phonological nor prosodic annotation was available for the ICE components.[28] This would have enabled phonological and prosodic analysis and, thus, reduced the number of indeterminable cases, "as the discourse marker LIKE is generally unstressed and has little prosodic prominence (and is often pronounced with a slightly different diphthong from that of the verb/preposition, [εI] vs [αI])" (Andersen 1997:39). Although not all indeterminate cases could be resolved using phonological analysis, since non-discourse marker *like* variants may also be unstressed at times, it would

[28] An exception is the latest version of ICE Ireland (version 1.2.1). This revised version does, in fact, include phonological annotation to a degree of accuracy not yet present in the other regional components.

have allowed for a more detailed analysis and increased both the quality and the quantity of the data.

All in all, the coding procedure applied in the present study was straightforward. The coding consisted of two phases: an automated phase in which non-discourse marker cases were removed from the data, and a subsequent manual coding phase during which the instances of LIKE were classified and other non-discourse marker cases were removed. The coding consisted of the following steps: First, all instances of the orthographic sequence *like* were retrieved from the ICE components. Secondly, all cases in which *like* was only a part of a word were excluded, e.g. *likely*. Then, instances which were followed by a *to* or a word ending in –ing were excluded, as they are instances of the verb *like*. Next, all instances of *like* which were preceeded by *-ould*, *–ould not* or *–ouldn't* were excluded as they represent verbs. Then, instances were excluded when *like* was preceded by a word ending in *-thing* and the subsequent word began with *th-* because in such cases *like* is part of a general extender, e.g. *something like that*. Next, all instances of *like* which preceded quantities or numeric expressions were removed from the data set, e.g. *like five*, because in this study such instances of *like* are considered adverbs (cf. section 4.7.1.2). The final step in automated coding consisted of removing all instances of *like* which were preceeded by personal pronouns, .e.g. *I like*, because these instances of *like* were most likely verbs.

Once all these cases of *like* were removed from the data, manual coding was applied. Manual coding was straightforward as well: each instance was inspected in context and coded as being either an instance of a discourse marker, in which case the instance was retained in the analysis – or not – in which case the instance was removed from the data. If the instance was indeed a discourse marker, then it was classified as either an instance of clause-initial LIKE (INI), clause-medial LIKE (MED), clause-final LIKE (FIN), non-clausal LIKE (NON), or an instance of LIKE for which a proper classification was not possible (NA). The decision to classify LIKE as representative of one of these

categories was based on the context and syntactic environment, e.g. LIKE was considered an instance of:

(i) clause-initial LIKE, when it occurred in a pre-subject position followed by a complete clause or beginning of a clause;

(ii) clause-medial LIKE, when it occurred in post-subject position preceding a phrasal consituent, but not at the end of a clause and not surrounded by repetitions, restarts, interruptions, or pauses;

(iii) clause-final LIKE, when it occurred at the end of clauses or speech units and was preceeded by phrasal or clausal constituents;

(iv) non-clausal LIKE, when the instance of LIKE was surrounded by repetitions, restarts, interruptions or pauses;

(v) unclassifiable LIKE, when non of the above criteria applied or – more likely – when more than one classification was appropriate and the context did not favour one over the other classifications.

For a more elaborate description of these types of LIKE, including their properties and a more fine-grained description of their classification, see the following section which focuses in more detail on the considerations underlying the manual coding process.

5.7.1 Types of LIKE

Though it may superficially appear as if the discourse marker LIKE is a single, homogenous form which happens to occur in various utterance positions and syntactic environments, only more fine-grained analyses are able to provide an adequately detailed picture and reveal its multifaceted nature. On closer inspection, one finds that LIKE use comprises a heterogeneity of quite distinct situations which occur under specific conditions and in well circumscribed contexts (e.g. D'Arcy 2005:ii; Tagliamonte 2005:1897). Depending on the linguistic context in which LIKE occurs, it fulfills a variety of more or less

distinct (pragmatic) functions. The following section exemplifies various uses of LIKE and additionally provides a classification which allows the systematization of seemingly unrelated instances of LIKE.

According to Andersen (2000:272), all instances of LIKE can be subsumed under either of two categories: clause-internal and clause-external uses of LIKE. Clause-internal LIKE is "syntactically bound to and dependent on a linguistic structure […] a pragmatic qualifier of the following expression" (Andersen 2000:273). Clause-external LIKE is "syntactically unbound (parenthetical) […] external to and independent of syntactic structure" (Andersen 2000:273). The instances provided in (62) and (63) exemplify this distinction.

(62) Clause-internal, syntactically bound LIKE (clause-medial LIKE[29])

a. I thought Jews'd always been very *like* stringently against divorce. (ICE Ireland:S1B-005$D)
b. And she obviously thought she was *like* with the delivery people as well. (ICE Ireland:S1A-006$D)
c. It's got *like* chocolate chip cookie base and lovely lime juice. (ICE Ireland:S1A-036$A)

(63) Clause-external, syntactically unbound LIKE (scopeless, non-clausal LIKE)

a. Mine aren't bifocal but I find *like* that if you wear if they're for reading and you wear them out there's I don't know it's sort of *like* uhm they're uncomfortable. (ICE Ireland:S1A-059$B)
b. But there's lots of uhm *like* I mean say if you were going to analyse a a rock face I mean there's probably only one way you can actually analyse it. (ICE Ireland:S1A-028$C)
c. UCD *like* first of all well well UCC supposedly it's meant to be easier to get into second year Psychology. (S1A-048$B)

Instances of clause-external LIKE require further sub-classification, because certain instances of clause-external LIKE are highly functional in so far as they introduce specifications best glossed as *that is* (cf. (63a)); establish coherence relations by linking higher level constructions as, for example, entire clauses or clausal elements as in (63b); or serve to indicate restarts as in (63c).

[29] Clause-medial *like* is equivalent to Andersen's (1997: 38) category labeled clausal *like.*

Instances of clause-external LIKE as in (63d), on the other hand, merely indicate processing difficulty and function as floor-holding devices while neither modifying elements nor establishing coherence relations.

(64) Clause-external, clause-initial LIKE (clause-initial LIKE)

a. *Like* will your job still be there when you if if you do come back. (ICE Ireland:S1A-014$D)

b. [I]t was a bit of a cheat. *like* it was a bit like wings of desire. (ICE New Zealand S1A-026#265:1:A)

c. *Like* we were sitting any time there was a bit of music we were sitting clapping away and like everybody starts clapping along to the music you know. (S1A-012$A)

An additional difference between (63) and (64) is that the instances of LIKE in (64) have scope over the entire higher-level construction following while the instances in (63) appear to lack scope altogether.

Related to the clause-internal versus clause-external distinction is the observation that the instances of LIKE in (62) modify a single element, while the instances in (63) index planning difficulty or serve as floor-holding devices, repair indicators, and discourse links, but do not modify individual elements. Except for cases in which LIKE signals planning difficulty or precedes an utterance termination, all of the above types of LIKE have forward scope, i.e. they relate to whatever follows to their right.

Nonetheless, LIKE may differ with respect to the direction of its scope. While the instances in (62) and (64) are bound to the right and thus have forward scope, the occurrences of LIKE in (65) are bound to their left and, therefore, have backward scope.

(65) LIKE with backward scope (clause-final LIKE)

a. It's a bit of a difference now from him going to Manchester and you going to a kibbutz in Israel *like*. (ICE Ireland:S1A-014$C)

b. They're in their bedroom *like*. (ICE Ireland:S1A-036$A)

c. He's from Wexford so he's probably no good but we'll sign him up anyway *like* you know (ICE Ireland:S1B-050$D)

This difference in direction of scope probably reflects their diverse origin. While LIKE with forward scope as in (62) probably originated from the comparative preposition (Buchstaller 2001:22; Meehan 1991; Romaine & Lange 1991), instances of LIKE with backwards scope probably originated from the suffix -*like* (Jespersen 1954:407, 417).

Instances of LIKE with backward scope do not, however, necessarily represent instances of clause-final LIKE, but may occur either in non-clausal constructions as in (66), or in clause-medial position as in (67).

(66) LIKE with backward scope in non-clausal constructions

a. Yeah after Mass *like*. (ICE Ireland:S1A-022$B)
b. How long how long *like*. (ICE Ireland:S1A-051$B)
c. A wee girl of her age *like*. (ICE Ireland:S1A-002$D)

(67) LIKE with backward scope in clause-medial position

a. [In Bergen *like*] it rains a lot.
b. There's [John *like*] standing by the stairs.

Furthermore, instances of LIKE which were clearly discourse marker uses but which could not be classified satisfactorily due to missing or ambiguous context (cf. (68) and (69)) have been classified as NA (not available). In fact, the relevance of analyzing the wider context of *like*-instances proved to be crucial for disambiguating problematic cases, and its importance cannot be overstated.

(68) Unclassifiable instances of LIKE

a. [H]e changed to a petrol just before my last lesson so I've had *like* ... and everything was fine but now getting used to the petrol's really hard (ICE Ireland:S1A-003$C)
b. NANCY: But *like* ... (Santa Barbara Corpus SBC050:NANCY)
c. I was *like*… (ICE Ireland:S1A-066$C)

Instances of *like/*LIKE such as those in (68) have been particularly difficult to classify because the direction of scope is ambiguous. This ambiguity arises when cues which enable the identification of scope direction – e.g. pauses or

metalinguistic information provided by the transcribers – are missing (compare (69a) and (69c) to (69b) and (69d)).

(69) LIKE with ambiguous scope

a. If you haven't found another job [within five years *like*] there must be something seriously wrong. (ICE Ireland:S1A-014$D)
b. If you haven't found another job within five years [*like* there must be something seriously wrong]. (ICE Ireland:S1A-014$D)
c. We were like oh for fuck sake [*like* Jesus]. (S1A-011$NA)
d. We were like [oh for fuck sake *like*] Jesus. (ICE Ireland:S1A-011$NA)

Since the instances of LIKE in (68) are clearly instances of the discourse marker LIKE, but ambiguous with respect to their clausal status, they are also classified as "NA" to indicate that further classification was "not available".

According to the typology of uses described above, the present study distinguishes between the following types of LIKE:

- INI: clause-initial with forward scope as in (64);
- MED: clause-medial with forward scope as in (62);
- FIN: clause-final with backward scope as in (65) and non-clausal LIKE with backward scope as in (66);
- NON: syntactically unbound, i.e. non-clausal and without scope as in (63);

Limiting the analysis of the discourse marker LIKE to its relation to the clause offers the advantage that results allow not only for directly meaningful comparability to other contemporary studies, but also remain viable for future research (Macaulay 2005:189). Tagliamonte, in particular, points out the importance of guaranteeing comparability in sociolinguistic research (2005:1912):

> On a more methodological note, these results highlight the value of pursuing a quantitative analysis of proportion and distribution when it comes to innovating features, even when they may have a number of different functions in the grammar. It is only when the high frequencies of individual

> forms are calculated from the total number of words spoken by individuals or groups (or some other normalizing measure) that number of forms can be compared accountably (whether across different sub-groups of the speech community or across studies).

Accordingly, the present study concerns itself exclusively with the positioning of the discourse marker LIKE, while for the most part disregarding the pragmatic functions of each individual occurrence to increase reliability and enable replication:

> [T]here have been different interpretations of the meaning of features such as […] focuser *like*. While notions of shared knowledge or similarity may not have affected the approach taken by investigators to these two items, such assumptions may make it more difficult for other investigators who do not share them. An ascetic approach in which discourse features are first of all treated as units of form avoids introducing controversial interpretations at an early stage. (Macaulay 2005:189-190)

Having classified all instances of the discourse marker LIKE accordingly, each instance was assigned to its respective speaker. In a subsequent step, all speakers present in the data were assigned the type and number of occurrences of LIKE they had used. This allows for a speaker-based analysis, which is a more suitable method than proceeding on an item-by-item basis.

The next step consisted of computing the per-1,000-word frequencies of each type of LIKE for each individual speaker.

In the final phase of editing, outliers were removed from the data as these speakers would have disproportionately affected the analysis.[30] Thirty-six outliers were identified and removed.[31] The elimination was criteria-based, i.e.

---

30 These outliers have not been deleted as we will come back to them when analyzing the leaders of change within each variety.

31 One outlier was a white university-educated Californian female aged thirty, who used 50 instances of LIKE in a total of 438 words. Her per-1,000-word frequency of LIKE use was thus outstanding, i.e. 114.16 (in comparison, the second highest frequency value

when the LIKE use of the respective speaker differed substantially from the distribution observable in his or her regional variety, or if speakers used relatively few words which led to an overestimation of his or her per-1,000-word frequency of LIKE. To exemplify, Figure 9 shows two box plots displaying outliers in PhiE and AmE which were removed from the data set. In PhiE the upper two data points have been removed, and in the case of AmE, the upper three.

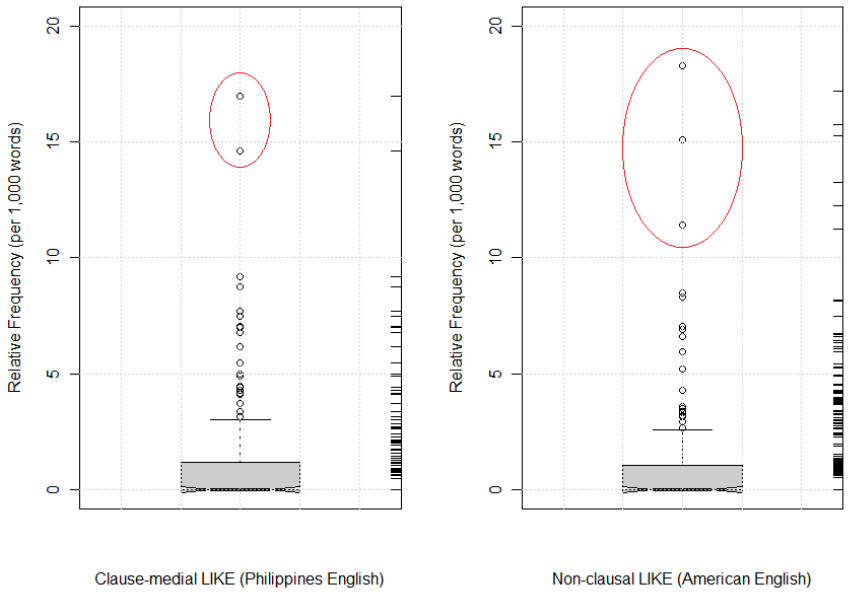

Figure 9: Examples for outliers in PhiE and AmE

It has thus been possible to compile an appropriate database for our analysis containing 1,925 speakers across eight varieties of English who produced 4,661 instances of the discourse marker LIKE (cf. Table 5).[32]

was 34.30). The speaker with the second highest frequency value was removed due to her low word count (175) and a high frequency of LIKE use. She too was a white university-educated US American female, but slightly younger at age 22.

32 Only 30 instances of discourse marker LIKE could not be assigned to the speaker who uttered them due to missing annotation in the corpus itself.

Table 5: Overview of the data base for the present analysis

Variety (ICE component)	Words (SUM)	Speaker (N)	INI (N)	MED (N)	FIN (N)	NON (N)	NA (N)	ALL (N)
Canada	194,574	244	368	381	26	112	13	900
GB	201,372	320	37	59	2	29	---	127
Ireland	189,787	309	249	237	318	118	14	936
India	211,646	236	107	64	21	132	7	331
Jamaica	207,807	228	138	288	3	86	11	526
New Zealand	229,193	227	209	183	20	115	2	529
Philippines	193,077	198	156	199	10	77	10	452
Santa Barbara C.	246,258	163	220	390	1	234	15	860
Total	1,673,714	1,925	1,484	1,801	401	903	72	4,661

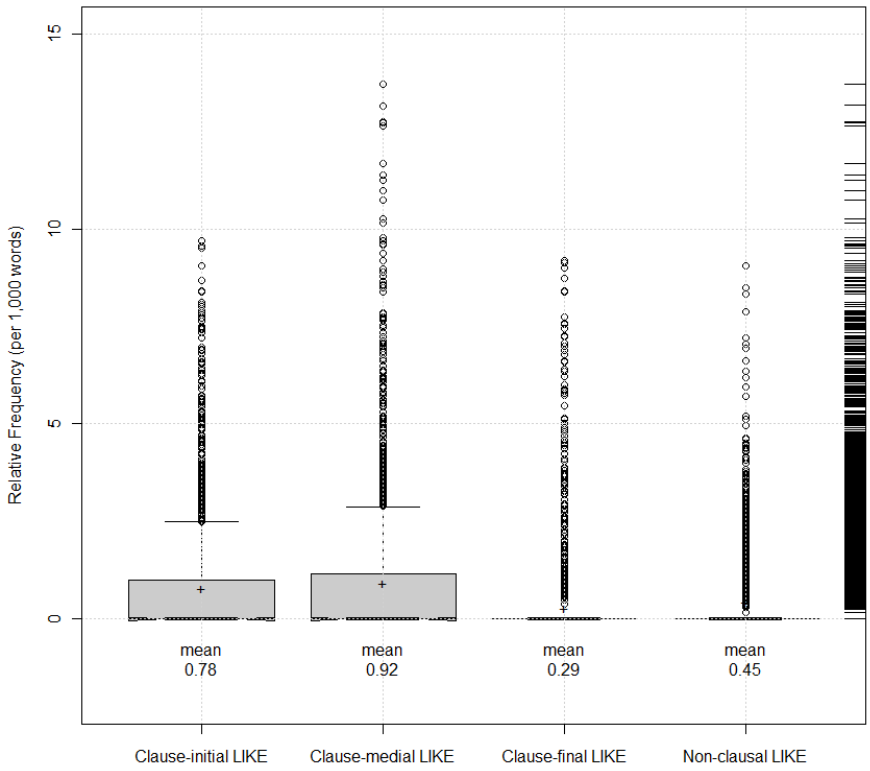

Figure 10: Frequency of LIKE variants in the final dataset[33]

[33] Box plots are very advantageous for displaying the general structure of data as they contain substantial valuable information (Gries 2009: 119). The bold horizontal lines

Figure 10 shows that the most frequent variant of LIKE across varieties of English is clause-medial LIKE, while clause-initial LIKE is slightly less frequent. The missing boxes of clause-final and non-clausal LIKE indicate that these variants are very infrequent among speakers of English.

5.8 Description and motivation of variables

Although a large body of literature on LIKE has accumulated, only a few studies have approached this marker with elaborate statistical evaluation; a notable exceptions are Andersen (1998, 2000, 2001), D'Arcy (2005), Dailey-O'Cain (2000) and Tagliamonte (2005). In light of this research gap, the present investigation will perform elaborate statistical evaluations to assess the correlation between the use of LIKE, i.e. the dependent variable, and up to nine independent variables such as age, gender, and occupation.

5.8.1 Dependent variables

The dependent variables of the analysis are straightforward, as they constitute the frequencies of types of LIKE as discussed in section (5.6). In addition, the variable ALL is added, which comprises all occurrences of LIKE without regard

represent the median, i.e. the value which separates the higher from the lower half of a given data set (Gries 2009: 119). The regular horizontal lines that make up the upper and lower boundaries of the boxes represent the hinges, i.e. approximately the 75 percent and the 25 percent quartiles (Gries 2009: 119). The whiskers, i.e. the dashed vertical lines extending from the box to the upper and lower limits, represent the largest and the smallest values which are not more than 1.5 interquartile ranges away from the box (Gries 2009: 119). Each outlier that would lie outside the range of the whiskers is represented as an individual dot, and "the notches on the left and right side of the boxes extend across the range ±1.58*IQR/sqrt(n): if the notches of the two boxes overlap, then these will most likely not be significantly different" (Gries 2009: 119).

to their clausal status. ALL thereby provides an estimation of the rate of overall LIKE use within a certain variety of English.

Table 6: Overview of the dependent variables included in the analysis

Variable name	Variable type	Definition (per 1,000 words per speaker)
ALL	numeric, ratio-scaled	overall frequency of LIKE
INI	numeric, ratio-scaled	clause-initial LIKE with forward scope
MED	numeric, ratio-scaled	clause-medial LIKE with forward scope
FIN	numeric, ratio-scaled	clause-final LIKE and non-clausal LIKE with backward scope
NON	numeric, ratio-scaled	syntactically unbound LIKE without scope

The independent variables included in the analysis are based on the speaker information provided by the respective ICE teams and comprise various highly relevant sociolinguistic factors.

5.8.2 Independent variables

The independent variables mostly comprise traditional sociolinguistic variables, such as the age and gender of speakers. In order to probe more deeply into the matter, this study includes additional variables to enable a maximally accurate and optimally fine-grained analysis of LIKE use across and within varieties of English. The variables included here are discussed either by previous research of discourse marker LIKE, or theoretical considerations concerning the study of language change on a more general level, or evidence from related fields such as psycholinguistics. The following section will thus discuss each independent variable and state how it has been operationalized in the present study.

Table 7: Overview of the independent variables included in the analysis

Variable name	Description	Variable type	Levels	Definition
VAR	Regional variety of English	categorical	USA CAN GB IND IRE JAM NZ PHI	U.S American English Canadian English British English (mainly London area) Indian English Irish English Jamaican English New Zealand English Filipino/Philippine English
SEX	Gender	nominal	F, M	F: female (reference variable), M: male
AGE	Age	nominal (dummy variable)	0, 1	As the age classification differs from corpus to corpus, the individual age cohorts are described in the respective sections. (A1= reference category).
OCC	Current occupation	nominal (dummy variable)	0, 1	SML: (un-)skilled & manual labor, CLM: clerical & managerial professions, ACA: academic & professional career
L1	First language	nominal (dummy variable)	0, 1	ENG: English only, (reference) ENG+: balanced bilingual with one L1 being English, OTH: first language is not English
ETH	Ethnicity	categorical	FIL, IND, JAM, MAO, WHI, OTH	Filipino, Indian, Jamaican, Maori, Caucasian none of the above
REG	Geographic region	categorical	N, S, CA, OTH	Northern Ireland Republic of Ireland, California, US but not California
DATE	Date of data collection	nominal (dummy variable)	0, 1	D1: 1990-1994, D2: 1995-2001, D3: 1002-2005
PAI	Priming Accommodation Idiosyncratic bias	numeric, ratio scaled		Combined frequencies of all other variants of LIKE in the dialogue (not applicable in the case of ALL being the dependent variable)

5.8.3 Regional variety (VAR)

The geographical variety of speakers is particularly interesting in the present case, as it may serve to shed light on the popular notion that LIKE has its origin in the speech of Californian Valley girls. While the claim that LIKE has its origin California is probably inaccurate in the case of LIKE (cf. D'Arcy 2007; Schweinberger 2013; section 4.1), it is very plausibly true for the quotative complementizer BE LIKE (cf. section 4.1). Quotative *like* is presumably a later development than the discourse marker LIKE (Buchstaller 2006a:363; D'Arcy 2007:386; Blyth, Recktenwald & Wang 1990), but it shares several properties with other vernacular uses of LIKE as, for example, its sociolinguistic profile. Accordingly, it has been associated with the discourse marker in several studies (cf. e.g, Schourup 1982, 1985).

The origin of the discourse marker LIKE, on the other hand, are more controversial. With respect to these forms, Andersen (2000:216) suggests that they developed in the counterculture groups (i.e., jazz, cool, and Beat) of New York City during the 1950s and 1960s (Andersen 2000:216; see also D'Arcy 2007:398), while D'Arcy (2009:400) promotes the idea that both markers have been around much longer than proposed by Andersen:

> Not only is the marker used by speakers older than 65 years of age, but the examples do not represent random occurrences. As noted, the marker is highly productive among Torontonians in their eighties, occurring nearly as often as all other discourse markers combined in the speech of this cohort. This suggests that the marker was already a feature of the vernacular before it was associated with the Beat and jazz groups of the 1950s and 1960s. In fact, working from the apparent-time hypothesis, in the 1930s, when these 80-year-olds were teenagers, *like* must have been relatively frequent in the ambient language as a discourse marker, a usage inherited by these speakers from the previous generations. (D'Arcy 2007:400)

This observation by D'Arcy is corroborated by Schweinberger (2013), who found substantial quantities of clause-medial uses in the speech of 65- to 75-

year-old speakers of Northern IrE in data compiled in the early 1970s. This finding seems implausible given the claim that LIKE originated in the US during the 1950s and 1960s. The presence of these forms in the data of older cohorts strongly suggests that the discourse marker LIKE has been around much longer than suspected and clearly undermines the hypothesis that LIKE is an American innovation.

On a more methodological note, regional variety is based on the respective ICE component and is thus taken to represent the origin of the data; that is, all linguistic output of speakers present in the refined ICE New Zealand data are assumed to reflect the variety of this region. This variable has eight levels and is operationalized as a categorical variable.

Table 8: Operationalization of regional variety of English in this analysis (VAR)

Variable name	Description	Variable type	Levels	Definition
VAR	Regional variety of English	categorical	USA	U.S American English
			CAN	Canadian English
			GB	British English (mainly London area)
			IND	Indian English
			IRE	Irish English
			JAM	Jamaican English
			NZ	New Zealand English
			PHI	Filipino/Philippine English

5.8.4 Gender (SEX)

With respect to gender-related LIKE use, the popular notion that LIKE – in all its forms – is more frequently employed by female adolescent speakers has been analyzed by quantitative sociolinguistic (e.g. Andersen 1998, 2000; D'Arcy 2005; Tagliamonte 2005) and attitudinal studies (Buchstaller 2006a; Dailey-O'Cain 2000:69). According to this popular view, LIKE has its origin in AmE (Andersen 2001:216), or more specifically in the speech of Californian Valley girls (cf. Blyth, Recktenwald & Wang 1990; Siegel 2002) who

presumably promoted its spread around the English speaking world (D'Arcy 2007:397; Meehan 1991; Romaine & Lange 1991:269; Siegel 2002:37; Underhill 1988; Valentine 1991).[34] This notion concerning the origin of quotative LIKE is fairly plausible, given that contemporary theories on language change and its interaction with social variables corroborate the fact that women are the forerunners in language change (Andersen 2000; Eckert 1989; Labov 1990, 2001, 2002; Milroy & Milroy 1993a; Müller 2005; Wodak & Benke 1998): "it is women who are ahead–usually by a full generation" (Labov 2002:12). Despite building on popular opinion as well as scientific theorizing, the considerable amount of research on this issue over the past 25 years has failed to provide unambiguous results. In fact, the emerging picture remains rather mixed and fails to provide a clear-cut answer as to whether LIKE use is significantly gendered. According to Dailey-O'Cain, this parallels conflicting "findings about *you know* and other discourse markers labelled as 'hedges' by early scholars of pragmatics and discourse analysis" (2000:63).

While a notable number of studies support the popular belief that LIKE is more pervasive among females and, in particular, among teenage girls (Andersen 2000; Iyeiri, Yaguchi & Okabe 2005; Fuller 2003; Romaine & Lange 1991; Siegel 2002; Tagliamonte 2005), other studies, for example, Dailey-O'Cain (2000:66) and Schleef (2004), failed to validate a statistically significant impact of gender on LIKE use. Moreover, this inconsistency remains, even though other variables such as the gender of addressees were taken into account. For instance, Schleef's (2004) results suggest that female students

[34] It is interesting to note in this context that, whatever the reality of the matter may be, LIKE is perceive to be more frequently employed by females (Dailey-O'Cain 2000, Recktenwald and Wang 1990). Associated with this perception is the fact that the use of LIKE seems to suggest that the speaker is less educated, but more attractive, cheerful, friendly and successful, while the correlation between LIKE and the attributes 'intelligent' and 'interesting' was not statistically significant (Dailey-O'Cain 2000: 73-74). Although not significant in Dailey-O'Cain's study, the popular press (cf. Mehren 1999) claims that LIKE makes students sound less confident and less intelligent.

employed LIKE with a significantly higher frequency than their male peers, but he failed to validate this observation when other factors such as the conversational roles and the educational level of speakers where considered. Thus, Schleef had to conclude that the "more obvious gender trend among students might be due to peer group socialization. […] For instructors, the peers of their own discipline are an important focus, and gender lines are of relatively little importance in the context of the discipline" (2004:372).

Part of the confusion may have arisen due to gender-specific differences in use of distinct types of LIKE. In other words, functionally and positionally distinct variants of LIKE use might be preferred by speakers of different genders. If all variants are treated as realizations of one underlying form, this could have skewed the results and led to the contradictory claims. An indicator of this is the female preference for quotative LIKE (e.g. Tagliamonte & D'Arcy 2004), which is more often than not studied in isolation, thus avoiding commingling with alternative variants. To avoid intermixture of distinct variants of LIKE, D'Arcy's (2005, 2006) large-scale quantitative study on CanE provided a detailed analysis finding that LIKE, when used in clause-initial position significantly correlated with females, while clause-medial LIKE significantly correlated with males, with both patterns remaining stable over apparent-time (D'Arcy 2007:396):

> [T]he gender puzzle is finely articulated: the question of men versus women depends on which vernacular form of *like* is at issue. In the case of the quotative, women use *be like* significantly more than their male peers do overall (N = 6,364; Tagliamonte and D'Arcy 2007). Concerning the discourse marker, women use this form more frequently than men do as well, and despite the narrow margin in the overall results, this too is significant (N = 3,363; D'Arcy 2005:97). (D'Arcy 2007:396)

Gender as a variable is operationalized based on the sex of a speaker, i.e. it is a binary nominal variable and not a ratio-scaled variable, reflecting a

continuum between masculine and feminine. Hence it is operationalized as follows:

Table 9: Operationalization of the gender of speakers in this analysis (SEX)

Variable name	Description	Variable type	Levels	Definition
SEX	Gender	nominal	F, M	F: female (reference variable), M: male

5.8.5 Age (AGE)

Research on language change and variation has produced a detailed picture of general patterns and mechanisms which underlie the seemingly chaotic patterns of linguistic behavior (cf. Labov 2001). One of the most salient and consistent aspects of this systematicity is that adolescents play a crucial function in triggering and driving ongoing change. Tagliamonte (2005), for instance, emphasizes that teen-talk represents a "tremendous breeding ground for linguistic innovation that exists […] in contemporary, urban speech communities" (2005:1913). Similar to other forms of cultural practice, adolescents are the main innovators and the driving force for linguistic change: "As with any form involved in change, adolescents are in the vanguard. They are not the only members of the community using these forms, but they use them at higher frequencies than older age cohorts within the population" (D'Arcy 2007:402).

To investigate language change at its forefront, a number of studies have thus almost exclusively focused on the speech of younger age groups and their use of LIKE (e.g. Andersen 2000; Fuller 2003; Miller & Weinert 1995, 1998; Müller 2005; Sharifian & Malcolm 2003; Siegel 2002, Tagliamonte 2005). Paying attention to speech patterns during adolescence is particularly promising, because it "provides greater motivation than at any other time in life to adapt linguistic patterns to community structure. This age group, therefore, provides an important key to the study of the mechanisms of such adaptation" (Eckert 1988:206). Indeed Tagliamonte (2005) argues that the

study of adolescents is "a gold mine for innovative linguistic features, revealing evidence for both grammatical, as well as sociolinguistic change" (2005:1913).

This also applies to the discourse marker LIKE as, in contrast to gender, the impact of age on the use of LIKE is far less controversial, and the results are far more homogenous than those for gender. With regard to age, the literature provides a comprehensive and fine-grained account of its correlation with changes in language use. The vast majority of studies conducted during the past twenty years confirm that adolescents show a higher frequency of LIKE use than their elders. In other words, age has a statistically significant impact on LIKE use which has been described as peaking during the teenage years (Andersen 2000; Dailey-O'Cain 2000; Müller 2005; Peterson 2004; Schourup 1982; Schweinberger 2012). Consider, for instance, D'Arcy (2007:397):

> It is a common assumption that vernacular uses of *like* are age-graded, frequently marking the speech of adolescents and younger adults only to be outgrown in adulthood. In other words, *like* use is presumed ephemeral and temporally banded, appropriate for a certain stage of life and then shrugged off when its suitability wanes.

This mainstream notion has been validated empirically by D'Arcy (2005) in her study of LIKE use in Toronto English. However, her research raises the question as to why this tendency does not wane like other short-lived fashions. As an answer to this question, D'Arcy (2005, 2007) proposed that this consistency may be related to an increase in syntactic flexibility, which leads to an increase in slot positions in which LIKE is allowed to occur.

Based on the results of previous studies, the present analysis expects young speakers to use LIKE more frequently than older speakers. Accordingly, the youngest age cohort will serve as the reference variable, in comparison to which the frequencies of other age groups are evaluated.

The operationalization of this variable requires a little more attention, since the age groups are not coherent across different ICE components. For example, there are six age groups in the original unedited Irish ICE component (age

group 0 = 0 - 18 years of age; age group 1 = 19-25 years of age; age group 2 = 26-33 years of age; age group 3 = 34-41 years of age; age group 4 = 42-49 years of age; age group 5 = 50 and older), while there are only three age groups in the British component of the ICE (age group 1 = 18 - 25 years of age; age group 2 = 26 - 45 years of age; age group 3 = 46 and older). Because of this heterogeneity, the variable of age does not lend itself to cross-varietal comparisons and is used only for intra-varietal analyses. Since the youngest age group will serve as the reference category in the regression models, it is always coded as A1, although the age of the speakers within that category may differ between varieties based on the age groups defined in the respective ICE components. The specifications of the age groups are provided in the respective sections. In addition, the *Santa Barbara Corpus* throughout and the ICE GB in parts offer the exact age of speakers, which will be used in graphical displays if advantageous, but these exact details are transformed into age groups during regression modelling. In general, the age of speakers is operationalized, as displayed in Table 10 below.

Table 10: Operationalization of the age of speakers in this analysis (AGE)

Variable name	Description	Variable type	Levels	Definition
AGE	Age	nominal (dummy variable) OR numeric (inUS data and parts of GB data)	0, 1 (if dummy) OR age-in-years (in US and parts of GB data)	As the age classification differs from corpus to corpus, the individual age cohorts are described in the respective sections. (If age is coded as dummy: the youngest cohort (A1) serves the reference category).

5.8.6 Occupation, social-class and socio-economic status (OCC)

Social class has acquired a central role in sociolinguistic research ever since Labov (1966, 1972) confirmed the existence of a reliable correlation between social stratification of linguistic features. Despite the fact that social class is commonly considered a crucial factor in explaining the mechanisms of language change and variation, only a few studies have concerned themselves

with the analysis of the interaction between LIKE use and social class, a notable exception being Andersen (1997, 2000). In accordance with Blyth et al., who propose that LIKE is "indicative of middle-class teenage girls" (1990:224), Andersen's (1997) study on LIKE in London Teenage speech notes that LIKE is distributed rather evenly, but that it is most common among the "higher social groups" (Andersen 1997:46). This trend holds true across both sexes and indicates "that *like* appears to be gaining ground in a fairly wide range of speaker's groups and across registers and is used by speakers who otherwise speak a standard variety of English (Andersen 2000:290). According to Andersen (1997a), the fact that LIKE is especially common among members of the highest social groups can best be accounted for in terms of a higher adoption rate of this 'American' feature by teenagers belonging to the middle and upper-middle classes. Strangely, the curvilinear pattern neither holds for quotative LIKE nor the grammaticalized construction *it's like* in Andersen's data, as the correlation between these forms and social class turned out to be insignificant.

With respect to methodology, it should be stressed that in the present study, the assignment of social class is not based on the combined SEC index of occupation, education and house value (Labov 2001:297), but solely on speakers' occupations. As the occupation of speakers correlates particularly closely with the socio-economic status of speakers (cf. Labov 1990, Macaulay 1977), occupation will serve as an indicator of both social class and socio-economic status. This approximation is corroborated by recent analyses, which found that occupation is a very accurate approximation of the SEC index: "occupation was correlated more closely with new and vigorous changes in progress than the other two, and at some points occupation gave even more significant correlations than the combined index" (Labov 2001:297).

In contrast to classifications common in other variationist research such as Chambers and Trudgill's (1998) five-point scale of social class (consisting of middle-middle class, lower-middle class, upper-working class, middle-working class, and lower working class), the present analysis utilizes a three-point

scale: (i) academic and professional careers; (ii) clerical and managerial professions; and (iii) skilled and unskilled manual labor, similar to the tripartite categorization employed by Macaulay (1978:138).

According to the curvilinear hypothesis, social stratification of linguistic variation is concentrated in the middle classes (Labov 2001:186). In other words, the highest ratio of advanced forms exists in the speech of speakers of high, but not the highest, socio-economic status. Hence, speakers in clerical and managerial professions serve as the reference variable in the present context.

In the present study, the socio-economic status of speakers is coded as a nominal variable: speakers who hold occupations of medium social prestige typically associated with office work, such as secretaries and accountants, are the reference variable – this variable level is referred to as CLM, which stands for *clerical and managerial professions*. The other two levels which are tested against this variable encompass: (i) speakers in occupations which typically do not require secondary education such as waiters, janitors or cleaning personnel – this variable level is called SLM which stands for *skilled and manual labour* and; (ii) speakers in occupations typically associated with university education, such as lawyers, teachers and lecturers – this variable level is called ACD, which stands for *academic and professional careers*. These levels are not based on groups defined by the ICE teams, but are based on manual classification whenever the current occupation of a speaker was provided in the speaker information data.

Table 11: Operationalization of the occupation of speakers in this analysis (OCC)

Variable name	Description	Variable type	Levels	Definition
OCC	Current occupation	nominal (dummy variable) OR categorical	0, 1 OR if coded as factor: see definition	SML: (un-)skilled & manual labor, CLM: clerical & managerial professions, ADC: academic & professional career (ADC serves as the reference category)

5.8.7 Native and non-native speakers of English (L1)

Only very few studies have investigated the difference in LIKE use and its frequency among native and non-native speakers of English. A notable exception is Müller (2005:229-230), who found that non-native speakers of English (L1 German speakers) used LIKE significantly less than native speakers of AmE. In fact, native speakers of AmE used LIKE in various functions up to four times more frequently than non-native speakers of English.

An additional study analyzing differences between native and non-native use of LIKE in language contact situations is Sankoff et al. (1997). This study investigated differences in LIKE use by L1 English and L1 Montreal French speakers in Québec, Canada – a region included in the present analysis, and it offers particularly insightful findings with respect to LIKE use in language contact situations. According to Sankoff et al. (1997), the frequency of the discourse markers such as *you know*, *so*, *like* and *well* correlates negatively with their fluency in the L2: "The overall picture that emerges from our analysis is that the ability to express oneself fluently and confidently in a second language entails the use of those discourse markers that native speakers produce so effortlessly" (Sankoff et al. 1997:213). Furthermore, this study strongly suggests that "only L2 speakers with a high degree of contact with native speakers will master the use of discourse markers" (Sankoff et al. 1997:193). Thus, only speakers whose native language is English will serve as a reference variable in the following statistical analysis.

In terms of operationalization, the variable first language (L1) is coded as a nominal variable with speakers whose only first language is English as the reference variable. The other two levels are tested against this variable and encompass (i) balanced bilinguals with one L1 being English (ENG+) and (ii) speakers with a first languge other than English (OTH).

Table 12: Operationalization of the mother tongue of speakers in this analysis (L1)

Variable name	Description	Variable type	Levels	Definition
L1	First language	nominal (dummy variable)	0, 1	ENG: English only, (reference) ENG+: balanced bilingual with one L1 being English, OTH: first language is not English

5.8.8 Ethnicity (ETH)

According to Labov (2001), ethnicity differs from the variables discussed so far – gender, age, and social class – "in that it has little systematic effect on linguistic change in progress" (Labov 2001:257).

As most researchers appear to presume that LIKE is predominately used by white adolescent speakers, ethnicity has largely been neglected. Only two systematic studies have included ethnicity into their analyses (Andersen 2000; Ferrara & Bell 1995). Ferrara and Bell (1995) found that black and Hispanic speakers contribute to the spread of the quotative construction BE+*like* in AmE – with both genders equally taking part in this change. Andersen (2000:290-291) who examined the use of LIKE among London teenagers provides a more fine-grained analysis and was able to show that although LIKE is a common feature in the speech of ethnic minorities, it is occurring with a significantly lower frequency than in the speech of comparable white adolescents. Furthermore, Andersen's (2000) results support the assertion that females have been responsible for introducing the discourse marker LIKE into minority communities, but he did not find supporting evidence for this trend concerning quotative *like* and, hence, failed to corroborate Ferrara and Bell's findings for AmE.

Based on the findings reported by Andersen (2000), and Ferrara and Bell (1995), the reference group in this study is white Caucasian speakers who are expected to use LIKE more frequently than other ethnic groups.

With respect to operationalization, the variable ethnicity (ETH) is coded as a categorical variable with speakers whose ethnic identity is white/caucasian as the reference variable. Other levels are tested against this variable and encompass levels listed in Table 13 below.

Table 13: Operationalization of the ethnicity of speakers in this analysis (ETH)

Variable name	Description	Variable type	Levels	Definition
ETH	Ethnicity	categorical	FIL, IND, JAM, MAO, WHI, OTH	Filipino, Indian, Jamaican, Maori, Caucasian none of the above

5.8.9 Region (REG)

The variable "region" may serve to shed light on more fine-grained pathways of diffusion within varieties of English. Unfortunately, this variable is only applicable to IrE and AmE, since the relevant information has not been available for other varieties. Nonetheless, regional differences with geographical varieties may indicate change in progress, since innovations are assumed to spread outward from distinct epicenters, typically urban areas, where innovations commonly originate. In the case of AmE, regional differences may shed light on whether LIKE is still undergoing change. For instance, if LIKE exhibited higher frequencies and LIKE in California compared to other US American regions, this could potentially indicate that LIKE is still in the process of diffusing through AmE. However, one has to keep in mind that region in the present context is not geared toward representing social networks or communities of practice, as it is far too indiscriminate for this purpose: it merely serves to evaluate the existence or absence of possible regional differences in LIKE use on a rather coarse-grained level. This is particularly true of the Irish data.

With respect to IrE, this study differentiates between northern and southern varieties of IrE. Although these two varieties exist in geographical

proximity, they are distinct not only in terms of phonology, but also in terms of morpho-syntactic properties. Should the statistical analysis show significant differences in LIKE use between these two varieties, it might shed light on factors which counter the spread of incoming forms and may best be accounted for in terms of distinct socio-cultural identities which are reflected in more or less intentional differences in linguistic behavior.

However, the category region in the present context is only a rough approximation of social networks or communities of practice, and is far too indiscriminate for this purpose; it merely serves to evaluate the presence or absence of possible regional differences in LIKE use. This is particularly true of the Irish data.

Since region applies only to Irish and the American data, it is operationalized as a nominal variable with two levels: Northern IrE vs Southern IrE for the Irish data, and California vs. US but not California for the American data.

Table 14: Operationalization of region in this analysis (REG)

Variable name	Description	Variable type	Levels	Definition
REG	Geographic region	nominal	N, S, CA, OTH	Northern Ireland Republic of Ireland, California, US but not California

5.8.10 The date of data collection

The date of data collection is of particular relevance in the present study as it serves to complement the apparent-time analysis, based on the age of speakers, with a real-time analysis. If LIKE is truly innovative and is currently undergoing change, as proposed by various sources (cf. Andersen 2000; D'Arcy 2005, 2007), then the earliest data, i.e. that compiled from 1990 to 1994, should exhibit lower frequencies of LIKE than data collected more recently, i.e. 2002 to 2005. Although the time span captured by the date of data compilation

covers a mere fifteen years, the apparent-time findings of previous studies indicate a rapid change, which may be reflected in the results of the complementary real-time analysis even though this period is brief. This real-time study would thus confirm the hypothesis that LIKE continues to spread within and across speech communities.

So far, LIKE has only been studied from an apparent-time perspective, which has been shown to accurately reflect ongoing changes in real-time. Nevertheless, apparent-time analyses are often inadequate to distinguish between different types of change – such as generational change versus communal change – or even the difference between change and mere stable variation, as in age-grading. As Labov notes: "[t]he main problem then is to distinguish age-graded stable variation from generational change in progress" (2001:77). Hence, the apparent-time distributions will be complemented by real-time analyses thereby providing a more detailed picture of LIKE use.

Unfortunately, the American data taken from the *Santa Barbara Corpus*, as well as the data taken from the British, the New Zealand and the Indian components of the ICE do not provide the necessary information to evaluate the apparent-time results from an additional real-time perspective. Hence, the present study can only provide apparent-time analyses of these components.

The date of data compilation is operationalized as a nominal variable, which is tested against a dummy variable – in the present case the dummy represents the oldest data, compiled between 1990 and 1994. If DATE significantly affects LIKE use and has a positive coefficient, then this confirms that LIKE use has significantly increased in comparison to the LIKE of 1990 to 1994.

Table 15: Operationalization of date of data collection in this analysis (DATE)

Variable name	Description	Variable type	Levels	Definition
DATE	Date of data collection	nominal (dummy variable)	0, 1	D1: 1990-1994, D2: 1995-2001, D3: 2002-2005

5.8.11 Priming effects, accommodation and idiosyncratic overuse (PAI)

Thus far priming effects, i.e. that speakers themselves or their interlocutors re-use linguistic items or constructions which have occurred in the preceding discourse (cf. Pickering & Garrod 2004), have not as yet attracted much attention with respect to possible influences on language change in general (Jäger & Rosenbach 2008a:86, 91), or the historical development of discourse markers in particular. Although priming effects are well established with regard to phonological and syntactic features in synchrony, this is certainly not the case with respect to its effect on language change, let alone the diffusion of pragmatic markers.

Nevertheless, it appears very plausible that priming is relevant in the context of both the study of discourse markers and the study of ongoing language change. . Consider, for instance, Szmrecsanyi (2006:212):

> [P]ersistence could also have implications for historical linguistics: the multiplicative and self-enforcing effect of persistence, coupled with logarithmic forgetting functions, might very well be involved in the s-curve patterns so often observable in language change. This is an intriguing issue which, needless to say, would be worth exploring in future research.

Only recently has this topic become the focus of linguists, especially psycho- and historical linguists. This newly awoken interest manifests itself, for instance, in a recent issue of *Theoretical Linguistics* devoted to the relationship between priming and language change (cf. Jäger & Rosenbach 2008a).[35] Nonetheless, Jäger and Rosenbach (2008a:108) emphasize that their discussion of priming and its relation to language change is think-piece, but not a fully fledged theory of how precisely this is to be understood.

[35] The interested reader is referred to Jäger & Rosenbach (2008a) and the sources quoted therein for a more thorough discussion of the effects of priming in the context of language change.

To illustrate: if priming is neglected, then particularly high frequencies of the discourse marker LIKE, for example, within a given dialogue would almost certainly be attributed to extra-linguistic features of the speaker such as his or her gender or age. This may lead to an erroneous interpretation according to which the elevated frequencies reflect ongoing change, while the more appropriate explanation may well be priming. This is, however, not to say that the two explanations are mutually exclusive, but rather complementary. In fact, increased frequencies of a given form among a population of speakers, i.e. the situation we observe in the case of ongoing change, will increase the probability of priming.

Accommodation refers to the fact that during conversation, the linguistic styles of interlocutors incline to each other (Le Page 1998:22). This phenomenon has primarily been addressed with respect to the observer's paradox, i.e. the fact of observation changes the behavior of the observed entity. In other words, the linguistic behavior of interviewees may converge with the speech of the interviewer (Le Page 1998:22). The use of LIKE by one interlocutor may lead to its use by the other interlocutor due to accommodation rather than extra-linguistic factors such as age and gender. If accommodation is neglected, then the effect of extra-linguistic variables may be overestimated.

Idiosyncratic overuse denotes abnormally high rates of LIKE by individual speakers. Such speakers are similar to yet different from outliers, as their frequency is notably elevated, although not high enough to justify the elimination of this speaker from the data set. This bias towards employing vernacular LIKE is best attributed to the speaker's idiosyncratic linguistic performance rather than to his or her age or gender. Again, neglecting idiosyncratic bias would lead to overestimating the effect of extra-linguistic variables. Hence, if no measure for idiosyncratic performance was implemented, the multifactorial model would, a priori, account for the abnormally high rate as resulting from extra-linguistic factors, and their effect size would necessarily be over-estimated.

In order to account for priming, accommodation and idiosyncratic overuse, the analysis of individual subtypes of LIKE includes the cumulative frequency of all other LIKEs present in the dialogue. Although this cumulative frequency, the PAI index, reflects priming rather poorly, it serves to provide a data-driven approximation of the likelihood of priming effects and, hence, protects against over-estimating the effect of time, i.e. the rate of ongoing change. Note, however, that the PAI index – the combined measure of priming, accommodation and idiosyncratic overuse – is only a rough approximation. As illustrations of these shortcoming, keep in mind that the PAI index considers neigther phonetic nor phonological similarity, nor does it reflect the time interval between previous uses of LIKE and the individual target occurrence. In this sense its primary function is to prevent over-estimation of other extra-linguistic factors rather than being a meaningful factor in itself.

The PAI index is operationalized as a numeric, ratio-scaled variable which represents the normalized per-1,000-words-frequency of all instances of LIKE by all speakers combined in the respective conversation.

Table 16: Operationalization of the PAI index in this analysis (PAI)

Variable name	Description	Variable type	Levels	Definition
PAI	Priming Accommodation Idiosyncratic bias	numeric, ratio scaled		Combined frequencies of all other variants of LIKE in the dialogue (not applicable in the case of ALL being the dependent variable)

5.9 Statistical design

The following section introduces and discusses the main concepts which are important for understanding and interpreting the results. The first part of this section focuses on very general concepts which are not specific to the present study, while the second part concerns itself with concepts which are rather specific to the type of analyses employed here.

5.9.1 General remarks about quantitative analyses and statistics

Statistical analyses are more powerful than mere visual inspection, and they are a crucial and powerful tool in differentiating between real patterns and imaginary patterns caused by pareidolia. The detection of real patterns or real correlations rests on the following logic: any p-value-based statistic[36] aims to quantify the confidence with which we can assert that there is a real correlation between one or more independent predictor variables and the dependent variable. The assumption that there is a correlation is called the test-hypothesis, since this is the hypothesis that we aim to evaluate. However, we assume initially that a pattern is not real or, in other words, that there is no systematic correlation between an independent and the dependent variable. This initial assumption is called the null hypothesis. The statistical tests then provide a probability value, or p-value, which denotes the probability of obtaining the results[37] (test statistics) given that there is no correlation, i.e. that the null-hypothesis is true. Only if this probability for obtaining the results – given the premise that the null-hypothesis is true – is very small, we will then be allowed to assume that there is a real effect. In other words, if the results are very unlikely given that there was no real correlation, then we regard this as evidence for the existence of a real correlation and switch to the test hypothesis.

The likelihood of the null hypothesis is also called the level of significance and commonly provided in the form of asterisks (stars and points) to the left of the relevant test-statistic. The asterisks indicate that the probability of the result; given that the null hypothesis is true is less than 10 percent, 5 percent,

---

36 In contrast to p-value-based statistics, Bayesian inference includes both the null-hypothesis and the test-hypothesis and determines the degree to which the current results speak for either of these hypotheses and the degree to which the results are caused by type I and type II errors (cf. Eid, Gollwitzer & Schmitt 2010: 158-163, 195).

37 In fact, the p-value actually represents the probability of obtaining results which are at least as extreme as the present results given that the null-hypothesis is true.

1 percent or 1 per mille. The present analysis adheres to this convention when providing test-statistics (>.1=·+; >.05=*; >.01=**; >.001=***; n.s.=not significant).

In addition to evaluating the reality of correlations, statistical tests serve to uncover trends and interactions which have previously escaped detection. The main objective of the statistics employed here, however, is to determine the strength and direction of correlations between extra-linguistic variables and the frequency of (functionally distinct uses of) LIKE per 1,000 words.

An appropriate design for this purpose is a regression analysis. All forms of regressions are derivations of the most basic type of regression, the linear regression in the form $y_i = \beta_0 + \beta_1 X_1 + \varepsilon$. The type of regression employed here is a subtype of regression models which allows the inclusion of several predictor variables and their interactions into a single model. This way, regression models out-perform various other designs which are incapable of handling more than one predictor variable and one dependent or outcome variable at the same time. Thus, the regression models used here are multifactorial (or multivariate) and follow the common form $y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n + \varepsilon$, where y_i is the dependent variable (frequency of LIKE per 1,000 words), β_0 the intercept, i.e. the point where the regression line cuts the y-axis, X_1 refers to a predictor variable, e.g. AGE, X_2 refers to a second predictor variable, e.g. SEX, X_n to the n^{th} predictor variable. The ε is an error term which represents unexplained variance. The design which best fits the present purpose is a generalized linear regression model allowing to determine whether any of the independent variables influence LIKE use significantly and to what degree (effect size). In contrast to ordinary least square regression models (OLS) which try to minimize the squared values of the residuals, i.e. the difference between observed and expected values, generalized linear models iteratively adjust the coefficients of the predictor variables, i.e. the values of β_1 to β_n, to maximize the fit between observed values (y_i) and expected values (i.e. the values predicted by the model) (Baayen 2008:214-215). Another advantage of generalized models is that they allow for the implementation of

various subtypes of regressions, while OLS designs are more restrictive with respect to which regressions may be implemented in this design.

5.9.2 Statistical concepts specific to the present analysis

In the present context, we are dealing with count data, i.e. y_i represents non-negative integers, which follow a highly non-normal distribution. There are several possible procedures which would be appropriate in this case. The three most viable options which have been employed in the present study are the *Poisson Regression* (PR), the *quasi-Poisson Regression* (QPR), and the *Negative Bionimal Regression* (NBR). They are described and presented below.

5.9.2.1 The Poisson regression

Although not often used in linguistics, the Poisson regression (PR) is the most prototypical regression analysis when dealing with count data. Like all regression analyses, PRs are employed to detect possible correlations between the independent predictor variables and the dependent numerical variable (Baayen 2008:295-300). The great advantage of PRs is that they are capable of dealing with extremely non-normal distributed data, i.e. it is a non-parametric design. While the vast majority of regression models, whether they are linear or logistic, require normality of the distribution of the values of the dependent variable, PRsdo not require this normality of residuals as they pre-suppose and are thus based on a certain type of distribution (the Poisson distribution). The Poisson distribution describes the occurrence of rare events and is, therefore, perfectly suited to analyze cases in which the phenomenon under investigation is not present in the speech of every speaker, but only in the speech of few (Baayen 2008:296).

PRs have three main advantages compared to other non-parametric designs: Firstly, PRsare not based on ranks such as, for example, Mann-Whitney U tests, non-parametric T-tests, Wilcoxon Signed Rank tests, Kruskal Wallis, or Friedman tests. This offers the advantage that the information

contained in the values of the dependent variable is not as heavily reduced as it is in the case of the non-parametric tests mentioned above.

Hence, PRs – like all regressions – are able to include various independent predictor variables and interactions between them and, thus, out-perform non-parametric designs which cannot handle more than one predictor.

Secondly, Poisson regressions require the dependent variable to represent counts, i.e. non-negative intergers, as they are designed to analyze the occurrences of a phenomenon in a given interval, e.g. per 1,000 words. With respect to the PR, the expected frequencies of occurrences within a given interval are then evaluated with respect to their correlation with various independent predictor variables. In other words, the PRs test whether a certain variable leads to a distribution of occurrences which would significantly differ from the expected distribution if this variable had no influence on the occurrence of the phenomenon.

Thirdly, in contrast to (hierarchical) configural frequency analyses, i.e. (H)CFA designs, PRs are capable of accounting for the performance of individual speakers. Accordingly, PRs perform better than CFA because they are able to take the distribution of forms within a sample into account. For example, imagine two data sets, A and B: both data sets represent 20 speakers, both data sets consist of 100,000 words, and in both data sets we find 100 tokens of the linguistic item we want to study. In data set A, however, all 20 speakers use the item 5 times, while in data set B only a single informant is responsible for all 100 instances. A (H)CFA design is unable to detect the difference between the two data sets, while a PR would easily report that the investigated item is distributed homogenously in sample A, but heterogeneously distributed in data set B.

In order to apply this regression model, the normalized per 1,000 word frequencies are rounded to zero decimals in order to best match the fact that Poisson regressions are meant to evaluate counts, i.e. absolute numbers of occurrences, within a given interval. This procedure is not optimal, as the

frequencies already represent the estimated number of occurrences of LIKE per 1,000 words (the interval in the present study).

A disadvantage of PRs is that they presuppose that the observed values follow a Poisson distribution which they rarely do. The Poisson distribution is defined by the fact that its mean is equal to the variance, i.e. $\lambda = E(X) = Var(X)$. In this ideal case, the dispersion parameter is 1, since $E(X) / Var(X) = 1$. Unfortunaely, this assumption is almost always unjustified, as the variance is typically greater than the mean (http://www-m4.ma.tum.de/nbu1/modreg/PoissonReg.shtml). The state in which the mean is greater than the variance is referred to as overdispersion, i.e. the variances of values exceed their expected values ($\lambda>1$), and to account for over-dispersion, i.e. when the variances of values exceed their expected values ($\lambda>1$), the option *family=quasipossion* is implemented to loosen "the requirement that the dispersion parameter should be close to 1" (Baayen 2008:298). The *quasipoisson* option accounts for this shortcoming and causes the estimation of the effects of predictor variables and their interactions to be more accurate.

The initial saturated model contains only age and gender as main effects as well as their interaction. Subsequently, this saturated model is fitted in a stepwise procedure to arrive at a minimal adequate model that contains the minimal number of significant predictor variables and interactions (Baayen 2008:181-185). Thus, only when age or gender are marginally significant ($p<.1$), or are part of at least marginally significant interactions, do they remain in the intermediate minimal model. Although step-wise model-fitting is not unproblematic (Johnson 2010), it is the most convenient measure to prevent overfitting (including irrelevant variables). Analogous to underfitting, overfitting leads to either missing existing correlations or to misperceiving their strength (Eid, Gollwitzer & Schmitt 2010:678). To prevent the exclusion of significant factors, an ANOVA (analysis of variance) is used to determine whether the more saturated model differs significantly from the reduced model. According to Occam's razor, models which rely on fewer predictor variables are to be preferred when everything else is equal. Hence, if two

models do not differ significantly in explanatory power, that model which contains fewer predictor variables is better and, therefore, preferable. After eliminating all irrelevant predictor variables which are not part of significant interactions, this procedure leads to the first minimal adequate model. In a last step, predictor variables or interactions are eliminated from this minimal model if their coefficients are insignificant. This last step is advisable, as insignificant predictors reflect a minute correlation of the predictor variable, which means that the effect they have on the dependent variable is probably too small to be meaningful. There is, however, an exception to this procedure: If the variable is particularly relevant in terms of its theoretical implication, the variable remains in the model in spite of this insignificance. After this final stage of regression modeling, we arrive at the final minimal model (the model for the results which are reported).

Despite their obvious advantages, Poisson regressions have until now been largely neglected in quantitative linguistics and are far less common than other types of regression models such as logistic regression models or simple OLS designs. Unfortunately, this means that Poisson regressions do not as yet provide model fit values comparable to Nagelkerke's R^2. In order to evaluate the performance of the Poisson regressions, their predicted values are compared to base-line models which predict the expected values of the dependent variable solely based on the intercept. If the Poisson regression out-performs the base-line models, the test statistics of this comparison are provided. Outperforming the base-line model means that the (Quasi-) Poisson regression provides more accurate predictions, which implies that it is statistically significant. The appropriate test for this kind of model evaluation is an ordinary χ^2 design. Accordingly, the model performance is reported in terms of χ^2 values and significance levels. Evaluating the performance of a regression model with a base-line model only works when the intercept is positive, which is not always in case over-dispersion is not fully accounted for. In such cases, comparing the regression model with a base-line model is futile and will not be considered.

In addition to multivariate designs, the present analysis uses χ^2 statistics and non-parametric t-tests, which do not require variance homogeneity, to confirm or evaluate certain correlations. Using different statistical test-designs is very useful, not only because it allows validating correlations but it also offers additional perspectives on the data. In addition, employing different statistics answers discrete questions and allows different concerns to be addressed. Hence, using multiple tests provides a more refined understanding of LIKE use.

6 LIKE across varieties of English

The following section addresses the question of the degree to which the global spread of LIKE follows similar trajectories when introduced into different speech communities. In other words, this section investigates the existence of supra-locally consistent trends on the one hand, and the degree to which LIKE is modified and adapted to the local norms and social practices on the other. The following thus provides an overview of the use of LIKE in the distinct geographical settings of the respective regional varieties of English. The underlying aim of this chapter is thus to display and describe the social distribution of LIKE use to allow an empirically corroborated explanation of how and why the use of this vernacular form varies, or remains consistent, in distinct regional settings. Hence, the results of these analyses are intended to inform about the interplay of local social practice and globalized language change. This relates especially to testing the consistency of re-occurring tendencies of language change and variation established in sociolinguistic research over the past thirty years.

In addition to providing a general overview of LIKE use in the combined data, the following chapter analyses LIKE use in highly comparable datasets from eight geographically distinct regional varieties of English. This kind of approach is promising, as previous sociolinguistic analyses of linguistic behavior of age and gender have either focused exclusively on merely one or two regional varieties simultaneously, or used data sets which were not designed to guarantee comparability.

The subsequent survey combines graphical displays and multivariate analyses evaluating the correlation of LIKE and its subtypes with language-external, social variables. The Labovian framework, which constitutes the theoretical underpinning of this study, is adapted to match the requirements of both the cross-variety comparison as well as the more fine–grained analysis with each regional variety. In addition to Labov (1994, 2001), the approach taken here owes much to Eckert (1998, 2000, 2001) and to Buchstaller and

D'Arcy's (2009) analysis of *be like* in three discontinuous settings as theoretical and analytical models.

6.1 Surveying LIKE across varieties of English

Though lacking the depth provided by more detailed analysis of LIKE use within varieties of English of the subsequent sections, this section provides a survey of general tendencies of LIKE use which emerge from the combined data. Thus, throughout the following sections, we will identify the social and geographical factors that predict high and low frequencies of LIKE use. The underlying aim of identifying significant social and geographical factors is to enquire whether certain trends which have been claimed with respect to LIKE use in the respective literature emerge on a global level. These general tendencies may be related to and thus inform about the time when and under which circumstances the discourse marker LIKE emerged or entered different regional varieties; about the overall frequency of LIKE across varieties; about possible gender and age differences in its use; and the distribution and frequency of the types of LIKE.

Before analyzing the data in detail, we will look at the basic statistics of the data to get a first impression of its structure. Of the 1,974 speakers who have contributed to the data, 1,094 use LIKE at least once, i.e. over half of the speakers make use of this marker. This surprisingly high value clearly demonstrates that LIKE has spread across the globe and is prevalent as a functional linguistic item in all varieties of English included in the present study. Table 17 reports the basic statistics of the combined data and provides a first, though rather rough, impression of LIKE use across varieties of English.

Table 17: Overview of LIKE use across varieties of English

	N (total)	**Mean (per 1,000 words)**	**Median (per 1,000 words)**	**Standard deviation (s)**
LIKE	4661	2.469	0.814	3.795

Table 17 shows that LIKE is used on average 2.469 times per 1,000 words (mean) with the median being substantially lower, i.e. 0.814. The difference between mean and median indicates that LIKE use is not distributed homogenously across varieties and speakers, but that most speakers use LIKE occasionally or not at all, while a substantially smaller subset of the speakers makes frequent use of this discourse marker. Figure 11 supports this interpretation, as it reports that the vast majority in the present data have used LIKE only once, or not used it at all.

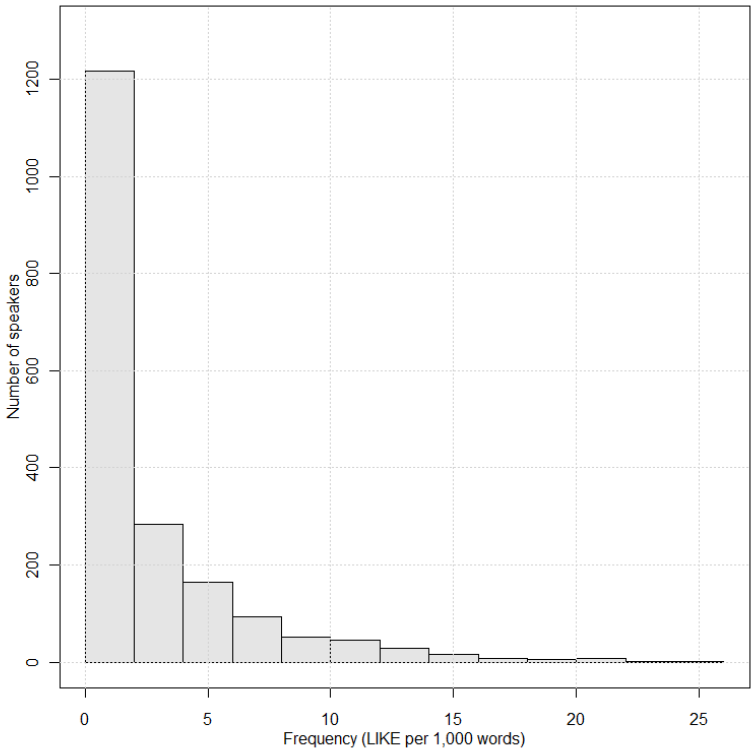

Figure 11: Histogram correlating the number of speakers with their rate of LIKE

The distribution reported in Figure 11 calls attention to the fact that LIKE use is distributed highly non-normally, which means that it does not follow a bell curve. However, this distribution approximates the Poisson distribution, which justifies the use of Poisson regression models in the subsequent elaborate statistical analyses. Concerning the present statistical analysis, it follows that parametric statistical tests are inadequate to analyze the present

data.[38] Figure 12 informs about how and to what degree LIKE use differs across varieties of English.

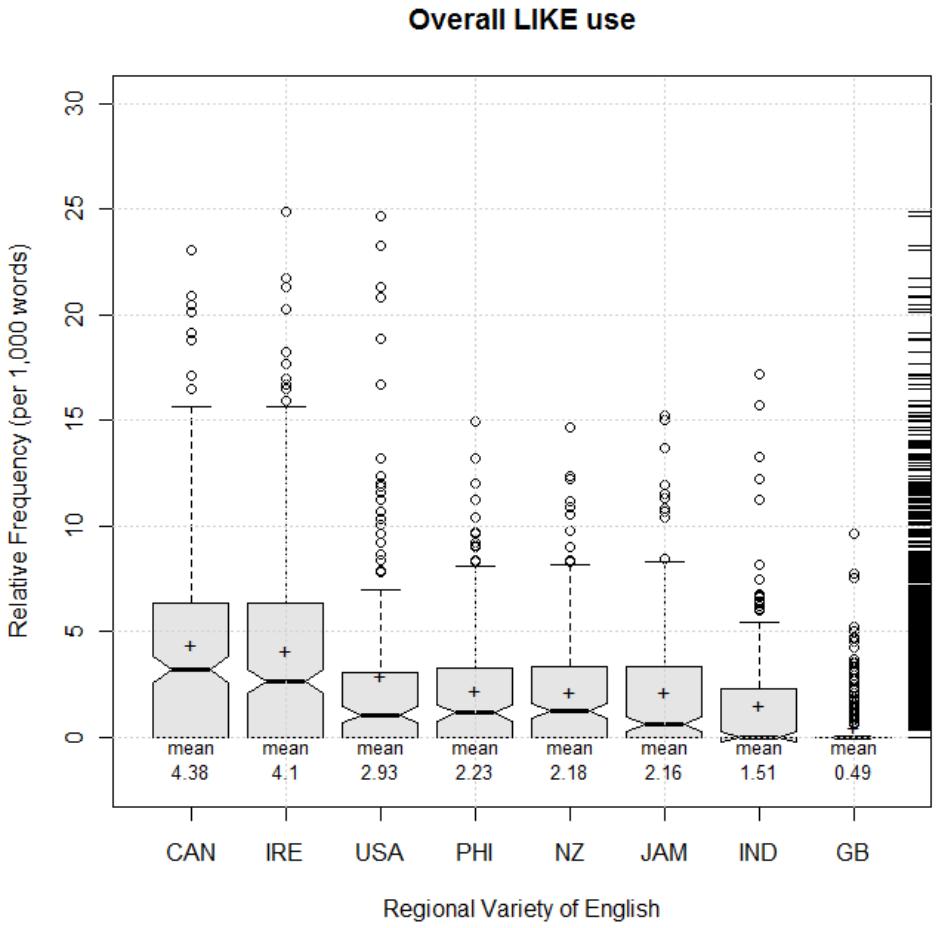

Figure 12: LIKE across varieties of English (in decreasing order according to their mean frequency)

The box plots in Figure 12 display the frequency of LIKE use by regional variety. Indeed, LIKE use differs notably across varieties, illustrating that the location strongly affects LIKE use. IrE and CanE show the highest frequencies

[38] The non-normality of the data is in fact not unproblematic, as most statistics employed in linguistic analyses – with exception of χ^2-tests – would be inadequate in the present case. The appropriate statistical tests which account for such high non-normality are a certain subtype of regression models which account for the fact that the phenomenon under investigation occurs infrequently. Section 5.9 offers a more detailed discussion of the appropriate statistical methods employed in the present data analysis.

of LIKE per 1,000 words. Despite their similarity in terms of average LIKE use, these two varieties exhibit an interesting difference: although the two varieties have almost identical means, 4.38 and 4.10 respectively, LIKE use in the Canadian speech community is slightly more homogenous than it is within Ireland: the median is closer to the mean, indicating a lower degree of dispersion than in the Canadian data. Surprisingly, the regional variety with the lowest frequency of LIKE per 1,000 words is EngE with a mean of only 0.49 instances of LIKE per 1,000 words and a median of 0.0, which indicates that the majority of EngE speakers do not use LIKE at all. This strongly suggests that LIKE in EngE is infrequent even in the present, highly informal register. Furthermore, the box plots in Figure 12 indicate which varieties differ significantly with respect to their median of LIKE use per 1,000 words: if the indentations on the left and right of two box plots, i.e. their notches, do not overlap, then these varieties will probably not differ significantly with respect to their frequency of LIKE use (Gries 2009:119). It follows that IrE and CanE do not differ significantly with respect to their relative frequency of LIKE, while both of them differ significantly from all other varieties which employ LIKE at a significantly lower rate.

While the box plots in Figure 12 display the basic characteristics of LIKE use across varieties of English, they fail to differentiate among the frequencies of the four types of LIKE introduced in previous section (5.7.1). Table 18 addresses this shortcoming by providing a more detailed description, according to both the regional variety and the respective type of LIKE.

Table 18: Overview of LIKE variants across varieties of English

Variety (ICE component)	**INI** per 1,000 words	**MED** per 1,000 words	**FIN** per 1,000 words	**NON** per 1,000 words	**ALL** per 1,000 words
Canada	1.78	1.93	0.11	0.56	4.38
GB	0.14	0.22	0.01	0.12	0.49
Ireland	1.03	1.01	1.51	0.49	4.10
India	0.48	0.29	0.09	0.62	1.51
Jamaica	0.55	1.18	0.04	0.35	2.16
New Zealand	0.84	0.77	0.09	0.46	2.18
Philippines	0.79	0.95	0.05	0.35	2.23
Santa Barbara Corpus	0.68	1.32	0.00	0.85	2.93
Mean	0.78	0.92	0.29	0.45	2.47

The overview in Table 18 summarizes the distribution of the types of LIKE as well as their combined frequencies across varieties of English. This more fine-grained depiction of the data makes differences immediately visible which otherwise would have escaped our attention. The most striking aspect which has emerged is the fact that the high frequency of LIKE in IrE is due to an extremely high frequency of clause-final LIKE. This finding causes us to restate the superficial similarity between CanE and IrE, as these two varieties exhibit quite distinct profiles when positional variants of LIKE are taken into account.

Other aspects rendered visible in Table 18 are the relatively high frequency of non-clausal LIKE in IndE and the preferences for clause-medial LIKE in all varieties except IrE, IndE, and NZE. To get a better understanding of LIKE use across varieties, we will for a moment disregard the frequency with which different types of LIKE are used and focus exclusively on the positional distribution. Hence, in contrast to Table 18, Table 19 depicts the percentages of each variant compared to all variants combined rather than the relative frequencies. Depicted in this way, the focus is on variety-specific usage patterns.

Table 19: Overview of LIKE variants across varieties of English

Variety (ICE component)	INI (%)	MED (%)	FIN (%)	NON (%)
Canada	40.64	44.06	2.51	12.79
GB	28.57	44.90	2.04	24.49
Ireland	25.12	24.63	36.83	11.95
India	31.79	19.21	5.96	41.06
Jamaica	25.46	54.63	1.85	16.20
New Zealand	38.53	35.32	4.13	21.10
Philippines	35.43	42.60	2.24	15.70
Santa Barbara Corpus	23.21	45.05	0.00	29.01
Mean	31.58	37.27	11.74	18.22

The percentage-based report in Table 19 emphasizes the impressions derived from the frequency-based report in Table 18. Particularly noteworthy are the high percentage of clause-final LIKE in IrE and the high percentage of non-clausal LIKE in IndE. Moreover, IndE and IrE differ notably from all other varieties with respect to their use of clause-medial LIKE: speakers of IrE and IndE appear to reject this variant compared to its use in other varieties. In addition, Table 19 shows that IndE and NZE exhibit a rather high amount of clause-final LIKE; although substantially lower than in IrE. The fact that neigther IndE nor IrE not seem to match the distributional pattern of LIKE use observed in other varieties corroborates a similarity between IrE and IndE previously reported by Siemund, Maier & Schweinberger (2009:29-30).

In summary, LIKE is a salient feature of present-day vernacular English around the globe, as it is attested in all varieties of English included in the present study. The use of LIKE is nonetheless not uniform, but differs markedly in distinct locations: while LIKE use is abundant in CanE and IrE, it is a rather marginal phenomenon in EngE. In addition, results confirm that LIKE use varies markedly not only across, but moreso within speech communities: although over half of the speakers in the present ICE data make use of this non-standard feature, the majority of speakers use it moderately or not at all. With respect to differences in the use of distinct forms of LIKE, clause-medial LIKE

is the most widely and most frequently occurring variant, while clause-final LIKE is extremely frequent in IrE, but almost negligible in nearly all other regional varieties.

6.1.1 Clause-initial LIKE

The following sections will elaborate on the data presented so far by focusing on each LIKE variant individually, paying particular attention to the context of its cross-varietal distribution.

The clause-initial discourse marker is the first variant of LIKE to be discussed in greater detail. Functionally, this variant introduces explanations, elaborations or specifications of preceding utterances or clausal constituents, as in (70). In addition, Miller (2009:330) also attests a highlighting function with respect to the clause which follows, although he admits that its primary function is that it "signals this clause [the elaboration headed by LIKE] as an addition to the preceding statement and an explanation [of the previous statement]".

(70) a. [B]ut *like* he went there now and we use to go there a couple of times. (ICE-New Zealand:S1A-015$A)

b. *Like* I can't make the usual adobo I'm not good at that nilaga but my Mom makes fabulous there's nilaga waiting for me at home right now. (ICE-Philippines:S1A-062$B)

c. But I don't know I think for me <,> *like* I mean once you learn a classical style you you're kinda kind of set in that in that mode. (ICE Canada:S1A-010$B)

Accordingly, clause-initial uses of LIKE share functional properties with certain uses of non-clausal LIKE, but differ with respect to the clausal status of the subsequent elaboration. Clause-initial LIKE is a prototypical discourse marker in the sense that it serves to enhance discourse cohesion (Schiffrin 1987:57-59), while non-clausal LIKE – if not used to buy processing time– is used to introduce phrasal specifications which do not form complete clausal structures. In contrast to both clause-medial LIKE and clause-final LIKE, clause-initial LIKE has not received similar attention and has mostly been disregarded

by scholars. This is unfortunate, as Figure 13 reports that this prototypical discourse marker is fairly frequent in almost all varieties of English investigated in the present study.

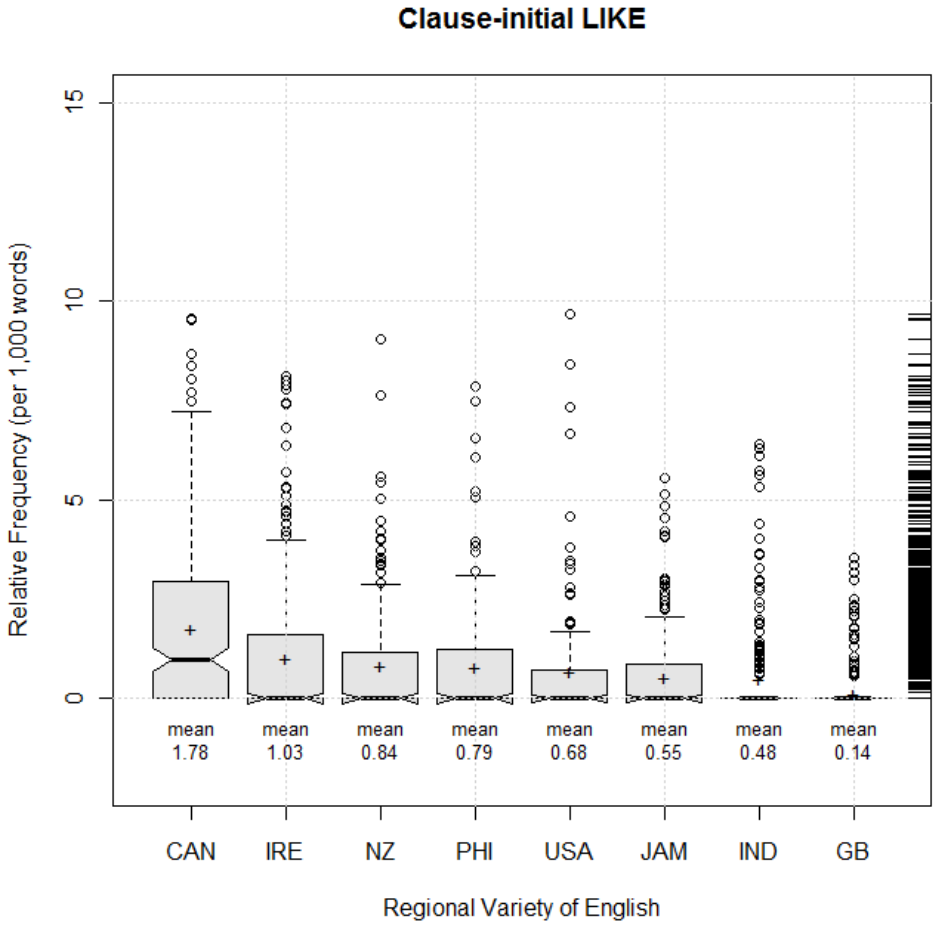

Figure 13: Clause-initial LIKE across varieties of English (in decreasing order according to their mean frequency)

According to Figure 13, the clause-initial discourse marker appears to be particularly frequent in CanE, and to a somewhat lesser extent in IrE, while it is only a marginal discourse feature in EngE. As the display of Figure 13 fails to take the overall frequency of LIKE use into account, it may obscure variety-specific preferences. Therefore,

Table 20 reports both the frequency and the proportion of clause-initial LIKE.

Table 20: Overview - clause-initial LIKE

Variety (ICE component)	**INI** per 1,000 words	INI (%)
Canada	1.78	40.64
GB	0.14	28.57
Ireland	1.03	25.12
India	0.48	31.79
Jamaica	0.55	25.46
New Zealand	0.84	38.53
Philippines	0.79	35.43
Santa Barbara Corpus	0.68	23.21
Mean	0.78	31.58

Patterns immediately emerge which remained undetected in Figure 13. While CanE still exhibits the strongest preference for this type of LIKE, it is NZE, PhiE, and to a lesser degree IndE, which also show elevated usage rates. The mean value of the percentages inform us that clause-initial LIKE is responsible for slightly little less than one third of all instances of LIKE. Indeed, the rates differ notably across varieties of English: while it accounts for more than forty percent of all instances of LIKE in CanE, clause-initial LIKE represents only about twenty percent of LIKE uses in AmE. This difference is unexpected and quite interesting, as it shows that LIKE has been integrated differently into communal grammars, even in geographically proximate varieties of English.

To summarize, clause-initial LIKE is present in all varieties of English though its rate differs notably. On average, instances of clause-initial LIKE account for about one third of all instances of LIKE, but its rate can exceed forty percent as in CanE.

6.1.2 Clause-medial LIKE

According to Miller (2009:332), clause-medial LIKE as in (71) is “possibly the most familiar” variant of all vernacular uses of LIKE.

(71) a. Oh no yeah you *like* have two more weeks before that. (ICE Philippines:S1A-013$A)

b. Cos he just won a place to *like* <,> Canterbury Cathedral Choir School. (ICE-Canada:S1A-051$A)

c. Yeah but it's supposed to be *like* the happiest day of your life. (ICE Jamaica:S1A-063$B)

This so-called 'innovative' clause-medial variant of LIKE is well attested across varieties. Both Table 18 and Table 19 confirm a general tendency of clause-medial LIKE to be the most frequent variant of LIKE across all varieties, with the exception of IrE, IndE, and NZE. It is probable that the functionality of clause-final LIKE, i.e. its ability for clause-internal modification (focusing and hedging of constituents), causes its high frequency. This cross-varietal tendency is particularly striking, as this variant has been described as an innovative type of LIKE that is entering varieties of English as an American borrowing (Andersen 2000:287). With respect to language change, it is this clause-medial type of LIKE that is expected to exhibit those (sociolinguistic) patterns associated with and prototypical of ongoing linguistic change. The box plots in Figure 14 display the frequencies of clause-medial LIKE across varieties of English.

Figure 14 qualifies assertions according to which clause-medial LIKE is solely "a feature of North American English" (D'Arcy 2007:390), as it reports substantial use of clause-medial LIKE in all areas with the exception of IndE and, as expected, EngE. To prevent arriving at conclusions prematurely, Table 21 reports not only the frequencies, but also the percentages of clause-medial LIKE (relative to all uses of LIKE in the respective location).

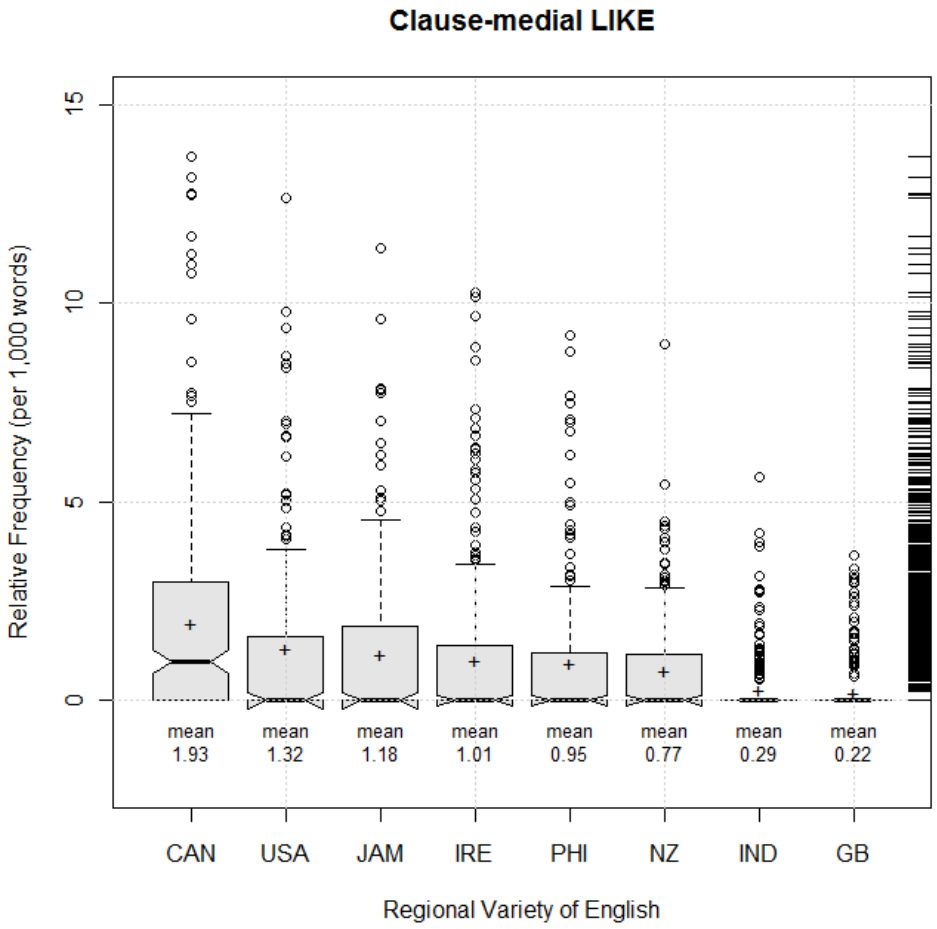

Figure 14: Clause-medial LIKE across varieties of English (in decreasing order according to their mean frequency)

Table 21: Overview - clause-medial LIKE

Variety (ICE component)	**MED** per 1,000 words	MED (%)
Canada	1.93	44.06
GB	0.22	44.90
Ireland	1.01	24.63
India	0.29	19.21
Jamaica	1.18	54.63
New Zealand	0.77	35.32
Philippines	0.95	42.60
Santa Barbara Corpus	1.32	45.05
Mean	0.92	37.27

The alternate display of Table 21 offers surprising insights into the variety-specific use of clause-medial LIKE. For example, it is striking that clause-medial LIKE is the most frequent LIKE variant in all varieties except IrE, IndE, and NZE. In all but these three varieties, clause-medial LIKE accounts for well over forty percent of all instances and even exceeds fifty percent in JamE. Despite being notably infrequent, the distribution of clause-medial LIKE in EngE converges to the group B pattern of LIKE use attested for in PhiE and East African English (EAE) in Siemund, Maier & Schweinberger (2009:29). According to Siemund, Maier & Schweinberger (2009), their selection of regional varieties of English, comprising IndE, IreE, PhiE, and EAE, could be classified as either group A varieties (LIKE occurs predominantly in clause-marginal position), or as group B varieties (LIKE occurs predominantly in clause-medial position). Schweinberger (2010) applied a similar approach to an extended selection of regional varieties and corroborated the validity of the group A vs. group B distinction. He further hypothesized that the distributional pattern of group B varieties is probably related to its use in AmE. The results of the present study confirm Schweinberger's (2010) assertion that both AmE and EngE approximate the typical group B distribution: in both varieties, clause-medial is the most frequent variant, followed by clause-initial LIKE which is used in a lower but still substantial proportion.

In addition, NZE which has not been included in Siemund, Maier & Schweinberger (2009) nor Schweinberger (2010), seems to form an intermediate link between the dichotomous patterns of group B (PhiE and EAE) and group A varieties (IndE and IrE). According to Siemund, Maier & Schweinberger (2009), the most distinctive feature of group B varieties is the high proportion of clause-medial LIKE, while the IrE and IndE distribution reflects a strong preference for non-clausal and clause-final LIKE. Taking NZE into account, this description may require refinement, as this regional variety apparently represents an intermediate in-between these two poles. This refined description thus assumes not two distinct groups of varieties, as

proposed by the data described in Siemund, Maier & Schweinberger (2009), but a continuum.

6.1.3 Clause-final LIKE

After reviewing the distribution of the so-called innovative clause-medial variant of LIKE this section focuses on the so-called archaic clause final use of LIKE. Clause-final LIKE as in (72) has been associated with traditional dialects of northern and Celtic varieties of English (Hedevind 1967:237, cf. also Andersen 2000:222), in which it is used "parenthetically to qualify a preceding statement" (Andersen 2000:222).

(72) a. [H]e <,> this yeah i just came up *like*. (ICE New Zealand:S1A-096$B)
b. He goes to Madras <,> yeah <,> for the weekend <,> so he is busy *like*. (ICE India:S1A-059$A)
c. And you've already got a couple of them *like*. (ICE Ireland:S1A-015$A)

Figure 15 below provides an overview of the use of clause-final LIKE across selected varieties of English.

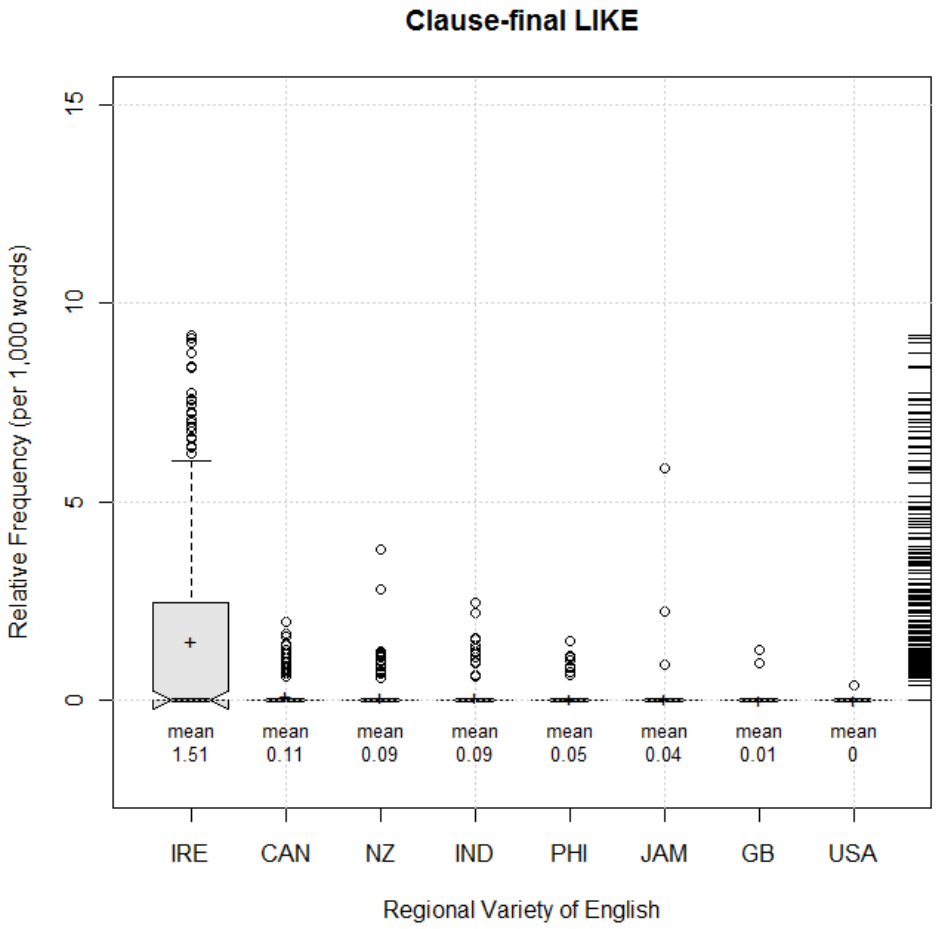

Figure 15: Clause-final LIKE across varieties of English (in decreasing order according to their mean frequency)

Figure 15 supports accounts in which LIKE with backward scope is depicted as a typically northern phenomenon. Indeed, the frequency of clause-final LIKE in IrE, as a representative of a Celtic English variety, vastly outnumbers its frequency in all other regional varieties. However, Figure 15 also illustrates that clause-final LIKE is not limited to IrE, as the data of all varieties of English examined here contain instances – although few – of this rather exotic form of LIKE. In fact, the AmE data contain only a single instance of clause-final LIKE. This is particularly noteworthy, as the tripartite categorical classification, deployed in Kortmann et al.'s "Handbook of Varieties of English", is not suitable for coping with infrequent features.

Table 22: Overview - clause-final LIKE

Variety (ICE component)	**FIN** per 1,000 words	FIN (%)
Canada	0.11	2.51
GB	0.01	2.04
Ireland	1.51	36.83
India	0.09	5.96
Jamaica	0.04	1.85
New Zealand	0.09	4.13
Philippines	0.05	2.24
Santa Barbara Corpus	0.00	0.00
Mean	0.29	11.74

Table 22 confirms the impression derived from Figure 15: IrE clearly stands out with respect to both the mean frequency and the proportion of clause-final LIKE. Despite most likely emerging in the British Isles between the late eighteenth and early nineteenth century (cf. D'Arcy 2005, 2007; Schweinberger 2013), clause-final LIKE is almost non-existent in the British data. It has to be borne in mind, however, that the British component of the ICE contains mostly EngE spoken in the London area by mostly educated speakers. Hence, regional dialects in which clause-final LIKE may still be in use – e.g. northern English dialects where it "is an emphatic device" (Beal 2004:136), i.e. as a "reinforcing element of right-dislocation (I'm Geordie, me, like)" (Kortmann 2004:1100) – are not reflected in this data.

The absence of clause-final LIKE in dialects of the London area nonetheless suggest that this form was introduced to Ireland when it was still commonly in use in Britain and subsequently survived in IrE as a fossilization or an archaism which subsequently became extinct or marginalized in the standard varieties of mainland Britain.

6.1.4 Non-clausal LIKE

The last form of LIKE to be surveyed in the present context is non-clausal LIKE. In contrast to clause-medial LIKE, it is not integrated into the syntactic

structure of the clause and functions as a hesitational device, or serves to indicate that what follows is a phrasal, non-clausal specification of a preceding element. The latter use of non-clausal LIKE clearly shares properties with clause-initial LIKE. But while the clause-initial discourse link indexes that what is about to follow is an exemplification or an elaborate specification of the preceding utterance or clausal constituent, non-clausal LIKE introduces phrasal specifications below a clausal level of complexity. In addition, such non-clausal instances of LIKE occur when speakers want to hold the floor but do not find the adequate expression. In such cases, LIKE buys processing time and is most accurately described as a filler, indicating that the speaker wants to continue but is currently searching for the adequate expression.

The distribution of non-clausal LIKE across varieties of English appears rather chaotic (cf. Figure 16). This seemingly unsystematic ordering may stem from the fact that this variant has an even higher degree of syntactic flexibility than related variants of vernacular LIKE.

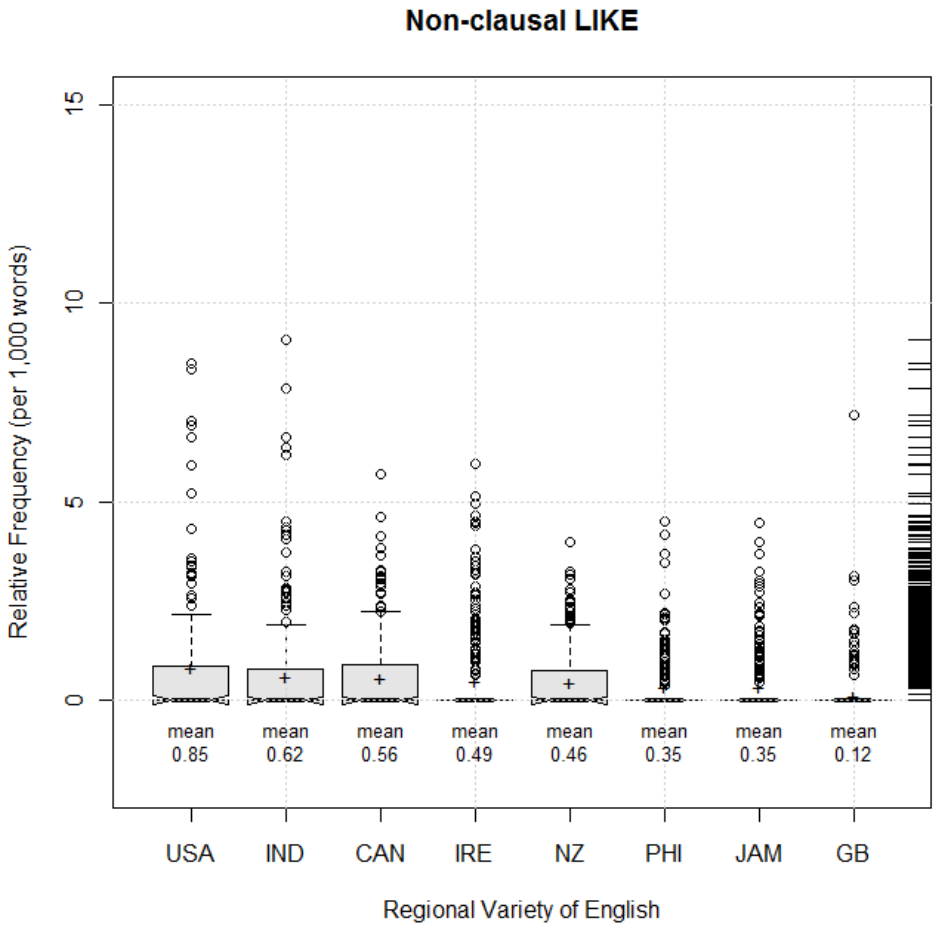

Figure 16: Non-clausal LIKE across varieties of English (in decreasing order according to their mean frequency)

Table 23: Overview - non-clausal LIKE

Variety (ICE component)	**NON** per 1,000 words	NON (%)
Canada	0.56	12.79
GB	0.12	24.49
Ireland	0.49	11.95
India	0.62	41.06
Jamaica	0.35	16.20
New Zealand	0.46	21.10
Philippines	0.35	15.70
Santa Barbara Corpus	0.85	29.01
Mean	0.45	18.22

In general, the distribution follows the pattern of overall LIKE use with the exception of IndE and IrE. The most notable change in the sequence is IndE for which Figure 16 shows a higher frequency than expected from the distribution of general LIKE use. The Indian bias towards non-clausal LIKE is even more surprising given that all other variants of LIKE are almost negligible in this variety. A speculative explanation for this rather odd scenario might be related to incomplete acquisition. One may hypothesize that speakers of IndE lacked the necessary input to acquire the adequate pragmatic functions or constraints thereof, and have thus developed an autonomous use of this form which differs markedly from its use in other varieties. A complementary factor may have been other markers which competed with LIKE on a functional level and caused LIKE use to adapt to different environmental conditions.

6.2 The sociolinguistics of LIKE use across varieties of English

So far, sociolinguistic factors have been ignored in the graphical displays. The following section will therefore report the data from a sociolinguistic perspective, enabling evaluation of general trends of LIKE use linked to extra-linguistic social factors.

Table 24: Age and gender distribution of LIKE across varieties of English[39]

			MALE		FEMALE			
AGE	**Speakers**	**Words**	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (10 – 25)	715	615,126	615	3.04	1714	3.36	2329	3.27
2 (26 – 33)	269	242,194	206	2.47	667	4.39	873	3.74
3 (30- 39)	371	337,480	325	1.46	487	2.05	812	1.79
4 (40+)	428	389,514	179	0.91	238	1.20	417	1.07
NA	142	89,400	27	0.97	171	2.70	230	2.03
SUM	1925	1,673,714	1360	1.88	1205	2.83	4661	2.47

The data summarized in Table 24 presents the absolute numbers and relative frequencies of LIKE use across varieties of English according to the age of speakers. They strongly suggest that LIKE has indeed undergone language change, as can easily be inferred from the decrease in the frequency of LIKE use with increasing age.

Figure 17 indicates that in speakers over the age of 30, the frequency of LIKE use decreases dramatically. This decline reinforces the impression derived from Table 24 that the discourse marker LIKE has been undergoing change, as it dramatically increased in frequency in English-speaking speech communities during the late 1960s and early 1970s. In addition, the fact that LIKE is attested among speakers over the age of 50, although in moderate numbers, indicates that LIKE has been in use for longer than commonly assumed. This preliminary finding corroborates D'Arcy's (2005:4) which presents examples of the discourse marker LIKE from the OED as early as the late eighteenth and early nineteenth centuries.

[39] Speakers for whom age and gender information is not available are excluded. Twenty-three speakers could neither be classified with respect to their age and gende and were therefore removed from the analysis.

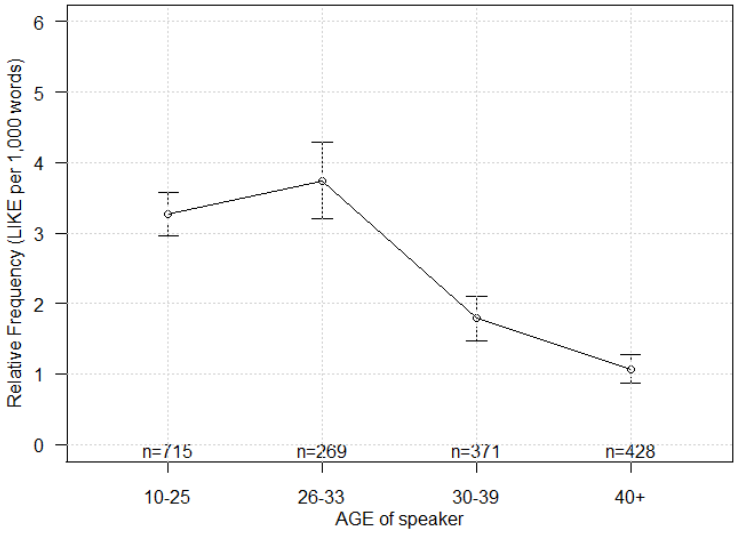

Figure 17: Age and gender distribution of LIKE across varieties of English[40]

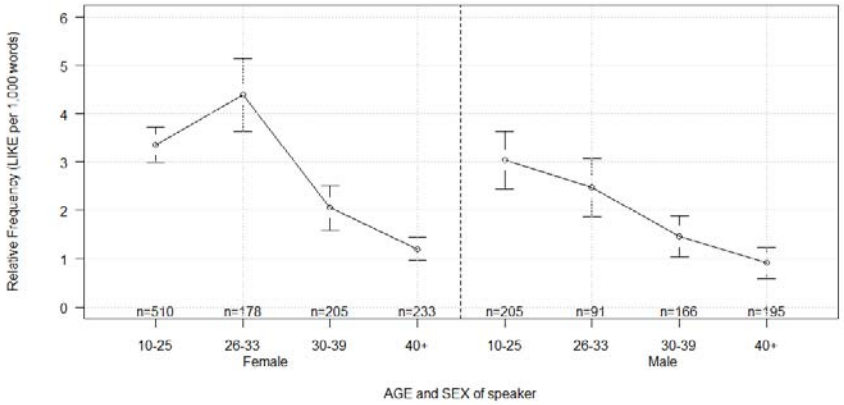

Figure 18: Age and gender distribution of LIKE across varieties of English

40 The confidence intervals (CI) included in this line graph are based on the assumption that the data follows a normal distribution, i.e. a bell curve. Unfortunately, this is not the case and the CIs are, thus, only rough approximations of the actual CIs. Nevertheless, they serve as approximate indicators of significance: if confidence intervals overlap, then this strongly suggests the absence of significant effects, while non-overlapping of the CIs is indicative of significant differences.

Before going into more detail, it should be stated that the trends visible in Figure 18 are tentative and not conclusive, as they might be affected by interactions between the regional LIKE use and the age of speakers. In other words, some age groups may be overrepresented in one ICE component where LIKE use is extremely high or low which could skew the displayed distribution.

However, Figure 18 strongly suggests that LIKE is gender sensitive particularly among speakers from their mid-twenties to mid-thirties, while there does not seem to be a gender difference among older or very young speakers. Moreover, the peak occurence among female speakers between 26 and 33 years of age challenges the popular belief that LIKE use is generally most prevalent among teenagers; the data indicate that this assumption is incorrect. Males, on the other hand, show a consistent declining pattern, with LIKE use peaking among the youngest age cohort indicating a teenager peak. The difference in LIKE use between females and males indicates that LIKE use has been initiated by or led by females, while males have apparently adopted this feature about 10 years later. This preliminary finding is interesting, as it appears to indicate that male speakers have lagged behind in a female-dominated change. Although, female speakers appear to have led the change, they are now in retreat, while male speakers appear to have cought up and have even exceeded females in the youngest age cohort. This youngest group of male speakers appears to be highly heterogeneous in terms of LIKE use, as the confidence intervals indicate a high degree of dispersion, i.e. a high degree of variance within this cohort.

Before proceeding to the within-variety analyses, it is advisable to check whether the date of data compilation corroborates the apparent-time hypothesis, i.e. if a real-time analysis supports the apparent-time analysis.

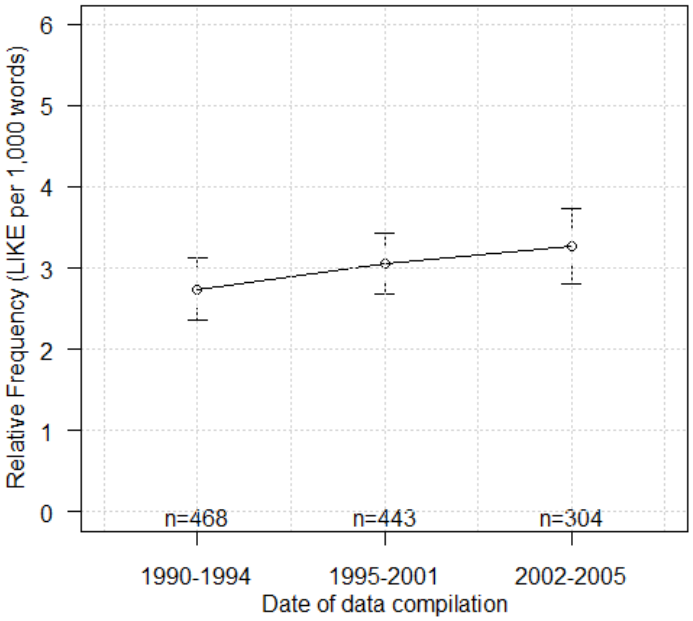

Figure 19: Mean frequency of LIKE according to the date of data compilation

Figure 19 does not report a significant difference in LIKE use with respect to the date of data compilation as the confidence intervals overlap, but a general tendency is clearly visible. There seems to be a positive correlation between the date of data compilation and the frequency of LIKE use. In other words, the frequency of LIKE has notably increased over time implying that LIKE has been undergoing change.

Table 25: Number of non-LIKE users to LIKE users and the resulting ratios according to the date of data compilation

	Mean (per 1,000 words)	**Non-LIKE users**	**LIKE users**	**Ratio**	**Percentage (LIKE users of all speakers)**
1990-1994	2.738	231	237	1.02	50.5
1995-2001	3.049	156	287	1.83	64.7
2002-2005	3.265	103	201	1.95	66.1

Table 25 confirms this impression, as not only the relative mean frequencies have increased in real-time, but also the proportion of speakers

who used LIKE in the present ICE data. The increase is indeed quite remarkable considering that the period of data compilation covers only fifteen years. Over these fifteen years, the percentage of speakers using this vernacular form has increased by more than 30 percent[41]. While only about half of the speakers used this non-standard variant in the data complied between 1990 and 1994, almost two-thirds of speakers used it in the data compiled between 2002 and 2005. Moreover, the increase is suggestive of communal change, as the rate with which the change proceeds is too high to be adequately accounted for by generational change. This is quite challenging as the apparent–time distributions do not seem to accurately reflect the rate and pace of ongoing change. Probing more deeply into the matter in the subsequent sections will provide a more accurate account of which types of change are present in the case of LIKE and, hence, help to shed light on this and related issues.

41 $^{66.1\%}/_{50.5\%\text{-}1}$ =30.9 percent

7 LIKE within varieties of English

The following chapter provides more fine-grained and in-depth analyses of LIKE use within the geographically distinct settings of the respective regional varieties of English. In contrast to the overview provided in section (4.4), the following analyses will include a broader spectrum of extra-linguistic variables and thus focus specifically on the contributive factors within regionally distinct speech communities. These more detailed examinations will serve to evaluate the extent to which the general tendencies rendered visible in the previous survey represent angloversal patterns of change, and the degree to which they underlie variety-specific constraints leading to modification and adaptation of LIKE use due to differences in socio-cultural norms and practices. Accordingly, the following chapter will assess not only previous claims concerning the use of LIKE, but also more general theoretical considerations relating to the mechanisms of linguistic change and variation, particularly with respect to the leading role of women and adolescents.

Each sub-chapter treats one regional variety and begins with introductory remarks on the respective variety of English and previous research on LIKE in this variety. In addition, each introduction describes the data and the predictor variables included in the statistical analysis.

7.1 LIKE use in US American English[42]

Besides CanE and EngE, AmE is probably the best-studied variety with respect to the discourse marker LIKE. In fact, the first professional linguists who paid attention to this pragmatic device were US American scholars (e.g. James 1983; Schourup 1982, 1985; Ross & Cooper 1979; Underhill 1988). Over the past years LIKE use in AmE has received a notable amount of scholarly attention,

[42] For a complete overview of the final data set of AmE used in this study, cf. Table 138 in the Appendix.

and it has attracted the interest of the wider public (for example Diamond 2000; Levey 1999; 1998; Mehren 1999). Nonetheless, LIKE has so far not been extensively studied from a sociolinguistic perspective and not with the elaborate quantitative means available to us today. This is rather surprising, given the "intricate lore surrounding *like*" (D'Arcy 2007:386). Examples of popular myths associated with LIKE are the wide-spread belief that LIKE is only one homogenous form and that this form is an American borrowing when used in another regional variety. Another popular myth relates to its association with the "Valley Girl" persona (D'Arcy 2007:404). Indeed, even scholars have claimed (probably erroneously) that Californian "Valley Girls" have been responsible for its spread throughout AmE (Siegel 2002:37).

Section 4, presents the notion that non-standard uses of LIKE have a long-standing history in the British Isles. Instances of LIKE dating back to the late eighteenth and early nineteenth centuries substantially predate and challenge the hypothesis that LIKE is a twentieth-century American development (cf. D'Arcy 2007:397-401; Schweinberger 2013). However, with respect to its historical development in AmE, it is not American data which offer the most intriguing insights into the location and time of LIKE's genesis in AmE, but data from New Zealand. D'Arcy (2007:401) presents examples of the discourse marker LIKE which occur in radio transcripts recorded between 1946 and 1948. Given that the vast majority of instances occurred in the English of speakers whose parents had emigrated to New Zealand from England, Ireland, and Scotland and not in the speech of speakers from the United States (D'Arcy 2007:401), the New Zealand data strongly suggest that LIKE entered the AmE speech community significantly later than those in the British Isles. Unfortunately, the problematic issue exactly where and when it emerged in AmE is still open to speculation. In fact, previous studies on LIKE use in AmE have mainly concerned themselves with its pragmatic functions, or the syntactic environments in which is occurs, while there are few studies systematically addressing issues relating to its sociolinguistic profile and spread across the North American continent. Siegel (2002), for example, claims

that LIKE "has persisted and spread in California and all over the country" (Siegel 2002:37), confirming the common myth that LIKE originated in Californian English. Siegel (2002) also reports anecdotal evidence according to which LIKE use "is most prevalent among very young women, [who] […] often seem not to be confident about their assertions" (Siegel 2002:43).

In the following, these assumptions will be tested by evaluating anecdotal evidence in light of the present data. However, before analyzing the data statistically and in greater detail, a look at the basic statistics of the summarized data will give a first impression of its structure. Of 165 speakers present in the AmE data, 110 have used LIKE at least once. This means that two thirds of AmE speakers have made use of this marker in informal speech. The high proportion of LIKE users seems to endorse the hypothesis that LIKE originated in the USA and that it has spread among varieties as an American borrowing (Andersen 2000:221-224, 286).

Table 26 reports the basic statistics of the AmE data and provides a first, rather rough impression of LIKE use in this regional variety.

Table 26: Overview of LIKE use in US AmE

	N (total)	**Mean (LIKE per 1,000 words)**	**Median (LIKE per 1,000 words)**	**Standard deviation (s)**
LIKE	860	2.928	1.037	4.734

LIKE is a highly salient feature of AmE vernacular. Although the difference between mean and median is substantial, both measures of central tendency are relatively high. For example, the median shows that fifty percent of speakers use LIKE more than 1.037 times per 1,000 words. These basic statistics provide valuable insights into the data distribution: The high proportion of LIKE users suggests that if LIKE has indeed been undergoing change, the process is probably nearing completion (Nevalainen & Raumolin-Brunberg 2003:55).

Before turning to a more fine-grained analysis of age- and gender-specific LIKE use, the frequencies of types of LIKE are displayed in order to clarify which types of LIKE are particularly frequent in AmE.

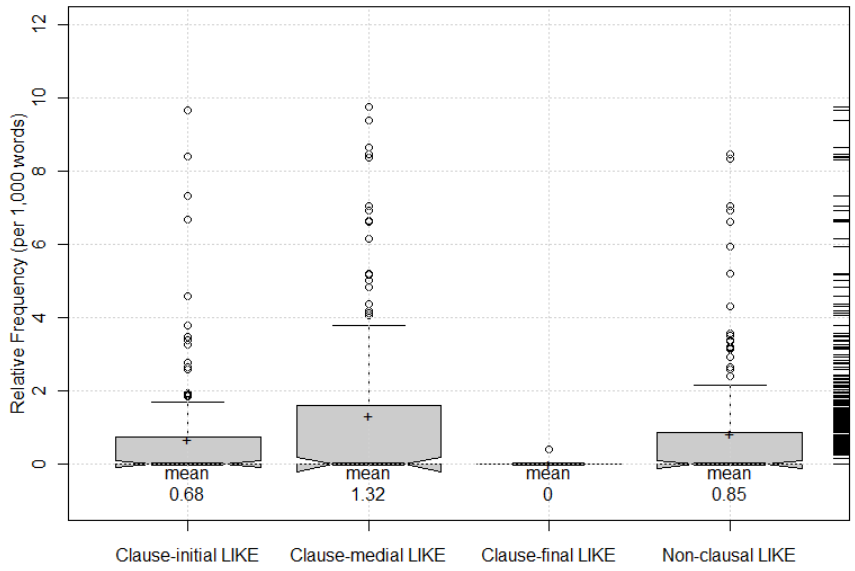

Figure 20: LIKE variants in AmE

Figure 20 shows that the most frequent variant of LIKE is – as expected – clause-medial LIKE. According to Andersen, it is exactly this clause-medial variant of LIKE which may have been borrowed by speakers of other regional varieties of English (Andersen 2000:221-224). The frequency of this variant substantially exceeds the frequencies of the other variants, particularly clause-final LIKE, which occurs only once in the AmE data. This strongly suggests that clause-final LIKE is almost non-existent in AmE. Clause-initial and non-clausal uses of LIKE are, in contrast, quite frequent in AmE, though significantly less so than clause-medial uses.

Table 27: Age and gender distribution of LIKE in AmE[43]

			MALE		FEMALE			
AGE	**Speakers**	**Words**	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (11 – 20)	19	25,391	10	1.92	174	6.02	184	5.16
2 (21 – 30)	40	62,373	93	4.32	231	5.17	324	4.79
3 (31-40)	27	52,064	118	3.14	59	3.23	177	3.17
4 (41+)	61	88,824	71	1.16	41	0.74	112	0.94
NA	16	17,606	20	3.97	34	1.90	63	2.79
SUM	163	246,258	312	2.61	539	3.12	860	2.93

Table 27 confirms the expected female lead, as women use LIKE more frequently than men in all but the oldest cohort. In addition, it reports a very distinct decrease in LIKE use with increasing age among females: this consistent decrease, however, only partially reflects male LIKE use, as males exhibit a peak in their twenties. This result is unreliable, due to the small number of speakers within this cohort. When the genders are combined, the data suggest a near-linear negative correlation between age and the frequency of LIKE. In other words, the older a speaker, the less likely it is that he or she uses this discourse feature (cf. Figure 21).

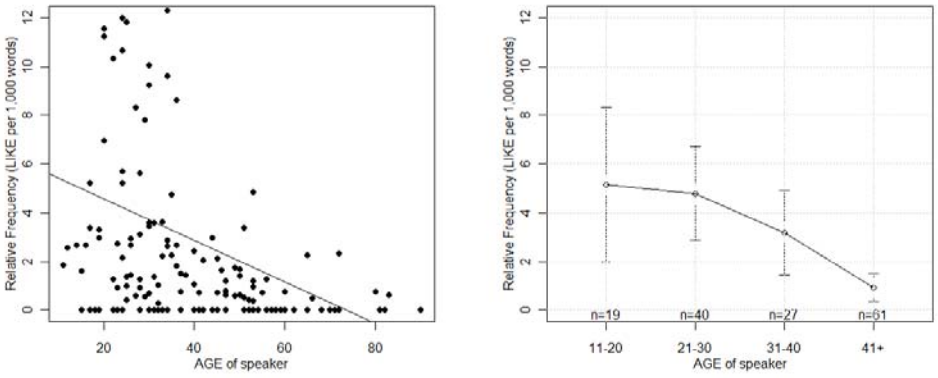

Figure 21: Age distribution of LIKE in AmE

The line graph in Figure 21 suggests a negative correlation between the frequency of LIKE and increasing age. In addition, Figure 21 indicates that

43 Speakers for whom age and gender information is not available are classified as NA.

speakers between the ages of 11 and 40 do not differ significantly in their use of LIKE, because the confidence intervals of the respective age groups overlap. The wide ranges of the confidence intervals indicate that the frequencies of LIKE use vary substantially. Indeed, the cloud of dots in the upper left of the scatter diagram corroborates this assumption. The dots represent speakers, while the position of the dot relates to the age and the frequency with which the speaker has used LIKE. Therefore, the cloud of dots to the upper left shows that it is younger speakers who make over-proportional use of this vernacular feature.

Based solely on apparent-time data, it is problematic to distinguish stable, age-based variation (age-grading) from generational change in progress (Labov 2001:77). To verify the assumption that LIKE is undergoing change, it is necessary to test whether LIKE use interacts with the socio-economic status of speakers. If this interaction exists, this would strongly support the hypothesis that LIKE use is socially stratified. Social stratification would in turn corroborate the assertion that LIKE use has not as yet diffused through all social strata.

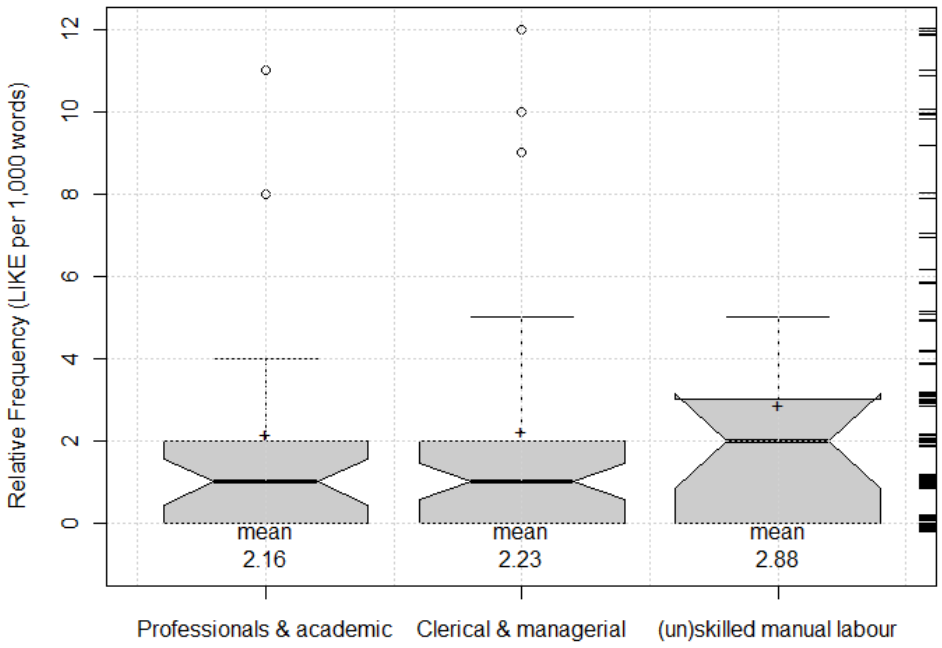

Figure 22: LIKE in AmE with respect to the occupation of speakers

Figure 21 indicates a moderate effect of social class, as the frequency of LIKE use appears to be substantially higher among speakers with lower social status. This correlation is suggestive, as different types of change are accompanied by distinct patterns of social stratification. Age-grading is commonly not accompanied by social stratification, as this type of change refers to situations in which individuals change their linguistic behavior over time while the communal grammar does not change (Labov 1994:84). Generational change, on the other hand, refers to situations in which speakers maintain their original grammar over time, "but regular increases in the values adopted by individuals […] lead to linguistic change for the community" (Labov 1994:84). In other words, if LIKE use is significantly socially stratified, this would corroborate the existence of ongoing change within the communal grammar, i.e. generational change.

7.1.1 Statistical analysis of LIKE usage in American English

The following statistical analysis evaluates the intuitions derived from the summaries and graphical displays. This statistical analysis is more powerful than a visual inspection; it aims to uncover trends which have previously escaped detection.

As discussed in section 5.9, a multivariate analysis is an appropriate design for this purpose – in this case, a (Quasi-) Poisson regression.[44] Like other types of regression models, Poisson regressions test whether an independent variable (age, gender, etc.) correlates significantly with the dependent variable (the frequency of LIKE). An advantage over simpler, more robust tests, such as non-parametric tests or $\chi 2$-tests, is that regressions also test for significant interactions between independent variables. In addition to reporting the level of significance, regression models provide the direction and strength of correlations or an estimate. The strength of correlations is commonly referred to as effect size or coefficient. An initial saturated model is fitted in a stepwise procedure to arrive at a minimal adequate model that contains only predictors which are either part of significant interactions or which have significant coefficients. Insignificant factors are eliminated to reduce "noise", enabling precise estimations of coefficients. This procedure will be applied to all variants of LIKE.

Before analyzing each variant of LIKE in isolation, we will focus on the overall use of LIKE. This initial inspection will provide more general understanding of the relation between LIKE use and social categories.

[44] Since the model requirements for both common linear models (e.g. ordinary least square regressions) and robust regression models have not been met, the statistical analysis uses a Poisson regression model, which is implemented as a subtype of a generalized linear model.

Table 28: Results of the multivariate regression[45] for LIKE use in AmE

ALL	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	1.65	0.279	5.90	<.001***
A2	-0.08	0.340	-0.24	.80
A3	-0.48	0.382	-1.25	.20
A4	-1.68	0.412	-4.08	<.001***

The final minimal adequate Poisson regression model predicts the frequency of LIKE per 1,000 words in AmE significantly better than a base-line model (χ^2 =21.737, df = 1, p<.001***). The output of the final minimal model is nevertheless unexpected: gender, ethnicity, occupation, and region are not significant. The only significant predictor of LIKE use is the age of speakers. Given the popularity of the assertion that female speakers, in particular, use LIKE (cf. Croucher 2004a:42; Siegel 2002:37and the sources therein), this result is quite unexpected: a notable female bias was predicted. In addition, the output does not show significant interactions between age and sex, indicating the absence of hidden gender differences. The increasing values of the age coefficients indicate an implicational hierarchy of the form A1>A2>A3>A4. In other words, the frequency of LIKE decreases steadily as age increases.

The absence of gender differences is validated by performing additional non-parametric t-tests. Non-parametric t-tests are more robust than regressions, but can only deal with one independent variable at a time.

Table 29: LIKE with respect to AGE and GENDER[46]

	A1	A2	A3	A4
LIKE	1.872*	0.530 (n.s.)	-0.149 (n.s.)	-0.566 (n.s.)

45 During the following statistical analyses, Poisson regression models serve as default statistics, while non-parametric t-tests are employed whenever model assumptions are not met, i.e. when the number of observations is too small to allow for a meaningful interpretation of the regression model.

46 The numeric values represent the t-statistics; the asterisks represent the level of significance (>.1=·; >.05=*; >.01=**; >.001=***; n.s.=not significant).

The t-tests detect a significant gender difference among young speakers of AmE. The low t-value suggests, however, that the effect of gender is quite moderate.

The absence of an effect of region corroborates D'Arcy's (2007:391) claim that "the perception of *like* as either an American or more specifically a Californian feature does not persevere" (D'Arcy 2007:391). The underlying assumption is that if California is the epicenter of the spread of LIKE throughout the US, then we should expect to appear more frequent there than elsewhere in the U.S.

In summary, the most notable result of the analysis suggests that most of the factors which were expected to be significant fail to affect LIKE use in any substantial way. The only variable which significantly correlates with LIKE use is the age of speakers, as older speakers use it less than younger ones. Furthermore, additional t-tests detected a significant albeit moderate effect of gender among young speakers of AmE.

7.1.1.1 Clause initial LIKE

This part investigates the use of clause-initial LIKE in AmE, as in (73).

(73) a. *Like* she was in the hospital in and out of the hospital all the time and stuff ever since she was little she was in and out of the hospital. (Santa Barbara Corpus:sbc001$Lynne)

b. So *like* what are they catching salmon and stuff? (Santa Barbara Corpus:sbc003$Pete)

c. *Like* they're they're trying to breed [...] a forty foot long tube chicken? (Santa Barbara Corpus:sbc003$Roy)

The number of young male speakers is too low to be representative; they will therefore not be included in the graphical displays.

Table 30: Age and gender distribution of clause-initial LIKE in US AmE[47]

AGE	MALE		FEMALE			
	LIKE (N)	LIKE (mean)	LIKE (N)	LIKE (mean)	SUM (N)	ALL (mean)
1 (11 – 20)	3	0.80	39	1.38	43	1.26
2 (21 – 30)	15	0.55	66	1.56	81	1.10
3 (31-40)	27	0.63	16	0.91	43	0.74
4 (41+)	24	0.32	13	0.18	37	0.25
SUM	70	0.48	134	0.88	204	0.70

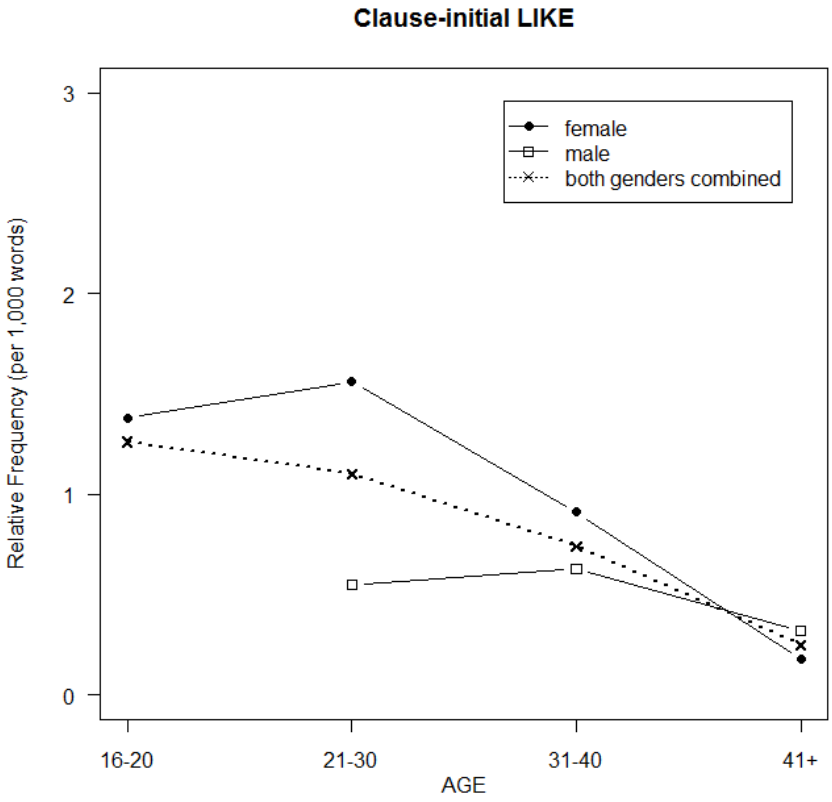

Figure 23: Age and gender distribution of clause-initial LIKE in AmE

If data from males and females are collapsed, the distribution shows a consistent recess of LIKE use with increasing age. This age stratification indicates age-graded use of clause-initial LIKE. If the genders are analyzed separately, this seemingly clear-cut picture changes. While the use of clause-

47 Speakers for whom age or gender information is not available are not included.

initial LIKE seems not to correlate significantly with age among males, this is not the case among females: older women clearly exhibit lower rates of clause-initial LIKE than younger ones; only the frequency of females below the age of 20 does not fully match this pattern. In contrast to the expected adolescent peak (Labov 2001:106), it is females in their twenties who make most use of this vernacular feature. This is notable, because a peak around the time of stabilization among females is one of the most striking reoccurring patterns in apparent-time studies of ongoing change (Labov 2001:454-456). Indeed, the existence of such a peak is considered a general requirement of female-dominated changes in progress (Labov 2001:455).

In fact, the effect of gender seems to be restricted to speakers who are in their twenties and thirties, while LIKE use among speakers older than 41 years of age does not seem to be gendered at all. With respect to claims that LIKE use is "rapidly increasing […] particularly in Western English-speaking countries" (Tagliamonte 2005:1898), the apparent-time distribution of clause-initial LIKE fails to disambiguate the situation: while the lack of age-grading among males indexes advanced stabilization, i.e. no change in progress, both the gender difference itself and its increase as speakers become younger suggest the opposite.

The following analysis, therefore, focuses on the question of whether the use of clause-initial is socially stratified, since this would support the notion that it is currently undergoing change. In addition, it will include the approximate measure of priming, accommodation, and idiosyncratic bias (the PAI index) to protect from over-estimating the effect of extra-linguistic factors such as gender and age.

Table 31: Results of the multivariate regression for clause-initial LIKE in AmE

INI	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.93	0.181	-5.14	<.001***
PAI	0.02	0.005	4.54	<.001***

The output reported by the minimal adequate model predicting the frequency of clause-initial LIKE is unexpected: the only significant predictor is

not a sociolinguistic one, but the PAI index. To interpret this finding adequately, it is essential to keep in mind that only significant coefficients are reported. This is crucial, as the age of speakers returned significant results in the final minimal adequate model (F-statistic=9.7615, df=1, 145, Pr(>|F|)<.01**). However the effect size of age was too minute to breach the level of significance. As in the case of overall LIKE use, it is the lack of significant effects that is remarkable: none of the classic extra-linguistic factors retuned significant. In contrast, it is the cumulative effect of priming, accommodation to the interlocutor and idiosyncrasies in LIKE use which best predicts the occurrence of clause-initial LIKE. This result draws attention to the fact that impressionistic analysis of data can lead to over-interpretation of variables. Based solely on Table 30 and Figure 23, we would have expected a stronger impact of age, while the more conservative – indeed perhaps too conservative – statistical analysis reports that the impact of age is negligible when priming, accommodation to the interlocutor, and individual bias are taken into account.

As Figure 23 indicates that age is significant among females, we will use χ^2-tests to check whether age is significant when males are removed from the data.

Table 32: Results of the χ2-tests for female use of clause-initial LIKE in AmE using the youngest females as reference

INI	χ2-value	df	p-value	Cramér's φ
A2 (21-30)	0.481	1	.48	0.00
A3 (31-40)	10.220	1	<.01**	0.01
A4 (41+)	40.300	1	<.001***	0.02

Indeed, if only female speakers are considered, age is significant, although its effect size (Cramér's φ) is extremely low. Moreover, when applying χ^2-statistics to test whether males differ significantly from females in their use of this feature, the more robust χ^2-test reports a significant difference, although the effect of gender is miniscule (χ^2 = 8.420, df=1, p<.01**, Cramér's φ=0.00). To validate this minute difference, additional non-parametric t-tests were performed.

Table 33: Clause-initial LIKE with respect to AGE and GENDER[48]

	A1	A2	A3	A4
INI	0.816 (n.s.)	1.404.	0.226 (n.s.)	-0.631 (n.s.)

The t-tests confirm a marginally significant gender difference among speakers in their twenties. While the use of clause-initial LIKE is not gendered in other groups, female speakers in their twenties use this form more than their male peers.

To summarize, sociolinguistic factors such as gender and age have only moderate explanatory power with respect to the use of clause-initial LIKE. Although females use it more than males, the difference – as with the impact of age – is marginal. The only significant predictor reported by the regression model is the PAI index, with a rather moderate effect size.

7.1.1.2 Clause-medial LIKE

The following section focuses on the use of t clause-medial LIKE as in (74).

(74) a. How come *like* the bread crumbs weren't already attached to the fish or the pan? (Santa Barbara Corpus:sbc003$Pete)

b. Should I make *like* tartar sauce or something really decadent? (Santa Barbara Corpus:sbc003$Marilyn)

c. You know if her skin weren't *like* really dark brown she'd probably be bright red. (Santa Barbara Corpus:sbc003$Sharon)

It is this variant of LIKE that has received most attention and which is commonly considered the prototypical form of vernacular LIKE typically associated with AmE (cf. D'Arcy 2007). Andersen (2000:221), for example, argues that this variant of LIKE in the speech of London teenagers represents an American borrowing. Hence, it is particularly interesting to explore its usage patterns within its supposed original habitat: nonetheless, the hypothesis that

[48] The numeric values represent the t-statistics; the asterisks represent the level of significance (>.1=·; >.05=*; >.01=**; >.001=***; n.s.=not significant).

this form originated in AmE is open to dispute[49] given that it is attested not only in the speech of 65- to 75-year-old NIE speakers in data collected during the 1970s (cf. Schweinberger 2013a; D'Arcy 2005, 2007 for equivalent attestations).

Table 34: Age and gender distribution of clause-medial LIKE in AmE[50]

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (11 – 20)	12	0.60	77	2.96	79	2.40
2 (21 – 30)	54	2.11	100	2.15	154	2.13
3 (31-40)	65	1.63	30	1.48	95	1.57
4 (41+)	24	0.51	16	0.29	40	0.39
SUM	145	1.20	223	1.46	368	1.34

Figure 24 represents further evidence corroborating assertions that LIKE is most common among younger speakers. Indeed, clause-medial LIKE appears to be highly sensitive to age, as it clearly shows a heavily age-stratified distribution. Gender, on the other hand, seems not to affect its use, which is surprising given the widespread belief that vernacular LIKE is typically a feature of female speech.

[49] The hypothesis according to which this particle "developed among New York City counterculture groups (i.e., jazz, cool, and Beat) during the 1950s and 1960s" (Andersen 2001: 216; D'Arcy 2007: 398) fails to explain its occurrence in the speech of sixty- to eighty-year-olds in Toronto English (D'Arcy 2005) and Northern Irish English (Schweinberger forthc.a).

[50] Speakers for whom age or gender information is not available are not included.

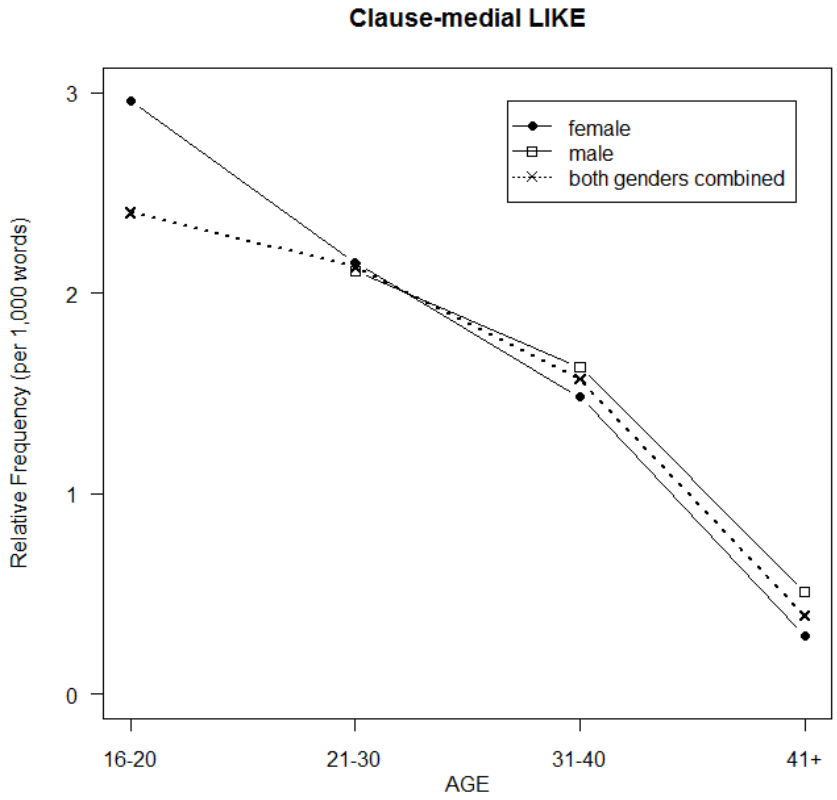

Figure 24: Age and gender distribution of clause-medial LIKE in AmE

We will now turn to the statistical analysis to confirm the trends emerging from the graphical displays.

Table 35: Results of the multivariate regression[51] for clause-medial LIKE use in AmE

MED	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	0.32	0.342	0.95	.33
A2	0.15	0.379	0.40	.68
A3	-0.03	0.437	-0.08	.93
A4	-1.45	0.416	-3.48	<.001***
PAI	0.01	0.005	2.90	<.01**

51 During the following statistical analyses, Poisson regression models served as default statistics while non-parametric t-tests were employed whenever model assumptions are not met, i.e. if the number of observations was too low to allow for a meaningful interpretation of the regression model.

The final minimal adequate model significantly out-performs the baseline model (χ^2 = 27.013, df=4, p<.001***). With respect to clause-medial LIKE, both the effect of age of speakers and the effect of approximate priming measure are significant. Although the negative correlation between the age of speakers and the use of this type of LIKE is highly significant, the effect is rather weak, though still stronger than the effect of the approximate priming measure.

Despite popular lore (cf. D'Arcy 2007), clause-medial LIKE is not overly abundant in California compared to other regions within the US. In addition, the results indicate that it is not more common among white middle-class speakers, that it is not a feature of the lower working class and it is not limited to the speech of adolescent females (Dailey-O'Cain 2000:68-69; Siegel 2002:37).

In contrast to what could be predicted based on the age stratification reported by Figure 24, the age coefficients do not reflect a steady increase from young to old. Indeed, the positive coefficient of age group 2 suggests a peak among speakers in their twenties. However, Figure 24 suggests that the use of clause-medial LIKE decreases with increasing age, i.e.: A1>A2>A3>A4. Again, we use non-parametric, one-tailed t-tests (cf. Table 36) to determine which of these assertions (peak among speakers in their twenties vs. monotone recess) is more likely to be true.

Table 36: Clause-medial LIKE with respect to AGE

	A1-A2	**A1-A3**	**A1-A4**
MED	0.322 (n.s.)	0.987 (n.s.)	2.835**

Although the results only confirm a significant difference between age groups 1 and 4, the t-tests lead to an improved understanding, as they validate the implicational hierarchy from young to old. The steady increase in the t-values reflects the trajectory expected based on Figure 24.

In summary, the only claims which can be asserted with confidence based on the present analysis are that younger speakers use LIKE more often than

older speakers and that if LIKE is used once in a dialogue, it will probably be used again.

7.1.1.3 Clause-final LIKE

There is only one instance of clause-final LIKE in the American data which is uttered by a female in a task-related interaction. This interaction consists of "an attorney preparing two witnesses to testify in a criminal trial [which has been] [r]ecorded in San Francisco, California. Rebecca is a lawyer, June and Rickie are the witnesses, and Arnold is Rickie's husband" (http: //www.linguistics.ucsb.edu/research/sbcorpus_citation.html; accessed March 5th, 2011).

(75) And he's like balding *like*. (Santa Barbara Corpus:008$RICKI)

Taking the context into account, the use of clause-final LIKE seems to focus the listener's attention to the fact that the person talked about is becoming bald. The person talked about is the suspect, and the immediately preceding context clarifies that Rickie identified the suspect at a police station. Indeed, the suspect's most prominent characteristics are that he is in his sixties and becoming bald. The fact that this instance of clause-final LIKE focuses or emphasizes a new piece of information is consistent with its use in other varieties of English. For example, Columbus (2009) asserts that clause-final LIKE in IrE commonly serves to focus the preceding element, phrase, or clause.

7.1.1.4 Non-clausal LIKE

A preliminary look at the age and gender distribution of non-clausal LIKE as in (76) shows that this form is age stratified and slightly used more often by males than by females (cf. Table 37).

(76) a. And it was *like*… (Santa Barbara Corpus:sbc003$Pamela)

b. And I'm *like* I'm trying not to vomit listening to this. (Santa Barbara Corpus:sbc003$Lenor)

c. I think maybe *like* you know how these chairs are. (Santa Barbara Corpus:sbc003$Ricki)

Table 37: Age and gender distribution of non-clausal LIKE in AmE[52]

AGE	MALE		FEMALE			
	LIKE (N)	LIKE (mean)	LIKE (N)	LIKE (mean)	SUM (N)	ALL (mean)
1 (11 – 20)	4	0.80	56	1.62	60	1.45
2 (21 – 30)	22	1.56	60	1.33	82	1.43
3 (31-40)	26	0.86	11	0.78	37	0.83
4 (41+)	22	0.31	12	0.25	34	0.28
SUM	74	0.80	139	0.88	213	0.85

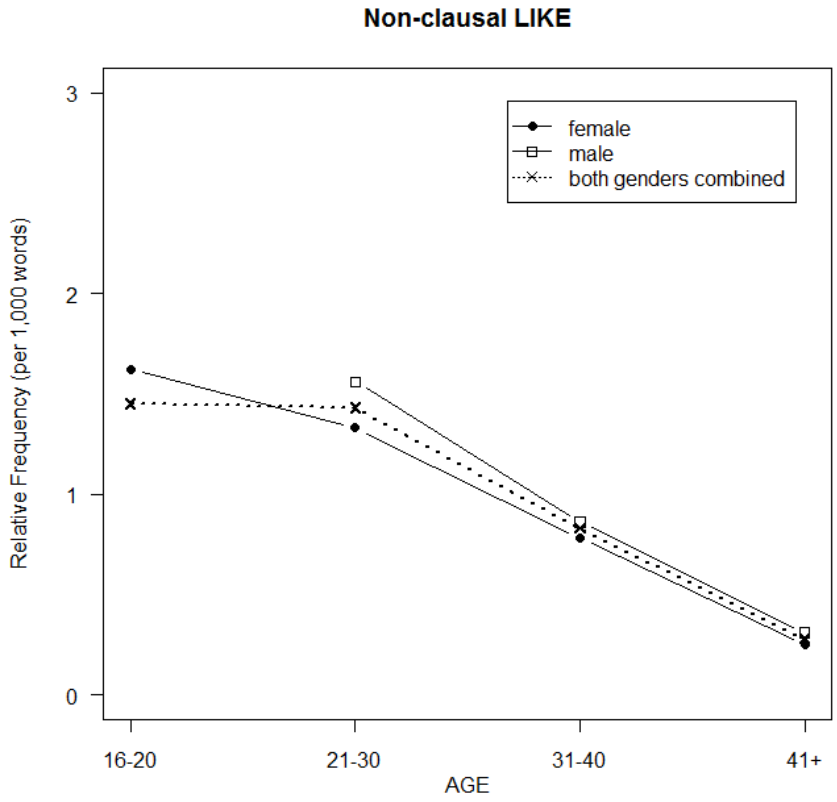

Figure 25: Age and gender distribution of non-clausal LIKE in AmE

Figure 25 corroborates a monotone declining age stratification, while the proximity of the lines representing the male and female mean frequencies

52 Speakers whose age or gender information was not available are not considered.

indicates the lack of significant gender differences. These impressions were tested using a multivariate design.

Table 38: Results of the multivariate regression [53] for non-clausal LIKE use in AmE

NON	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	0.85	0.456	1.87	<.1
A2	-1.23	0.431	-2.86	<.01**
A3	-1.76	0.573	-3.08	<.01**
A4	-2.66	0.499	-5.32	<.001***
PAI	0.02	0.008	2.39	<.05*
ADC	0.79	0.409	1.93	<.1

In contrast to all other variants of LIKE in AmE, non-clausal LIKE shows signs of social stratification, although not in the expected direction: academics and those working in the professions use non-clausal LIKE more than speakers in clerical and managerial occupations, as well as (un-)skilled manual labor. Furthermore, its use is extremely age sensitive: the younger the speaker is, the higher the rate of his or her use of this form. The pattern matches the implicational hierarchy observable for all common variants of LIKE in AmE (A1>A2>A3>A4).

The last significant predictor variable is the frequency of LIKE in a dialogue. This result matches observations about the accommodation of discourse styles and lexical priming effects; it also confirms that non-clausal LIKE heavily clusters. It follows that beyond its dependency on psycho-linguistic factors (priming), social interaction, and co-operation (accommodation), LIKE use is also likely to be affected by the individual speaker's idiosyncratic linguistic preferences.

[53] In the following statistical analyses, Poisson regression models served as default statistics while non-parametric t-tests were employed whenever model assumptions were not met, i.e. when the number of observations was too low to allow for a meaningful interpretation of the regression model.

7.1.2 Summary and discussion of LIKE use in American English

The results of the present analysis confirm that vernacular uses of LIKE are a salient feature of contemporary AmE, as it was used by two thirds of AmE speakers with an average of 2.9 instances per 1,000 words. This confutes the hypotheses stating that LIKE was already becoming archaic in AmE as early as 1988 (Underhill 1988). The high proportion of speakers who use LIKE suggests that it has stabilized and that it is common across all social strata. Based on the lack of social stratification, it is likely that the process of diffusion is nearing completion (Nevalainen & Raumolin-Brunberg 2003:55).

Despite the results of the present analysis superficially appearing to confirm the popular belief that this form is typically a feature of "teenage speak", this conclusion is too simplistic considering LIKE's complex sociolinguistic profile. Although LIKE use is notably age-stratified and thus age-graded, it is far from being exclusively used by adolescents. The analysis of clause-initial LIKE, which appears to be gender sensitive, shows that the effect of gender, though significant, is too weak to be considered substantial. In addition, the prototypical form of LIKE, i.e clause-medial LIKE, as well as non-clausal uses of LIKE show no sign of a female preference.

The monotone recess of LIKE use with increasing age is the most consistent pattern in the present study: all forms are age sensitive, though clause-initial LIKE showed age-grading only among females and not males. While the analysis validates the notion that clause-medial LIKE is more common among younger speakers (Dailey-O'Cain 2000:66), it fails to corroborate a substantial female preference. Hence, the results challenge the validity of popular lore which holds that LIKE is particularly common in the speech of women and specifically female adolescents (cf. D'Arcy 2007:391; Dailey-O'Cain 2000). In addition, the analysis refutes the widespread belief according to which LIKE is most prevalent in "Valspeak", i.e. the sociolect associated with female adolescents from California (Blyth, Recktenwald & Wang 1990:224; Dailey-O'Cain 2000:70; Siegel 2002:37). Indeed, LIKE use is not particularly frequent

in Californian English, as REGION does not significant correlate with its frequency. Indeed most factors which were expected to be informative show no significant effect.

Considering that Dailey-O'Cain (2000:74) found that LIKE is perceived to be associated with lower intelligence, we should expect it to be more common among speakers with lower socio-economic status. This, however, is not the case. On the contrary: the only significant effect of occupation surfaces with respect to the use of non-clausal LIKE which is preferred by speakers pursuing academic careers, or who are in the professions – i.e. the opposite of what we would conclude based on Dailey-O'Cain's (2000) perceptional study. The results, therefore, strongly suggest that LIKE use is less socially stratified than commonly assumed (Dailey-O'Cain 2000). This finding is particularly interesting, as the absence of social stratification does not match the assumption that social patterning of linguistic behavior is parasitic on stylistic linguistic variation (cf. Labov 2002).

As the results of the multifactorial analysis confirm that the overall frequency of LIKE in conversations is a significant factor, the introduction of the PAI index appears appropriate to counter over-estimating the effect of extra-linguistic factors such as age or gender.

The difference between the common lore and the reality of quantitative analyses (this study and Dailey-O'Cain 2000) require further discussion. The question arises as to why LIKE is associated with the Valley Girl persona (D'Arcy 2007:397) when the evidence does not support this assertion. D'Arcy (2007:397) makes a compelling case in approaching this question:

> It is possible, for example, that vernacular uses of *like* were recycled as a Valley Girl phenomenon once their initial association with the counterculture groups waned among subsequent generations of speakers. As Milroy (2004: 169) states, "different groups may be foregrounded at different times." In other words, 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.

According to D'Arcy (2007), the association of LIKE with "Valspeak" probably arose once the association between LIKE and New York counterculture groups (cf. Andersen 2001) faded. In this view, "Valley Girls" have not been responsible for initiating the spread of LIKE throughout the US, but served as the social reference category once LIKE was already well in use. This argumentation is in line with Eckert's (2001) analysis of stylistic variation and identity marking. Following Eckert (2001), certain linguistic practices are akin to and correlated with other social practices, such as choice of dress, adornment, and demeanor. In this perspective, speakers intentionally use linguistic means to construct and define mutually distinctive social categories (Milroy & Gordon 2003:208).

An alternative interpretation relates to the results of attitudinal studies of LIKE use (Dailey-O'Cain 2000). Dailey-O'Cain (2000) found that LIKE is not only associated with certain social groups, e.g. teenagers, but it is also perceived to correlate with certain personality traits. According to Dailey-O'Cain, speakers using LIKE are perceived to be "more attractive, more cheerful, more friendly, and more successful […] [while its] absence causes the listener to think of the speaker as less polite and less friendly" (2000:73). Hence, LIKE may not only serve as a marker of social identity, but also as a device deployed to manipulate the perception of oneself by interlocutors. Akin to the interpretation above, this take on LIKE use expands on Eckert's (2001) analysis of stylistic variation, as it too relates to identity marking. Indeed, it expands the social psychological concepts underlying this argument by complementing it with theoretical aspects of the psychology of personality: the difference consists in focusing on the self image rather than marking group membership.

In conclusion, it appears that LIKE has successfully diffused through all social strata. The lack of significant extra-linguistic social factors supports this

interpretation: the effect of gender and social class wanes (Labov 1994:79-82) during the final stages of change. Thus, the view taken here assumes that any processes of ongoing change are nearing completion while the popular beliefs and prejudices associated with this vernacular feature are more conservative. The conservatism of social stereotypes is at odds with the present data, as it reflects earlier stages of ongoing change, while in reality the linguistic bebavior has outgrown these stereotypes.

7.2 LIKE in Canadian English[54]

CanE is particularly interesting with respect to the study of LIKE, as it is the most extensively studied variety of English in terms of systematic, quantitative analyses (cf. D'Arcy 2005, 2006, 2007; Kastronic 2011; Tagliamonte 2005; Tagliamonte & D'Arcy 2004). Previous studies have provided a detailed picture of the syntactic and sociolinguistic factors affecting which type of LIKE is used by which subsection of this speech community. The following section will, therefore, aim at evaluating previous findings and results in light of the present data.

Before analyzing the data in detail, we will have a look at the basic statistics to obtain a first impression of its structure (cf. Table 17). Of all 252 CanE speakers, 186 have used LIKE at least once, i.e. nearly three quarters. This high proportion shows that LIKE is abundant in CanE.

Table 39: Overview - LIKE use in CanE

	N (total)	Mean (LIKE per 1,000 words)	Median (LIKE per 1,000 words)	Standard deviation (s)
LIKE	900	4.377	3.180	4.797

Table 17 suggests that LIKE is a highly salient feature of CanE vernacular. And although the difference between mean and median is substantial, both measures of a central tendency are remarkably high. Indeed, the median shows that 50 percent of speakers use LIKE more often than 3.180 per 1,000 words. These basic statistics provide valuable insights into the data distribution: as in AmE, the high proportion of LIKE users suggests that if LIKE has indeed been undergoing change, it is probably nearing completion (Nevalainen & Raumolin-Brunberg 2003:55).

54 For a complete overview of the final data set of CanE used in this study, cf. Table 139 in the Appendix.

Before turning to a more fine-grained analysis of age and gender-specific LIKE use, the frequencies of types of LIKE are displayed in order to clarify which types of LIKE are especially frequent in CanE.

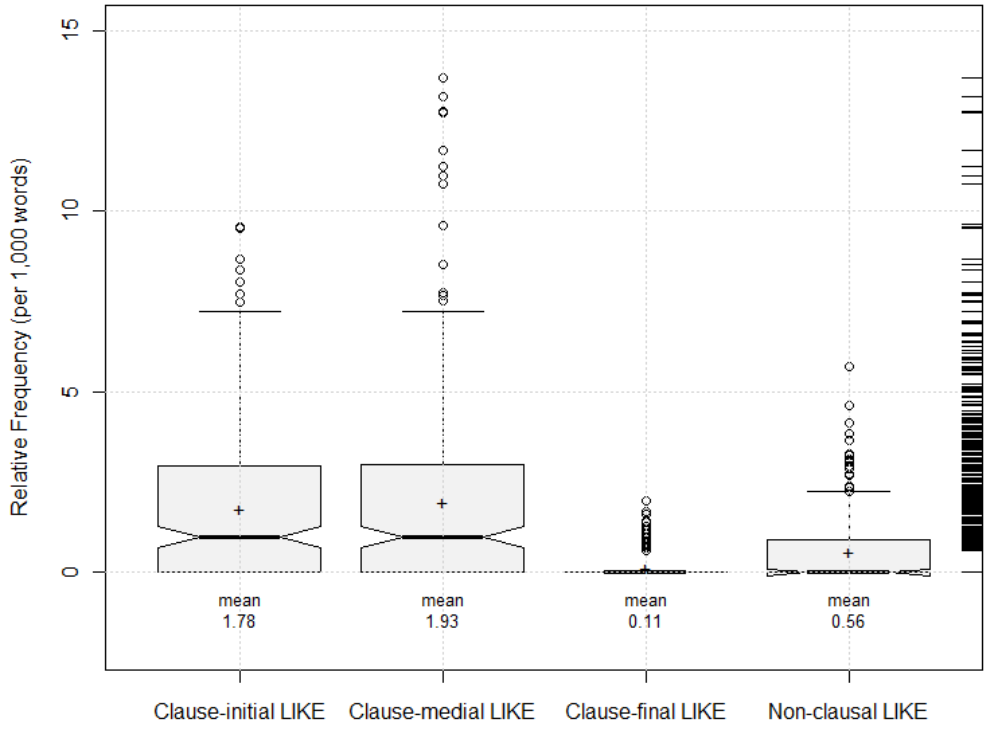

Figure 26: Rate of LIKE variants in CanE

Figure 26 shows that clause-initial LIKE and clause-medial LIKE are the most frequently used variants. In contrast to AmE, clause-final LIKE is also used in this variety although rather infrequently. While being used substantially more than the clause-final variant, non-clausal LIKE is significantly less frequent than clause-initial and clause-medial forms.

In her analysis, D'Arcy (2005:221) found an apparent-time increase in the use of both clause-initial and clause-medial LIKE over time. An increase in LIKE use over time is substantiated by the increase in frequency displayed in Table 40.

Table 40: Age and gender distribution of LIKE in CanE

			MALE		FEMALE			
AGE	**Speakers**	**Words**	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (10 – 24)	28	23,293	115	7.96	106	10.08	221	8.72
2 (25 – 30)	56	41,038	65	4.26	182	7.50	247	6.46
3 (31 – 40)	78	65,065	82	3.08	235	4.62	317	4.03
4 (41+)	82	65,178	22	1.24	93	2.03	115	1.81
SUM	244	194,574	284	3.83	616	4.69	900	4.38

Furthermore, the data reported in Table 40 show a consistent female lead in LIKE use across all age groups. The extent of the female lead is surprising given that D'Arcy found that "[a]s a marker, LIKE is significantly correlated with females. […] By contrast, clause-medial LIKE is significantly correlated with males" (2005:222). Applied to the present analysis, this would lead to the prediction that overall LIKE use lacks a notable gender difference because the gender differences in the use of clause-initial and clause-medial LIKE should cancel each other out. According to D'Arcy (2005), we would furthermore predict that clause-medial LIKE is preferred by male speakers, while we should expect females to prefer clause-initial and clause-final LIKE.

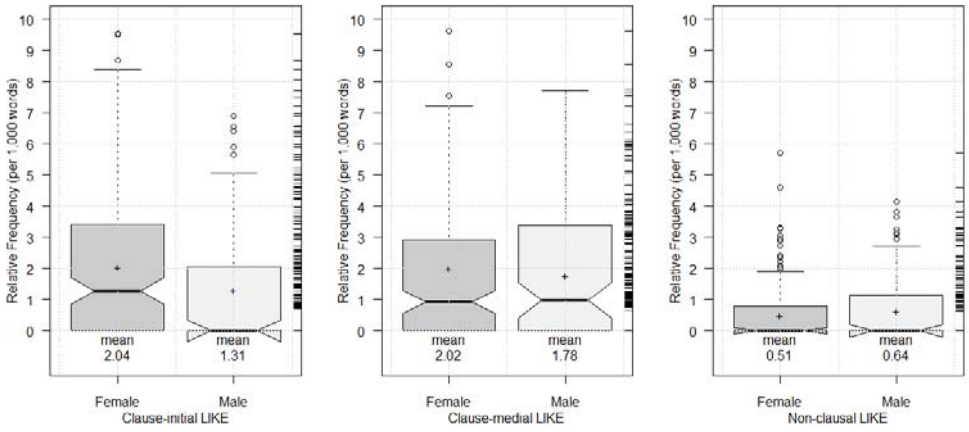

Figure 27: Box plots showing the distribution of clause-initial (left), clause-medial (middle), and non-clausal LIKE (right) in CanE with respect to gender

Figure 27 corroborates D'Arcy's claim concerning a female lead in the use of the discourse marker and reports the expected lack of a gender difference

regarding non-clausal LIKE. The data fails, however, to substantiate D'Arcy's (2005:222-223) finding that males exhibit a preference for clause-medial LIKE. According to the ICE data, females use both clause-initial and clause-medial LIKE more frequently than male speakers do.

This pattern remains stable when the frequencies of clause-initial and clause-medial LIKE are displayed with respect to the age and the gender of the speakers. The additional information can be derived from Figure 28 and Figure 29 is that that this tendency is particularly notable among young female speakers aged 10 to 24.

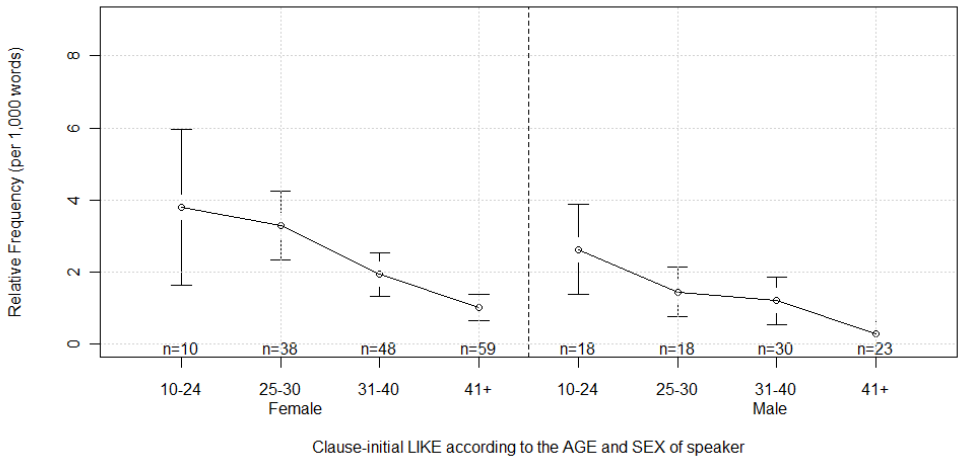

Figure 28: Age and gender distribution of clause-initial LIKE in CanE[55]

55 The confidence intervals of the group containing female speakers aged 10 to 24 are extremely wide, as this group consists of only 13 speakers who use LIKE quite distinctly: of these 13 speakers, 4 use LIKE with a frequency exceeding 20 times per 1,000 words, while 5 speakers use it less than 10 times per 1,000 words. The combination of both factors, the low number of speakers plus the extremely large variance of LIKE use within this cohort causes the wide range of the confidence intervals.

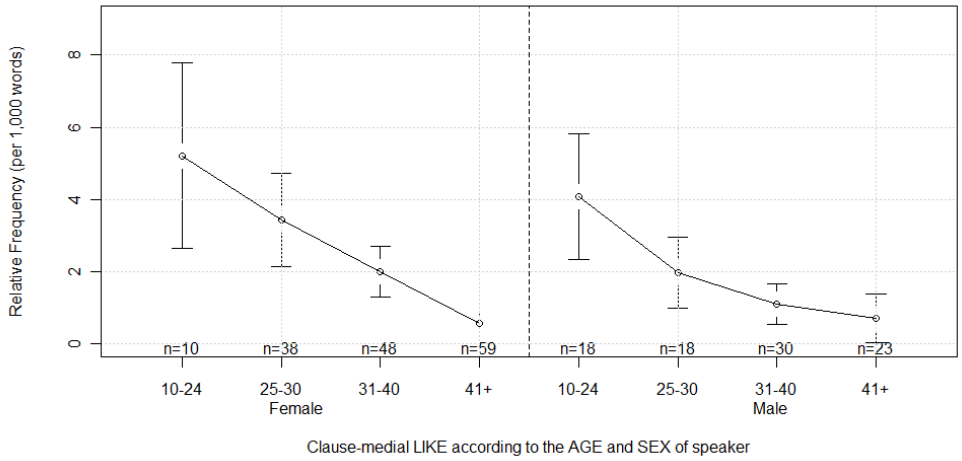

Figure 29: Age and gender distribution of clause-medial LIKE in CanE

In contrast to the AmE data, the CanE data seem to confirm the assumption that women use LIKE more frequently than men. Indeed, both Figure 28 and Figure 29 indicate a slight female preference for LIKE, irrespective of the clause-initial versus clause-medial distinction, i.e. the marker vs. particle distinction in D'Arcy's (2005, 2007) study. In addition, the distributions of clause-initial and clause-medial LIKE across genders and age groups are very similar: both figures exhibit a peak in the youngest age cohort and show a steady but consistent decrease in use with increasing age. This age stratification suggests not only age-grading, but it also indicates the same implicational hierarchy which was observable in AmE (A1>A2>A3>A4).

7.2.1 Statistical analysis of LIKE usage in Canadian English

The following statistical analysis evaluates the intuitions derived from the graphical displays and seeks to uncover correlations which previously escaped detection. As in the case of AmE, the appropriate design for this purpose is a multivariate regression, complemented by non-parametric t-tests to ascertain ambiguous findings.

The first regression model is applied to the combined values of all variants of LIKE. This rather coarse-gained analysis aims at providing a preliminary impression of correlations between LIKE use and extra-linguistic factors.

Table 41: Results of the multivariate regression [56] for LIKE use in CanE

ALL	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	2.15	0.198	10.81	<.001***
A2	-0.16	0.238	-0.67	.50
A3	-0.51	0.256	-2.01	<.05*
A4	-1.34	0.250	-5.35	<.001***
ADC	-0.39	0.162	-2.46	<.05*

The Poisson regression model performs significantly better than the base-line model (χ^2 =56.534, df=4, p<.001***) and reports a more complex interplay of variables than in AmE. With respect to age, AmE and CanE exhibit very similar distributions, as the implicational hierarchy observable in AmE also seems to apply to the CanE data. In addition, the age stratification is suggestive of age-grading. However, AmE and CanE differ notably with respect to social stratification. While the socio-economic status of speakers in AmE did not correlate with LIKE use, it does significantly in CanE. This shows that LIKE use in Canada is socially stratified, as speakers with a higher socio-economic status, i.e. speakers pursuing academic or professional careers, use it less often than speakers in clerical and managerial occupations (the reference category). This correlation between the socio-economic status of speakers and their use of LIKE suggests that LIKE has not (yet) diffused through all social strata. A plausible explanation for this finding is that LIKE use in CanE is more stigmatized among certain social layers than it is in AmE. This would imply that register sensitivity and thus usage constraints are more prominent in CanE. If this explanation were shown to be valid, then this would mean that LIKE use in CanE has attracted more social awareness in the sense that it suffers from more intense stigmatization. Although this interpretation initially seems plausible, it is at odds with the consistent female lead as women conform more closely to

[56] During the following statistical analyses, Poisson regression models served as default statistics while non-parametric t-tests were be employed whenever model assumptions were not met, i.e. when the number of observations was too low to allow for a meaningful interpretation of the regression model.

overtly prescribed norms and use fewer non-standard forms than men (Labov 2001:266). One way to resolve this issue relates to the onset or extent of stigmatization. While women adopt incoming or non-standard forms more readily than men, the situation is reversed when these forms become stigmatized (Labov 2001:293). It may be hypothesized that women lead the change and, therefore, exhibit higher frequencies of LIKE, but that either stigmatization began only recently, or that it is constrained to certain social strata, i.e. strata relatively high in the socio-economic scale.

7.2.1.1 Clause-initial LIKE

The following section focuses on the clause-initial discourse marker as in (77) and its correlation with extra-linguistic social variables.

(77) a. You know but uhm *like* I'd be willing to do if you wanted to do two sets. (ICE-Canada:S1A-010$A)

b. *Like* the little organizer that you put in your door drawer. (ICE-Canada:S1A-011$A)

c. *Like* do you watch those <,> uh home improvement things on uh <,> P B S uh. (ICE-Canada:S1A-012$A)

As was the case when analyzing the distinct types of LIKE in AmE, the following analysis will include the PAI index to countervail over-estimating the effects of extra-linguistic variables.

Table 42: Age and gender distribution of clause-initial LIKE in CanE

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (10 – 24)	38	2.63	41	3.80	79	3.05
2 (25 – 30)	25	1.45	86	3.28	111	2.69
3 (31-40)	33	1.21	94	1.93	127	1.65
4 (41+)	7	0.29	44	1.02	51	0.82
SUM	103	1.13	265	2.74	368	1.77

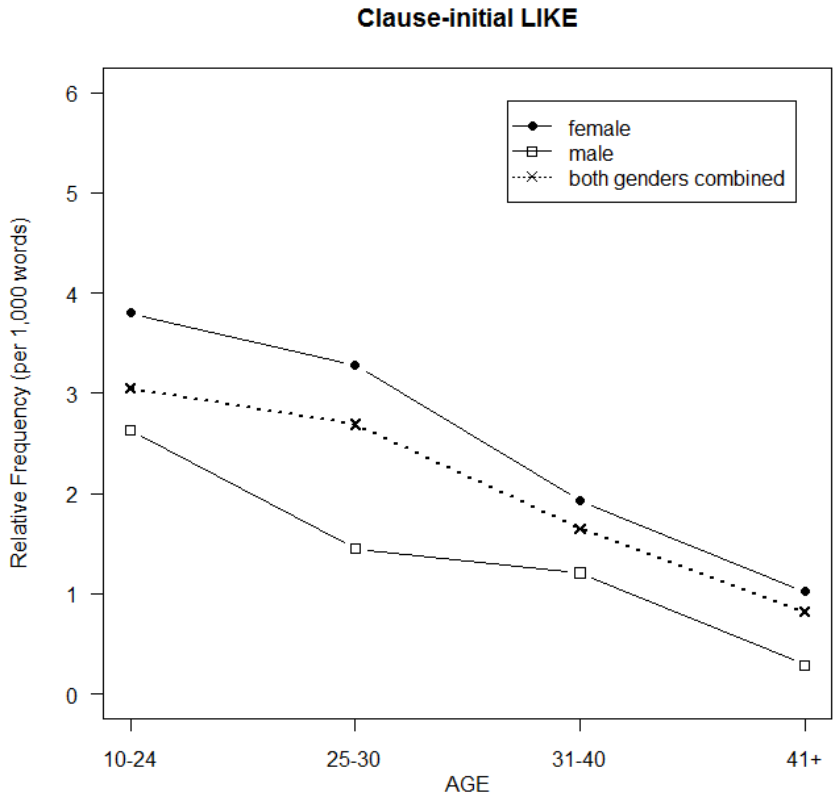

Figure 30: Age and gender distribution of clause-initial LIKE in CanE

The age stratification in Figure 30 suggests age-grading and gendered use of clause-initial LIKE. Indeed, the gender difference is consistent across all age groups with women ahead of men. The most pronounced difference between female and male speakers is, however, not among teenagers, but among speakers in their twenties.

Table 43: Results of the multivariate regression for clause-initial LIKE in CanE

INI	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	1.16	0.547	2.12	<.05*
SEX: Male	-0.77	0.238	-3.24	<.01**
A2	-0.61	0.525	-1.17	.23
A3	-0.63	0.534	-1.19	.23
A4	-1.35	0.533	-2.54	<.05*
PAI	0.03	0.005	5.41	<.001***
ACD	-0.31	0.184	-1.69	<.1

Again, the final minimal model performs significantly better than a model based solely on the intercept (χ^2 = 86.189, df=6, p<.001***). The regression report validates that the use of clause-initial LIKE is significantly gendered as it is more frequently used by females. In order to obtain a more precise understanding of exactly where males and females differ, non-parametric t-tests are applied to each age-group in isolation.

Table 44: Clause-initial LIKE with respect to GENDER

	A1	A2	A3	A4
INI	0.891 (n.s.)	3.131** (<.01)	1.513 (<.1)	3.450*** (<.001)

The t-tests report that the gender difference is particularly pronounced among speakers in their twenties and forties, while there is only a marginal difference among speakers in their thirties: the gender difference among speakers below the age of 25 is not significant, suggesting that the association of LIKE with female speech seems to disappear as it continues to diffuse throughout the speech community as a whole.

The regression model also confirms a near linear negative correlation between the frequency of clause-initial LIKE and the age of speakers. In other words, the frequency of this form decreases as the age of speakers increases. However, this result is tentative, as speakers aged 25 to 40 do not differ significantly from speakers aged 10 to 24. Indeed, only speakers above 41 years of age use clause-initial LIKE significantly less often than speakers under 24 (the reference group). The steady increase in the age coefficients from young to old seems, nonetheless, to corroborate the assertion that LIKE use decreases with increasing age. In addition, the values of the coefficients conform to the expected order: A1>A2>A3>A4. The robustness of this implicature is evaluated using non-parametric one-tailed t-tests (cf. Table 45).

Table 45: Clause-initial LIKE with respect to AGE

	A1-A2	A2-A3	A3-A4
INI	0.429 (n.s.)	2.431**(<.01)	2.937**(<.01)

The t-tests confirm the regression results and validate the implicational hierarchy from young to old. The results also confirm significant differences between age groups 2 and 3 as well as between age groups 3 and 4.

The occupation of speakers also affects the use of clause-initial LIKE: speakers in clerical and managerial occupations, i.e. the reference group, use this variant significantly more often than speakers pursuing academic and professional careers. This finding is not particularly surprising in itself, given that the overall model also reported this tendency. It is notable, however, that LIKE use is apparently socially stratified in the way we would expect if it was indeed stigmatized in higher social strata.

The present regression model does not report a significant difference between monolingual native speakers of English and speakers who have been brought up bilingually. As is the case in AmE, the PAI index returns significant and weakly correlates with the use of clause-initial LIKE.

7.2.1.2 Clause-medial LIKE

D'Arcy asserts that clause-medial LIKE, as in (78), "is significantly correlated with males" (D'Arcy 2005:222). This atypical gender bias has left D'Arcy rather perplexed; she concludes that "LIKE does not fit into typical categories of linguistic variability, whether the variation signals change or not" (D'Arcy 2005:224).

(78) a. And they they they had more of a feel of *like* apartment buildings even though they were only duplexes. (ICE-Canada:S1A-007$B)
b. [A]nd that's why he has *like* these exercise books (ICE-Canada:S1A-010$B)
c. That guy *like* provides the cash. (ICE-Canada:S1A-020$B)

The present data seems not to follow the pattern described by D'Arcy (2005) as both Figure 27 and Figure 29 indicate a substantial female lead for clause-initial as well as clause-medial LIKE. The following analysis will attempt to clarify this issue.

Table 46: Age and gender distribution of clause-medial LIKE in CanE

AGE	MALE		FEMALE			
	LIKE (N)	LIKE (mean)	LIKE (N)	LIKE (mean)	SUM (N)	ALL (mean)
1 (10 – 24)	57	4.07	52	5.21	109	4.48
2 (25 – 30)	28	1.97	71	3.44	99	2.97
3 (31-40)	29	1.09	104	2.00	133	1.65
4 (41+)	10	0.71	30	0.56	51	0.60
SUM	124	1.77	257	2.01	381	1.93

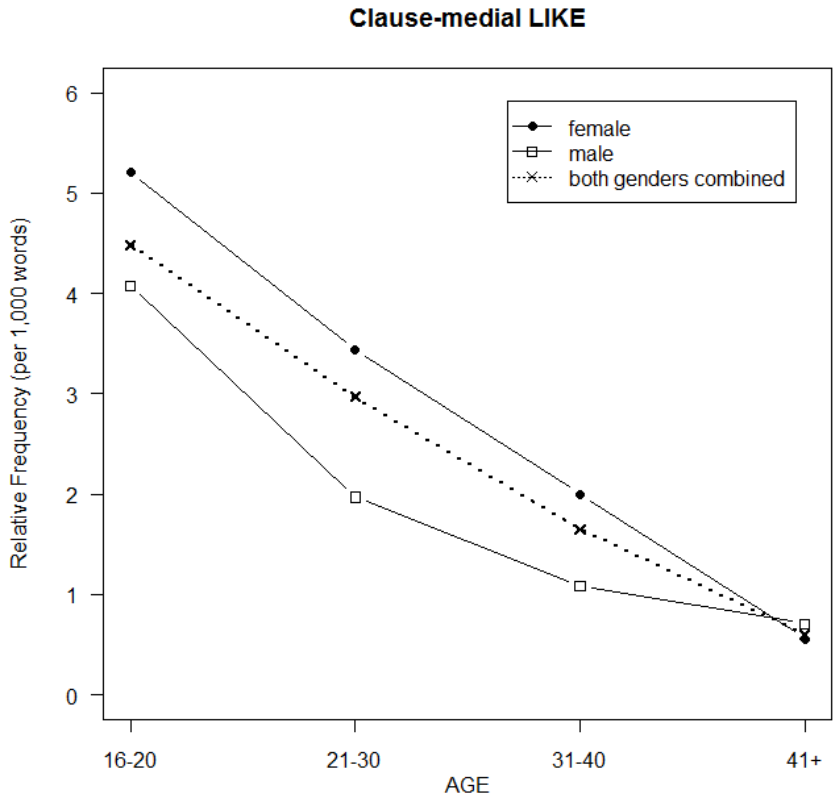

Figure 31: Age and gender distribution of clause-medial LIKE in CanE

Figure 31 corroborates a substantial decrease in the use of clause-medial LIKE from young to old. This type of age stratification suggests age-grading, i.e., that individuals stop using this variant as they grow older. In addition, Figure 31 indicates that until about age 40, females make notably more use of clause-medial LIKEthan males.

Table 47: Results of the multivariate regression for clause-medial LIKE in CanE

MED	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	0.52	0.256	2.04	<.05*
SEX: Male	-0.00	0.192	-0.01	.98
A2	-0.36	0.268	-1.36	.17
A3	-0.32	0.195	-1.67	<.1
A4	-1.50	0.275	-5.44	<.001***
PAI	0.04	0.006	7.28	<.001***
SEX: Male*A3	-0.74	0.377	-1.93	<.05*

The Poisson regression model predicting the frequency of clause-medial LIKE per 1,000 words in CanE is statistically significant (χ^2 = 127.16, df=6, p<.001***) and reports age and the PAI index as significant predictor variables. Gender does not have a significant effect, but it is part of a significant interaction with age. As expected, age is the strongest predictor in the present model. Speakers over age 41 use clause-medial LIKE significantly less often than speakers under 24, i.e. the reference group. It is notable that the multifactorial model indicates that the age stratification is not near-linear. Rather, the model seems to reflect a step-like progression with three levels: the youngest speakers in the data show the highest rate of non-clausal LIKE use and, accordingly, form the first level. The second highest level form age groups 2 and 3, as they use this form very similarly and with almost identical rates. Speakers above age 41, i.e. the oldest age group, have the lowest rate of clause-medial LIKE and thus form the third level: A1>(A2/A3)>A4. This result is, however, at odds with the visualization in Figure 31, which suggested a near-linear negative correlation between clause-medial LIKE and age. Accordingly, non-parametric one-tailed t-tests are applied to evaluate the robustness of this implicature indicated by the regression analysis (cf. Table 48).

Table 48: Clause-medial LIKE with respect to AGE

	A1-A2	A2-A3	A3-A4
MED	1,962*	2.400**	3.762***

The t-tests provide a more accurate picture of the relationship between age and the use of clause-medial LIKE. The results confirm the existence of an

implicational hierarchy of the form A1>A2>A3>A4 and thus validate the trend visible in Figure 31.

The model reports that males aged between 31 and 40 are particularly unlikely to use clause-medial LIKE. This could be attributed to the phase of change when LIKE was most strongly associated with females and could thus be reminiscent of a notable sex difference accompanying the most vigorous phase of its spread in CanE. Hence, t-tests are applied to each age group in isolation to validate this finding and to look for hidden gender differences.

Table 49: Clause-medial LIKE with respect to GENDER

	A1	A2	A3	A4
MED	0.797(n.s.)	1.771*	2.043*	-0.306(n.s.)

Revisiting the data using t-tests has paid off as they report previously undetected gender difference among speakers in their twenties. Furthermore, the t-tests validate the regression output as they confirm the gender difference among speakers in their thirties. Hence, the results of the present study fail to substantiate D'Arcy's (2005) assertion that clause-medial LIKEcorrelates with male speakers. When revisiting the data using χ^2-statistics to obtain yet another insight into the data, the results confirm the opposite, i.e. that females make use of this form significantly more often than males, though the size of this gender effect is quite small (cf. Table 50).

Table 50: Results of the χ2-test results evaluating the effect of gender for clause-medial LIKE in CanE

INI	χ2-value	df	p-value	Cramér's φ
A2 (16-20)	2.905	1	.08	0.011
A2 (21-29)	4.988	1	<.05*	0.011
A3 (30-39)	5.334	1	<.05*	0.009
A4 (40+)	3.011	1	.08	0.006

The existence and direction of the gender effect is thus confirmed not only by the multivariate analysis, but also by non-parametric t-tests as well as χ^2-test. According to the ICE data, we can confidently ascertain that females in

their twenties and thirties use clause-medial LIKE significantly more often than their male peers.

The weak positive coefficient of the PAI index reported by the multifactorial analysis indicates that LIKE is unevenly distributed across conversations. In contrast to clause-initial LIKE, clause-medial LIKE is not socially stratified, as neither the occupation of speakers nor their first language significantly correlate with its frequency.

7.2.1.3 Clause-final LIKE

Despite being commonly associated with northern varieties of English (Hedevind 1967:237; cf. also Miller & Weinert 1995:368 or Andersen 2000:222), clause-final LIKE is also present in the Canadian data (cf. (79)).

(79) a. Can we go there *like*. (ICE Canada:S1A-044$B)
b. I I don't know *like*. (ICE Canada:S1A-061$A)
c. Oh maybe they thought they could *like*. (ICE Canada:S1A-075$B)

Despite being uncommon, this variant is intriguing, as it differs from clause-medial LIKEnot functionally, but with respect to its direction of scope. It follows that if clause-medial LIKEand its clause-final cousin do indeed share functional properties, then we might expect that the more common clause-medial form is replacing the "'traditional' dialectal usage" (Andersen 2000:226) and causing its demise in CanE. Although not particularly frequent, clause-final LIKE is used by 25 of 244 speakers of CanE. This rather substantial proportion of speakers is inconsistent with the assumption that LIKE with backward scope is becoming obsolete (D'Arcy 2005:5).

Table 51: Age and gender distribution of clause-final LIKE in CanE

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (16–20)	4	0.43	5	0.36	9	0.39
2 (21–29)	5	0.20	1	0.06	6	0.15
3 (30-39)	5	0.11	2	0.10	7	0.11
4 (40+)	3	0.07	1	0.05	4	0.06
SUM	17	0.14	9	0.13	26	0.13

The sociolinguistic profile of clause-final LIKE does not suggest that it is confined to the oldest cohort (D'Arcy 2007:413), as the distribution suggests that its profile approximates the monotone recess common to all functionally and positionally distinct uses.

Due to the overall low number of instances, we will not apply the same protocol as in the previous analyses, but we will only test for possible effects of the date of data compilation using Fisher's exact-tests. If clause-final LIKE is indeed becoming obsolete as suggested by D'Arcy (2005:5), than we would expect to find that it decreases over time, i.e. a negative correlation between clause-final LIKE and the of date of data compilation. The results are inconclusive, but seem to corroborate D'Arcy's claim, as clause-final LIKE is used significantly more from 1990 to 1994 as compared with 1995 to 2001 (D1 v D2: p-value< 0.01**, $CI_{95\%}$= 1.41, 10.22, sample estimates: odds ratio 4.028). On the other hand, this trend is not consistent, as the earliest data do not differ significantly from the latest data (D1 v D3: p-value = 0.497, $CI_{95\%}$= 0.42, 21.99, sample estimates: odds ratio 2.228). Another way to address this question is to take the number of speakers who use clause-final LIKE into account: not asking whether cause-final LIKE has decreased over time, but whether the number of speakers who use it has decreased. Again, the difference between the number of clause-final LIKE users from 1990 to 1994 and 1995 to 2001 differs significantly in the expected direction (D1 v D2: p-value<.01**, $CI_{95\%}$= 1.37, 13.18, sample estimates: odds ratio 4.442). Unfortunately, this difference is again not consistent over time, as the data compiled between 1990 and 1994

does not differ significantly from the data compiled between 2002 and 2005 (p-value = 0.264, $CI_{95\%}$= 0.52, 35.88, sample estimates: odds ratio 3.248)

In summary, clause-final LIKE is rarely used in CanE; it is used by roughly 10 percent of speakers (25 of 244). The analysis of clause-final LIKE fails to corroborate D'Arcy's (2005:5) claim that this form of LIKE with backward scope is becoming obsolete. While clause-final LIKE is used significantly more often in data compiled between 1990 and 1994 compared with data compiled between 1995 and 2001, there is no difference between the use of this LIKE variant in the earliest (1990 to 1994) and the most recent data (2001 to 2005).

7.2.1.4 Non-clausal LIKE

The final analysis investigates the use of non-clausal LIKE, as in (80):

(80) a. I mean I love American crap especially comedies *like* crap comedies that everybody thinks are crap. (ICE Canada:S1A-041$A)

b. It will apparently uhm <„> prolong the <„> *like* the the <„> cleaning that's required ev every two years or whatever <„> I don't actu at a vet's office… (ICE Canada:S1A-023$A)

c. Yeah you know *like* you know ridin around with a horse with the whip and that sort of thing (ICE Canada:S1A-093$A)

The examples in (80) confirm that non-clausal instances of LIKE are functionally indistinguishable from equivalent syntactically unbound uses in AmE. The following passage will probe more deeply into the matter and focus on its sociolinguistic profile.

Table 52: Age and gender distribution of non-clausal LIKE in CanE

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (10 – 24)	14	0.97	8	0.73	22	0.88
2 (25 – 30)	11	0.78	16	0.61	27	0.66
3 (31-40)	17	0.71	28	0.55	45	0.61
4 (41+)	4	0.18	14	0.35	18	0.30
SUM	46	0.64	66	0.50	112	0.55

Table 52 indicates moderate but consistent age-grading, as the frequencies correlate negatively with age in apparent-time. Gender, on the other hand, does not follow a consistent trend. While males use non-clausal LIKE slightly more often than females among speakers younger than 40, the reverse holds true for speakers aged 41 and above.

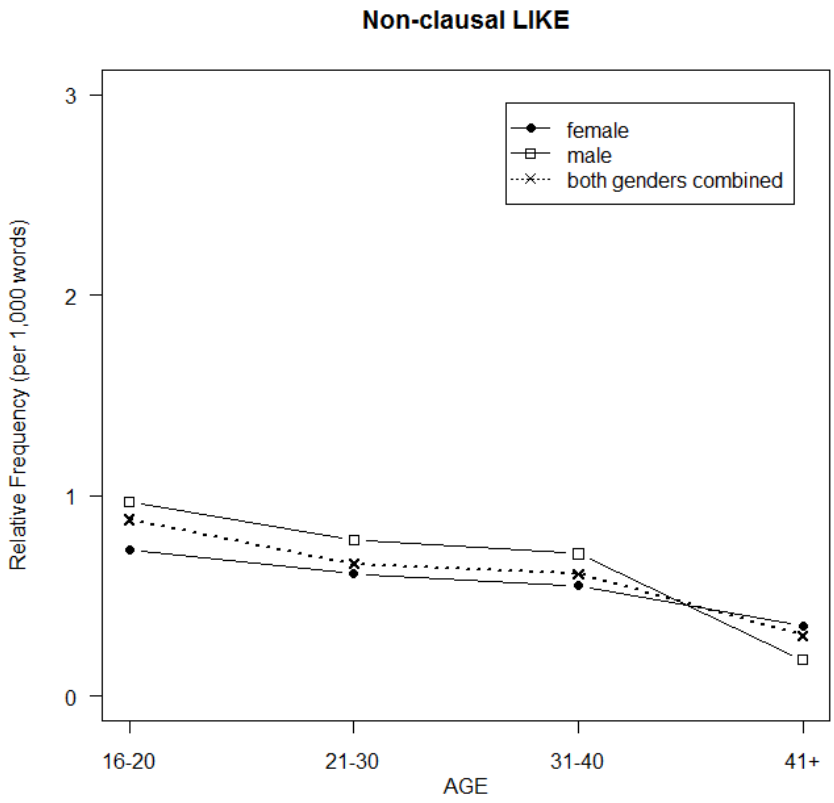

Figure 32: Age and gender distribution of non-clausal LIKE in CanE

Figure 32 shows that non-clausal LIKE is rather infrequent, particularly among speakers over the age of 40. Its use seems to decrease from young to old, but not as drastically as in the case of both clause-initial and clause-medial LIKE. In addition, females use it more frequently, although the proximity of the lines representing male and female rates suggests that gender is probably insignificant.

Table 53: Results of the multivariate regression for clause-initial LIKE in CanE

NON	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.07	0.223	-4.81	<.001***
ACD	-0.55	0.278	-2.00	<.05*
PAI	0.03	0.008	4.32	<.001***

The regression output reports that the strongest effect determining the use of non-clausal LIKE is the occupation of the speakers. When speakers pursue academic careers or work in the professions, they are significantly less likely to use non-clausal LIKE. This may reflect stigmatization of this form among speakers of higher socio-economic status, though other factors such as general fluency and intelligence cannot be dismissed as potential explanations for this choice. In other words, speakers with an overall better linguistic performance, and exhibits lower rates of processing time buyers, will also be situated higher in the socio-economic scale, as both factors correlate with higher general intelligence.

Furthermore, as the PAI index reflects the overall use of LIKE in the respective conversation, the higher the overall use of LIKE, the higher the probability that non-clausal LIKE is used. This indicates a clustering of LIKE use within certain conversations and suggests that the use of this variant depends on the individual bias towards LIKE use in informal conversations. Furthermore, the significant PAI index indicates that the use of non-clausal LIKE is affected by accommodation behaviour and priming effects.

7.2.2 Real-time and apparent-time

In order to evaluate the apparent-time construct, the following section will focus on differences in the data correlating with the date of data compilation to detect systematic differences between data collected from 1990 to 1994; 1995 to 2002; and 2001 to 2005. If the date of data compilation significantly affects the frequency of LIKE in this within-group analysis, this would shed light on the use of LIKE over time. From the apparent-time analysis, we would predict

that LIKE use was less frequent between 1990 and 1994 compared to 2002 to 2005. If this increase in LIKE use over time was the case, then this would support the assumption of the apparent-time construct that speaker age reflects different stages of the communal grammar.

Table 54: Results of the multivariate regression for LIKE use in CanE (real-time)

ALL	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	2.05	0.152	13.50	<0.001***
D2	-0.64	0.171	-3.72	<0.001***
D3	-0.79	0.270	-2.93	<0.01**

The negative coefficients indicate that the frequency of LIKE decreases over time, i.e. it decreases in real-time. This is surprising, as is suggests the opposite of what was expected, i.e. an increase in the frequency of LIKE use over time. Consider D'Arcy (2007:405) for example:

> Since this time, all vernacular forms of like have subsequently increased in frequency to the point where they are now significantly favored among speakers under the age of 30 and disfavored (though not absent) among older age groups. (D'Arcy 2007:405).

When taking the proportion of speakers who use LIKE into account (cf. Table 55) the data even seem to validate the startling finding that LIKE use is decreasing over time.

Table 55: Mean frequencies, number of non-LIKE users to LIKE users plus the resulting ratios according to the date of data compilation

	Mean (LIKE per 1,000 words)	Non-LIKE users	LIKE users	Ratio	LIKE users of all speakers (%)
1990-1994	9.91	1	22	0.05	95.6
1995-2001	4.48	59	150	0.39	71.7
2002-2005	4.12	6	14	0.43	70.0

The distinct decrease in LIKE use raises intriguing questions. However, to understand what is taking place, age must be taken into account (cf. Table 56). Hence, the data is split into different age groups and each age group is analyzed in isolation. Subsequently, the values of the data compiled earlier are compared

with the values of data compiled later. The values are tested for significance using non-parametric t-tests.

Table 56: LIKE use with respect to AGE and the date of data compilation[57]

	A1 (16-20)	A2 (21-30)	A3 (31-40)	A4 (41+)
All variants combined	n.s.	N.A.	n.s.	n.s.
Clause-initial LIKE	n.s.	N.A.	n.s.	n.s.
Clause-medial LIKE	n.s.	N.A.	n.s.	n.s.
Clause-final LIKE	n.s.	N.A.	n.s.	n.s.
Non-clausal LIKE	-1.607.	N.A.	n.s.	n.s.

The fact that the t-tests report insignificant values for all age groups strongly suggests that the previous negative correlation was the product of an underlying, hidden interaction between age and the date of data compilation. In other words, it is not a change in real-time which caused the initial negative correlation, but the fact that younger speakers, who used LIKE readily, were recorded between 1990 and 1994 while older speakers, who use LIKE significantly less often, have been reported between 2001 and 2005. This finding draws attention to the importance to seriously considering the possibility of hidden interactions, thus taking interactions into account even when using less sophisticated mono-factorial designs.

7.2.3 Summary of LIKE use in Canadian English

Vernacular uses of LIKE are a common feature of present-day CanE vernacular. In the ICE data, almost three quarters of CanE speakers used it at least once in informal conversation. Accordingly, this non-standard feature occurs at a substantially higher frequency in CanE than it does in AmE.

In contrast to AmE, LIKE use in CanE is affected more strongly by extra-linguistic variables. This suggests that LIKE use has not yet fully stabilized in this regional variety. The results for AmE showing that LIKE use is neither

[57] N.A.= not available. The data of age group 2 was compiled exclusively between 1995 and 2000.

affected by gender nor by the socio-economic status of speakers do not entirely apply to the Canadian setting, as the overall use of LIKE as well as the use of clause-initial and non-clausal LIKE appears to be class sensitive. The effect of occupation, which serves as a proxy for the socio-economic status of speakers, corroborates a recurring pattern in sociolinguistics more generally. Indeed, the direction of the social stratification is congruent with previous studies reporting that non-standard forms are "most strongly associated with lower working class speakers who do not use […] most advanced forms" (Labov 2001:375). A viable explanation for this result may be related to the overt stigmatization of LIKE within certain social cohorts, as has been noted by D'Arcy (2007:411).

The gender differences corroborate the assumption that LIKE use is undergoing change, as in most changes "women are a full generation ahead of men" (Labov 2001:283). Indeed, the effect of gender strongly suggests ongoing change beyond mere age-grading. Consider the gender paradox and the fourth principle of gender differentiation as proposed by Labov (2001):

> Principle 4
>
> "In linguistic change from below, women use higher frequencies of innovative forms than men" (Labov 2001:292).
>
> Gender Paradox
>
> "Women conform more closely than men to sociolinguistic norms that are overtly prescribed, but conform less than men when they are not" (Labov 2001:293).

Given that LIKE is overtly stigmatized only in the highest social strata, the results of the present analysis match the predictions of both principles. To elaborate: according to principle 4, we should expect to find women ahead of men, if LIKE is indeed undergoing change. This is exactly what we observe. Furthermore, considering the gender paradox, we would expect higher frequencies among males only in higher social strata, i.e. where LIKE is overtly stigmatized. Among older and more conservative speakers, the effect of gender

should thus be substantially weaker, as here the use of LIKE is indeed overtly proscribed. In other strata, i.e. those social strata where LIKE is not stigmatized, women would be expected to be in advance; again, this is exactly what we observe.

With respect to age, the results for the overall frequency of LIKE, clause-initial, and clause-medial LIKE are similar in terms of their degree of age grading, as all tests reported age to be either significant or to be the strongest predictor. In all three cases, the use of LIKE correlated negatively with age, which implies that the older a speaker is, the lower the probability that he or she uses this vernacular feature. Hence, the age stratification conforms to the implicational hierarchy also observable in AmE: A1>A2>A3>A4.

Moreover, the fact that LIKE use is more common in CanE than AmE challenges the widespread belief that this form is a feature typically associated with California, but not with CanE. Though exaggerated, we may say that CanE matches the popular lore surrounding LIKE (cf. D'Arcy 2007) better than AmE. For example, the gender differences observable with respect to clause-initial LIKE and clause-medial LIKEmatch the common assumption that LIKE is typically associated with females (cf. Dailey-O'Cain 2000; D'Arcy 2007; Siegel 2002). In contrast, no such effect of gender is observed for any LIKE variant in AmE. Moreover, the present study does not substantiate D'Arcy's (2005:222-223) assertion stating that clause-medial LIKEis preferred by males, while clause-initial LIKE is thought to be more common among females. Indeed, the present data suggest that both forms are more often used by women, as this observation is consistent across all age groups.

The most contingent patterning of LIKE use in CanE is undoubtedly age-grading. Indeed, all variants of vernacular uses of LIKE consistently decrease in frequency with increasing age. Although the effect of age is statistically insignificant with respect to non-clausal LIKE, Figure 32 clearly shows lower rates of such use as speakers grow older. Thus, the patterning of the present

study corroborates previous analyses of LIKE, which found similar age distributions in CanE (D'Arcy 2005, 2007, 2008).

Furthermore, the use of LIKE in AmE seems to have spread across all social strata, implying that the diffusion it has undergone is nearing completion or has already stabilized. This is, however, not the case in CanE. In CanE, LIKE use still shows signs of social stratification indicative of incomplete diffusion and thus change. Adding weight to this interpretation is the fact that the gender and age distributions of both clause-initial and clause-medial LIKE are almost perfect displays of the patterning which is, according to Labov (1994:65), prototypically associated with female-dominated changes.

Despite its pervasiveness in this variety, LIKE use has not yet entered all strata on the socio-economic scale. The finding that LIKE is not gendered among speakers over the age of 40, and that it correlates negatively with higher socio-economic status indicate overt stigmatization of LIKE in higher social strata. Indeed, D'Arcy's (2007:405) assertion that as with other vernacular forms, LIKE "draws overt attention and commentary" (D'Arcy 2007:405) corroborates this interpretation.

Nonetheless, the real-time analysis fails to confirm that these apparent-time distributions reflect an increase in the use this vernacular feature in real-time. Hence, the distributions could be interpreted as indicative of slow diffusion accompanied by stable age-grading rather than vigorous change. From this perspective, the apparent-time pattern would be interpreted to indicate age-grading in its narrow meaning, i.e. that "individuals change their linguistic behavior throughout their lifetimes, but the community as a whole does not change" (Labov, 1994:84).

Hence, the results of the present analysis challenge the assumption that LIKE use in Toronto English reflects its use in Northern American English more generally (D'Arcy 2007:389).

Although the additional real-time analysis failed to confirm an increase in LIKE use over time, it draws attention to the fact that significant correlations

between main effects may very well be caused by hidden interactions. Indeed, the real-time analysis may serve as a warning to those who neglect testing for interactions, thereby drawing erroneous conclusions about significant correlations.

7.3 LIKE in British English[58]

The earliest instances of LIKE provided in the OED represent instances of clause-final LIKE in EngE dating back the late eighteenth century (D'Arcy 2005:4). This implies that LIKE has a long history in EngE which can reasonably be assumed to be the birthplace of this vernacular feature. Adding weight to this assertion is the fact that in certain regional dialects in the North-eastern varieties of English in the British Isles LIKE is still "used as a focusing device, with different discourse functions according to its position in the sentence. The most traditional function is as an emphatic device in clause-final position" (Beal 2004:136) as in (81):

(81) *I'm* a Geordie, me, *like* (Beal 2004:135)

"In this position it can also be used in interrogatives, where it often conveys a sense of interest or surprise" (Beal 2004:136) as in (82), but it can also occur in clause-initial position to introduce or focus on a new topic as in (83):

(82) How'd you get away with that *like*? (Beal 2004:136)

(83) *Like* for one round five quid, that was *like* three quid, *like* two-fifty each. (Beal 2004:136)

According to Beal (2004), clause-medial LIKEhas entered the speech of "younger speakers, in the North-east […], [where it] is also used within clauses, often as an explanatory device" (Beal 2004:136) as in (84). Moreover, she notes that "discourse *like* can occur several times within one sentence in the speech of younger people North-east of England" (Beal 2004:136) as in (83) above.

(84) They were *like* lightning, as they say, …his legs. (Beal 2004:136)[59]

---

58 For a complete overview of the final data set of EngE used in this study, cf. Table 140 in the Appendix.

59 In fact, the cases of clause-medial LIKE are not considered discourse markers in this study – as is argued in section (4.7.1.2) – but are considered adverbial expressions, as they occur before quantifying expressions and can be replaced by other adverbials, such as *roughly*, *approximately* and *about*.

Anderwald (2004:192) corroborates Andersen's (2001:216) hypothesis that clause-medial LIKE is an American borrowing, at least in south-eastern dialects of England. Indeed, Anderwald hypothesizes that it "is not implausible however that London is the source for the outward spread of these new – imported – uses of *like*" (Anderwald 2004:193). In addition, Anderwald affirms that it is "used almost exclusively by adolescents and young adults" (Anderwald 2004:192). This is congruent with Andersen's finding that "83 percent of the tokens of the pragmatic marker like are uttered by speakers 41 or lower" (Andersen 2001:225).

With respect to its discourse-pragmatic functions, Anderwald (2004:193) remarks that

> "[t]he pragmatic marker *like* has a wide range of functions: it is used in 'ad hoc constructions', i.e. for purposes of approximation and exemplification [as in (85a) and (85b);] *like* is used to construct a metalingistic focus [as in (86)]; it is used as a quotative after BE [...], and, very frequently, it is a hesitational device or a discourse link (Andersen 2001:209-299)." (Anderwald 2004:192-193) as in (87) below:

(85) a. It's just like all sticking out all over the place. (Andersen 2001:237)

b. You know what I mean it's like all plotted (Andersen 2001:237)

(86) It's *like* one day developing, right (Andersen 2001:242)

(87) I know and *like* … on Friday yeah … (Andersen 2001:255)

Anderwald suggests that due to its recent origin, clause-medial LIKE "is not found in FRED material, which dates from the 1970s and the 1980s and contains the speech of mostly older speakers" (Anderwald 2004:192-193). In contrast, the "traditional" clause-final LIKE is attested in the south-east of England, which corroborates "Andersen's hunch that this dialectal *like* is not exclusively a northern phenomenon" (Anderwald 2004:192-193) as in examples (88) and (89) from FRED.

(88) but they're dead and gone now *like*. And eh, I went out with eh,… (FRED LND_003 from Anderwald 2004:193)

(89) Used to come down here *like* and have the day (FRED KEN_001 from Anderwald 2004:193)

With respect to the present ICE data, the cross-varietal overview of LIKE use has shown that LIKE is notably infrequent in EngE. In contrast to the low frequencies attested in the ICE, Andersen's (2000) analysis of LIKE use in London teenage speech has shown that this non-standard feature is abundant among speakers aged 10 to 19. Thus, the present apparent-time analysis is intriguing, as it may serve to answer questions related to the very beginning of LIKE's spread in EngE.

Furthermore, Andersen (2000) has chosen an approach to the analysis of discourse marker LIKE which is similar to the one taken in the present study; he has statistically evaluated the effect of various sociolinguistically relevant variables, such as the gender, age, ethnicity, social class and regional location of speakers. Hence, Andersen has provided a very detailed picture of the sociolinguistic as well as pragmatic factors affecting the use of LIKE among adolescents. The following section will, therefore, aim at adding to the already detailed picture of LIKE use in EngE and complement Andersen's analysis with findings based on a different data set.

Before analyzing the data in detail, we will have a look at the basic statistics of the data to get a first impression of its structure. Of the 320 speakers present in the data, only 73 make use of LIKE, i.e. less than one quarter of EngE speakers have used this marker in the present data. This proportion is remarkably low when compared to other L1 varieties. Nonetheless, this moderate number of LIKE users substantiates Andersen's (2000) hypothesis, according to which LIKE is infrequent among adults (2000:289).

Table 17 reports the basic statistics in the EngE data and provides a first rough impression of LIKE use in this regional variety.

Table 57: Overview - LIKE use in EngE

	N (total)	Mean (LIKE per 1,000 words)	Median (LIKE per 1,000 words)	Standard deviation (s)
LIKE	127	0.490	0.000	1.259

Table 17 shows that LIKE is only used 0.490 times per 1,000 words. The median value confirms that the majority of speakers do not use LIKE at all and that only a small fraction of speakers makes frequent use of this discourse marker (cf. Figure 33).

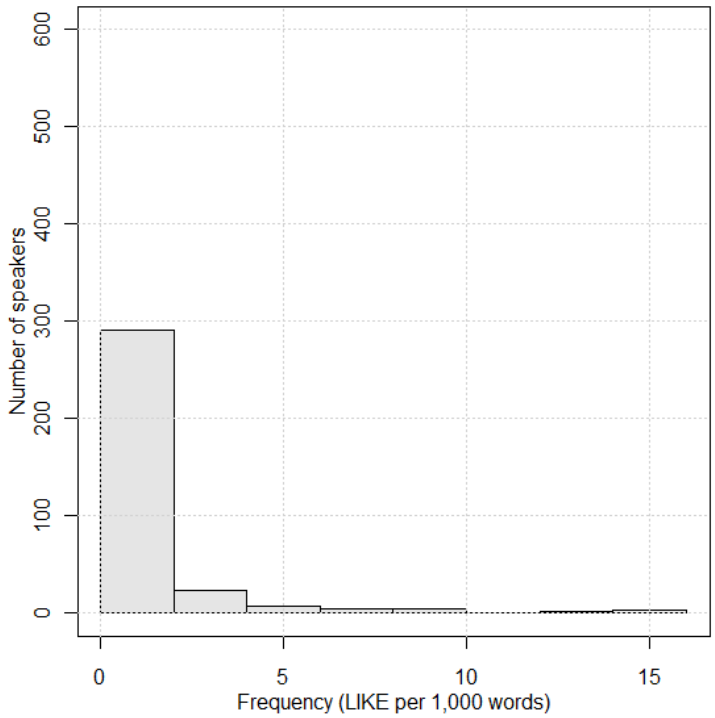

Figure 33: Histogram correlating the number of speakers with their rate of LIKE use

Turning to a more fine-grained analysis of age and gender-specific LIKE use, Figure 34 displays the frequencies of LIKE variants to show the distribution of LIKE variants in EngE.

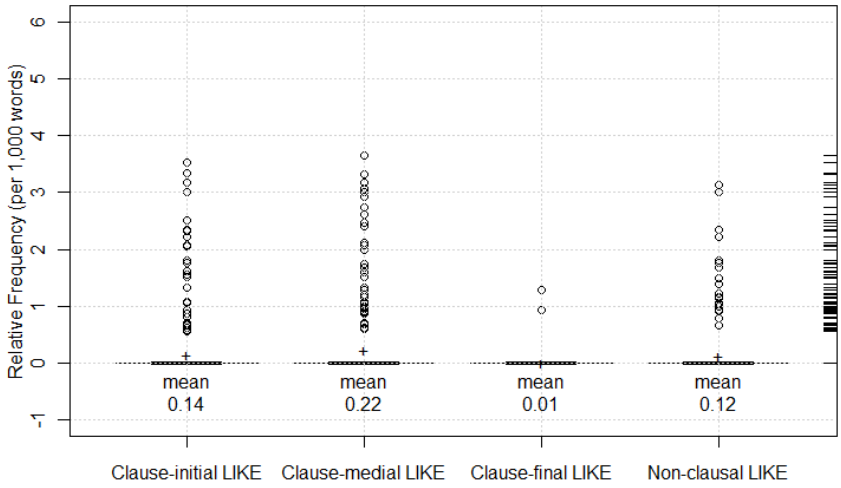

Figure 34: Rate of LIKE variants in EngE

Figure 34 shows that clause-medial LIKEis the most frequently used variant, while clause-final LIKE is rarely used. However, the differences between the mean frequencies are so minute that it appears problematic to draw any conclusions from this distribution other than that all variants of LIKE are extremely rare in the British data.

In his analysis of the COLT, Andersen (2000:221) found an increase in LIKE use in apparent-time, with a steep decline after age 20. According to the COLT data, adolescents between the ages of 10 and 19 use LIKE between 2.53 and 5.61 times per 1,000 words, while young adults (20-29) use it with a frequency of 0.88 instances per 1,000 words. Speaker over the age of 30 (older adults) use it only 0.35 times per 1,000 words (Andersen 2000:289). This finding is strongly supported by the comparatively low frequencies of LIKE use displayed in both Figure 34 and Table 58. Adding weight the stability of Andersen's (2001) finding can be seen in Table 58, which reports a moderate albeit distinct decrease in LIKE use with increasing age. Again, this age stratification is indicative of age-grading and suggests that ongoing change in LIKE use is incipient in EngE.

Table 58: Age and gender distribution of LIKE in EngE[60]

			MALE		FEMALE			
AGE	**Speakers**	**Words**	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (18 – 25)	108	70,376	20	0.52	29	0.61	49	0.57
2 (26 – 45)	134	82,453	25	0.37	23	0.37	48	0.37
3 (46+)	57	37,264	8	0.33	9	0.67	17	0.46
NA	21	11,279	--	--	--	--	13	0.90
SUM	320	201,372	53	0.39	61	0.50	127	0.49

In addition, Table 58 reports a female lead in the use of LIKE up to age 45, whereas male speakers are in the lead among older speakers, i.e. aged 46 and older. The gender difference is, however, minute, and probably insignificant. With respect to gender, Andersen (2000:288) found that female speakers use LIKE significantly more often than males (female speakers 3.24 and male speakers 2.78 instances per 1,000 words).

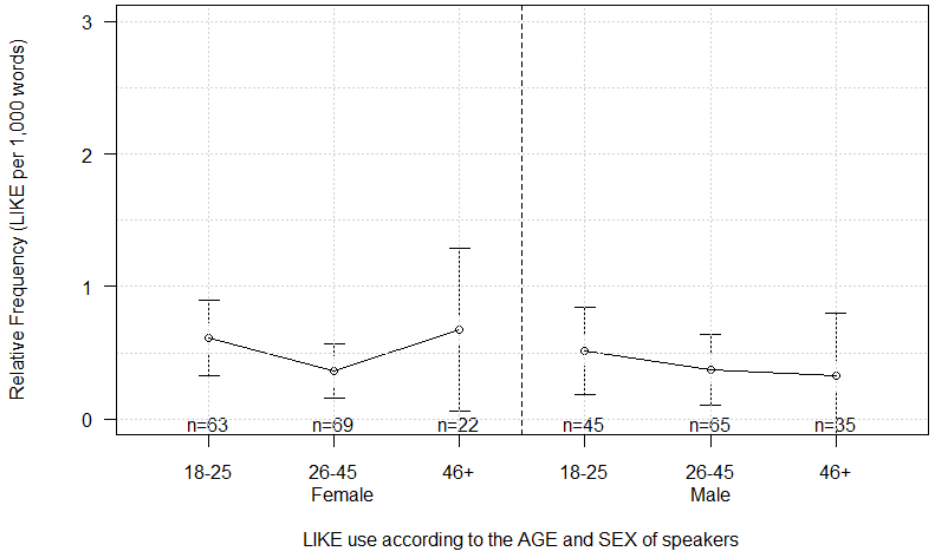

Figure 35: Age and gender distribution of LIKE in EngE

60 Speakers for whom age and gender information is not available have been classified as NA.

Figure 35 shows that the frequency of LIKE is very homogenous among female speakers, as there is no clearly visible or significant trend. The same holds true for males. Hence, Figure 35 strongly indicates an absence of both significant age or gender differences, as the confidence intervals overlap across genders and age groups. Thus, the present data suggest the absence of a female lead in LIKE use. This absence of gender and age differences complements Andersen's (2000) results, as a female lead among adolescents is expected only after a form has become associated with young female speakers as a social reference category. In Andersen's (2000) data, this was indeed the case, as he found a notable increase in LIKE use among speakers younger than 20. In contrast, the present analysis is limited to speakers over age 18. Hence, the ICE data probably reflect the EngE setting before LIKE began to diffuse through the EngE speech community. Nevertheless, Andersen's results are foreshadowed when only LIKE users aged 20 to 32 are considered: the smoothed regression line in Figure 36 clearly indicates a notable increase in LIKE use among these very young speakers.

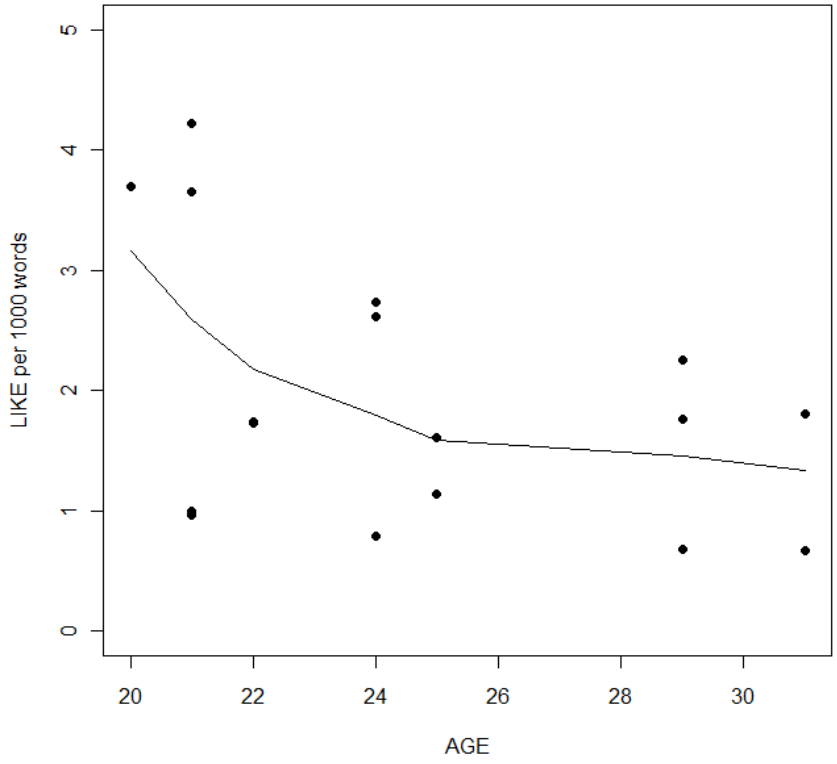

Figure 36: Scatter plot of young LIKE users in EngE including a smoothed regression line

7.3.1 Statistical analysis of LIKE usage in British English

The following statistical analysis intends to evaluate the intuitions derived from the graphical displays and aims to detect general tendencies of LIKE use in the EngE data. As this statistical analysis is more powerful than visual inspections, it may help to uncover correlations which previously escaped detection.

The statistical protocol applied in the present case differs from the protocol used in the analysis of CanE and AmE, as the overall number of LIKE occurrences is extremely low. The low frequency forces us to employ simpler and thus less versatile means to detect patterns in the data. In the present context, only the LIKE uses are considered, i.e. the data contains only the 73 speakers of EngE who have used LIKE at least once. This refined data set is

analyzed using χ^2-tests to detect possible correlations between the age (A1 = 18-25 years, A2 = 26 to 45, A3 = 46 years and older), gender, occupation (OCC), and the frequency of LIKE use and its subtypes.

The results confirm a significant albeit very weak correlation between the gender of speakers and the number of instances of LIKE use. This results suggest that female speakers use LIKE significantly more than expected by chance alone (χ^2=6,599, df = 1, p<0.05*, φ=0.01). The age of speakers proved to be insignificant, as the χ^2-tests reported no significant difference between the expected and observed rates of LIKE. Furthermore, neither the education nor the occupation of speakers significantly correlated with LIKE use.

7.3.2 Summary and discussion of LIKE use in British English

Vernacular uses of LIKE are notably infrequent in the EngE data reflecting speakers aged 18 and older. The proportion of speakers using LIKE (73 out of 320) amounts to about twenty-two percent and suggests that LIKE use is only in the initial incipient phase of change. However, the increase in LIKE use in terms of both frequency and percentage among the speakers aged 20 to 32 indicates that it is starting to gain ground among younger speakers of EngE. This finding complements Andersen's (2001) analysis of the COLT data, according to which LIKE was a frequent feature among young adolescents, while its use drastically declined above the age of 20. This implies that while LIKE use is frequent among adolescents as shown by Andersen (2001), it is marginal among older speakers, as shown in the present study. The absence of age stratification or effects of the socio-economic status of speakers corroborates this hypothesis. Indeed, the absence of social stratification strongly suggests that at the time of data compilation, LIKE had not yet entered the EngE speech community as a whole, but only the speech of EngE speakers younger than approximately 30 years of age. Hence, the ICE data provides an account of LIKE use which predates Andersen's (2000:289); he showed that,

within the COLT data, LIKE use is mostly restricted to adolescents aged 10 to 19. We will return to this point below.

The only significant predictor in the present analysis was the gender of speakers, as female speakers use LIKE slightly but significantly more often than males. This gender difference adds weight to the assumption that the ICE data reflect incipient change of LIKE use, because in its initial stage innovative variants are adopted over-proportionally by female speakers (Labov 1994:65).

Nonetheless, the data fail to differentiate between change on the communal and the individual level. In other words, based on the present results one cannot rule out stable age grading over ongoing generational or communal change. Thus, LIKE use among adolescents may merely serve as an identity marker among adolescents who stop using it as they grow older. On the other hand, this interpretation is problematic with respect to the gender difference unless one presumes that LIKE – at least in EngE – is indeed a marker of femininity or group membership among adolescent girls.

The low frequency of LIKE use among speakers beyond adolescence is somewhat surprising considering the long history of this marker in EngE. As shown in section 4, the earliest attestations of this form date back to the late eighteenth and early nineteenth centuries (D'Arcy 2005:4). The low frequency, therefore, poses a series of challenging questions: Why is it that LIKE is almost gone extinct in the regional variety in which it originated? Has its moderate use remained rather stable for almost two centuries, or has it indeed started spreading in the latter half of the twentieth century? Why is it only now that this form is entering regional varieties in which it was not used before?

A viable explanation for this starling observation is that LIKE has grammaticalized at least twice – once in EngE and once in AmE. According to this account, the spread of LIKE in AmE is independent of the earlier development in EngE. While this cannot be ruled out as a viable hypothesis, it is not sufficient to adequately explain early instances of clause-medial LIKE unless it presumes that LIKE has grammaticalized in similar ways several

times. This hypothesis also fails to explain why there are attestations of clause-medial LIKE "in the speech of septa- and octogenarians living in isolated, rural villages across the United Kingdom" (D'Arcy 2007:399; cf. also Schweinberger 2013).

An alternative explanation relates to what is called "colonial lag" (Görlach 1987). According to Trudgill, colonial lag denotes the "post-colonial survival of earlier phases of mother-country culture, taken in conjunction with the retention of earlier linguistic features" (1999:227). Assuming a possible colonial lag, clause-final LIKE, in particular, and to a lesser extent clause-medial LIKE, may have spread from EngE to SctE, IrE, AusE, and probably NZE. Consequently these forms may have been preserved as fossilizations in these former colonies (Hedevind 1967:237; Miller & Weinert 1995:37), while they almost became extinct in the southern varieties of EngE. In a next step, LIKE could have entered AmE and CanE through IrE via the influx of Irish immigrants. Although this hypothesis can explain the spread of LIKE from the British Isles to the North American continent, it is not sufficient to explain the rapid global spread we observe today.

Hence, the recent global spread is more appropriately attributed to a later development, i.e. the ongoing change initiated by the association of this vernacular feature with a social reference group – quite probably the New York counterculture movement, as proposed by Andersen (2001:216). During the intermediate period between these two phases of regional diffusion, in which non-standard LIKE must have existed as a rather marginal feature, it was not associated with a certain social reference group. Following Labov (2001:24) the association with a certain social reference group is, nonetheless, aprerequisite not only for the initiation of changes themselves and their continuation, but also for linguistic features to diffuse into all social strata:

> The orientation to the relation of language and society that is closest to my own point of view is that of Sturtevant (1947). He viewed the process of linguistic change as the association of particular forms of speaking with the

> social traits of opposing social groups. Those who adopt a particular group as reference group, and wish to acquire the social attributes of that group, adopt the form of speaking characteristic of that group. (Labov 2001:24)

According to Labov (2001), it is, therefore, quite likely that it was the association with certain US American counterculture groups which triggered the recent and rapid spread of LIKE across the English-speaking world today. Of course, this is rather tentative, but it is not implausible, considering the timeline of attestations of LIKE in the literature. The late increase in LIKE use among British adolescents may thus be seen as Americanization, as proposed by Andersen (2001). To speak of "Americanization" is, however, somewhat misleading: Americanization implies passive adoption, which does not seem to match observations concerning semantic and pragmatic modifications during the integration of innovative forms into local linguistic settings. In most cases, innovative features are not adopted wholesale, but they are transformed, re-interpreted, and adapted to the local system. This active process of adapting incoming forms to local practices challenges the assumption that a feature is simply borrowed (for example, Buchstaller 2008; Buchstaller & D'Arcy 2009; Meyerhoff & Niedzielski 2003; Tagliamonte & Hudson 1999).

Following Britain (2002), the present findings can be interpreted to mean that LIKE was until recently rejected outright by speakers of EngE. In other words, EngE speakers refused its implementation into the local system. As a result of prolonged contact, possible hostile attitudes (or the perception of this feature as typically non-British) eroded particularly among younger speakers until it became eventually nativised and re-allocated with locally specific associations. Indeed, this is quite likely accurate, as Buchstaller's (2006b) attitudinal study on *be like* seems to corroborate this hypothesis. While Buchstaller (2006b) did not find that *be like* was particularly linked to AmE (2006b:374), she confirmed that informants who expressed *overt attitudes* towards its local associations did predominantly report it as being a US feature (2006b:374). The extent to which a link between national or local identity and LIKE use may have impacted its use in EngE is, however, uncertain and thus

rather tentative. What is quite certain, however, is the fact that innovative variants undergo re-allocation of attitudes when implemented in local systems. Buchstaller (2006b), for example, states that

> [a]s linguistic resources are borrowed across the Atlantic, they may lose or gain associations during the process, or, alternatively, already existing percepts may be re-analyzed and re-evaluated. Consequently, for speakers of the borrowing variety, new associations interact with possibly second-hand ones and aspects of existing meaning can become more or less salient during the process.

To conclude, the analysis of EngE indicates that LIKE is rarely used among middle-aged and older speakers of this regional variety – possibly as a result of the perception that LIKE is markedly non-British. There are, however, features such as its frequent use among adolescents, and a slight but significant female lead, which point to an imminent increase in LIKE use in EngE. Thus, the results suggest that LIKE in EngE can be located at the lower left of the S-shaped curve representing the rate of change (Labov 2001:450).

7.4 LIKE in Irish English[61]

Despite the fact that IrE has been extensively studied with respect to non-standard features, research on the use of LIKE in this Celtic English variety is rather sparse. Nevertheless, the analysis of IrE is highly promising with respect to variety-specific uses of LIKE, as the IrE variety exhibits markedly idiosyncratic usage patterns such as the high frequency of clause-final LIKE.

With respect to its pragmatic functions and their distribution among the population of IrE speakers, Hickey notes that "[f]ocuser *like* is found in all age groups and is particularly common in *explanatory* contexts" (Hickey 2007:376) as in (90). With respect to its clausal position and without regard to its pragmatic function, LIKE most frequently occurs in clause-marginal position, vastly outnumbering clause-medial uses (Siemund, Maier & Schweinberger 2009:21-30) and is also allowed for in interrogative clauses, as in (91).

(90) a. They'd go into the houses, *like*, to play the cards. (Hickey 2007:376)

b. 'Tis quality now, *like*, and all this milk and everything. You're getting paid on the quality of milk, *like*, and you could lose, *like*, you know ... (Hickey 2007:376)

c. I'm just telling you what I heard, *like*. (Hickey 2007:376)

(91) Did you get to see him, *like*? (Harris 1993:176)

While the aforementioned authors view LIKE predominantly as a marker of focus, Amador-Moreno (2010b:531-544) points out that it serves additionally as a hedging device. In her study on IrE, Amador-Moreno (2010b) compared the use of the discourse marker LIKE in both non-fictional data (LCIE) and fictional texts (*The Curious Incident of the Dog in the Nightdress* by Ross O'Carroll) and concludes that the functional behavior of the discourse marker LIKE is very similar in both registers (Amador-Moreno 2010b:539). In both types of data, LIKE:

[61] For a complete overview of the final data set of IrE used in this study cf. Table 141 in the Appendix.

> indicates fuzzy thought; it is used as a hedge, often with a degree of hesitation involved. […] [M]id-position *like* is used when a speaker is searching for an appropriate expression, or an alternative term. It is also very often used with numbers, as a synonym of 'roughly', 'approximately' and it tend to be employed when a speaker/character is trying to emphasize the expression of certain feelings, when exaggerating, or describing unusual actions/surprising events.

Before analyzing the ICE data in detail, we will have a look at the basic statistics of the data to get a first impression of its structure. Of all 329 IrE speakers, 219 make use of LIKE at least once, i.e. two thirds of speakers have used this marker. This remarkably high proportion is particularly intriguing when compared to its use in EngE, but it also substantiates Schweinberger's (2012) analysis according to which LIKE is highly salient among younger speakers of IrE, in particular.

Table 59: Overview - LIKE use in IrE

	N (total)	Mean (LIKE per 1,000 words)	Median (LIKE per 1,000 words)	Standard deviation (s)
LIKE	936	4.102	2.642	4.893

Table 59 reports that LIKE is used on average 4.102 times per 1,000 words (mean) with the median being 2.642 instances of LIKE per 1,000 words. The difference between mean and median indicates that LIKE use is not distributed homogenously, but that the majority of speakers use LIKE moderately while a smaller subset of speakers makes frequent use of it.

Before turning to a more fine-grained analysis of age- and gender-specific LIKE use in IrE, the frequencies of LIKE variants are displayed to clarify which types of LIKE are particularly frequent.

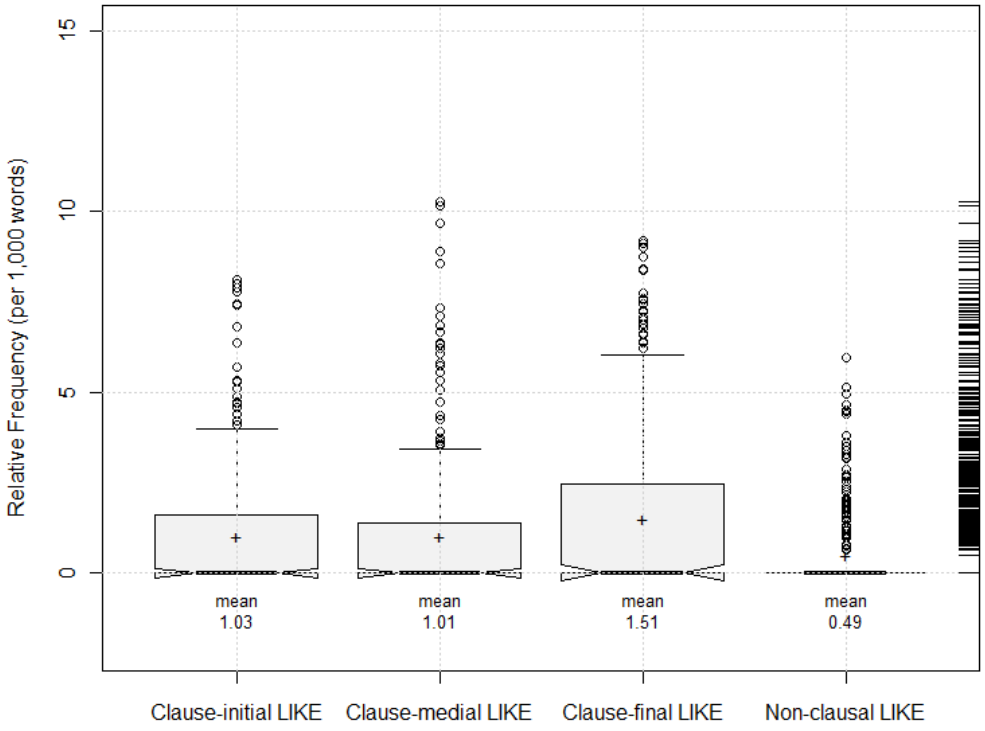

Figure 37: Rate of LIKE variants in IrE

Figure 37 provides a striking example of the diversity of LIKE use across regionally distinct varieties of English: unlike in other varieties, it is the so-called "traditional" clause-final use of LIKE with backward scope that is most frequent.

Table 60: Age and gender distribution of LIKE in IrE[62]

			MALE		FEMALE			
AGE	**Speakers**	**Words**	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (0 – 25)	138	85,963	122	5.20	385	4.48	507	4.88
2 (26 – 33)	52	32,446	17	3.86	207	5.80	224	5.58
3 (34 – 49)	21	13,621	21	4.07	14	1.97	35	2.67
4 (50+)	52	33,002	14	2.45	42	1.40	56	1.66
NA	46	24,755	14	1.71	100	4.39	114	3.52
SUM	309	189,787	188	3.77	748	4.20	936	4.10

Table 60 reports that LIKE use in IrE is remarkably frequent and age-stratified, as younger speakers seem to use it more than the oldest speakers do. Although there appear to be significant gender differences, the data indicate that these are age-specific. To illustrate, females aged 26 to 33 have a substantially higher mean than their male peers, while this trend is reversed among speakers aged 34 to 49.

Figure 38 supports these inferences, but it also draws attention to a strong gender and age bias in the data: the dotted line, representing age-specific use without regard of gender, conforms closely to (younger) female speakers. This strongly suggests that this subsection of the population is vastly over-represented. Accordingly, male speakers older than 26 years of age are only marginally represented, which renders a precise interpretation and generalization of the data difficult.

62 Speakers for whom age and gender information is not available are classified as NA.

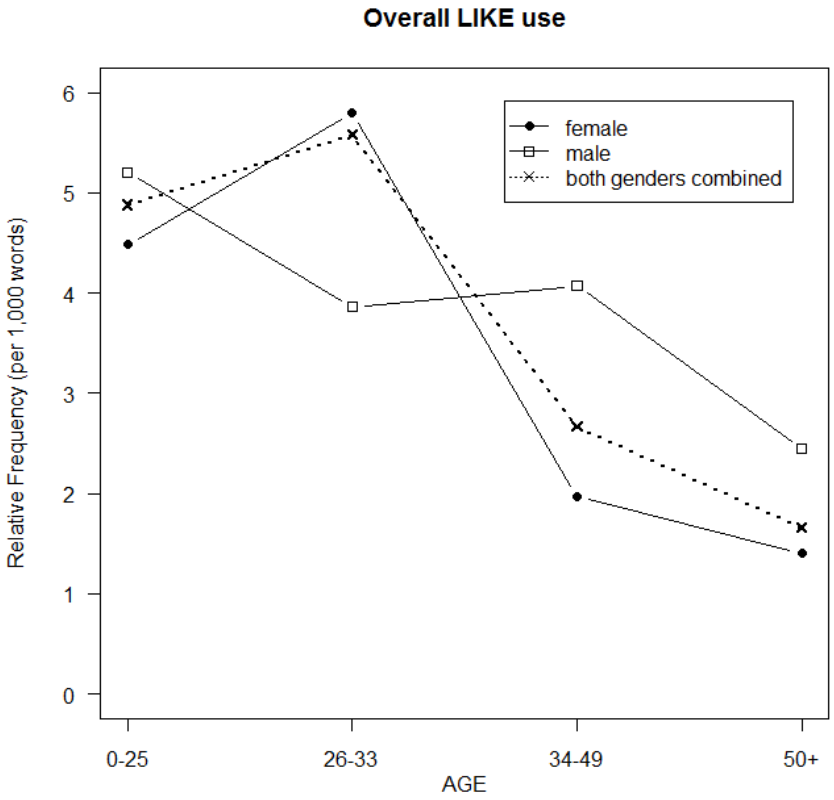

Figure 38: Age and gender distribution of LIKE in IrE

Figure 38 confirms both age stratification and gender differentiation. However, the emerging picture is complex, given that there is neither a consistent lead by males or females nor a linear correlation between LIKE use and the age of speakers. Indeed, among speakers older than 34 years of age it is males who use LIKE most, while this trend is reversed with respect to speakers in their late twenties and early thirties. This distribution can be interpreted in at least three ways:

Firstly, it may be viewed as suggesting that LIKE use is age-graded, i.e. that speakers stop using this element as they grow older. If this were the case, then the statistical analysis should show a lack of social stratification, which was the case when no predictor other than age correlates significantly with LIKE use.

Secondly, the distribution can be interpreted as reflecting ongoing communal or generational change, which would predict that clause-initial LIKE use is substantially socially stratified.

Thirdly, the distribution can be interpreted to reflect a complex interplay among different types of change. This would be corroborated if the statistical analysis showed a moderate degree of social stratification and an increase in LIKE use in real-time, particularly among younger speakers.

The analysis of the overall frequency of LIKE does not offer insight into differences between the use of distinct types. Taking the diversity of uses into account is, nonetheless, particularly important with respect to IrE, as Table 19 reported that LIKE use in IrE differs markedly from other regional varieties. One case in point is that IrE speakers make over-proportional use of clause-final LIKE, which is very infrequent in other regional varieties. In fact, this clause-final type of LIKE has been described as becoming obsolete in North American varieties of English (cf. D'Arcy 2005:5 and the sources quoted therein). This tendency does not, however, apply to IrE. Schweinberger (2013) points out that clause-final LIKE is, in fact, the most frequent use of LIKE among young speakers of IrE. Hence, there are neither signs that clause-final LIKE is becoming obsolete in IrE, nor is there evidence suggesting that the so-called innovative clause-medial LIKE may be replacing this supposedly traditional use of LIKE (Schweinberger 2012).

The present data add to Schweinberger's (2012) interpretation, as Figure 39 confirms that clause-final LIKE is used significantly more by speakers aged 25 and younger than by speakers aged 50 and over.

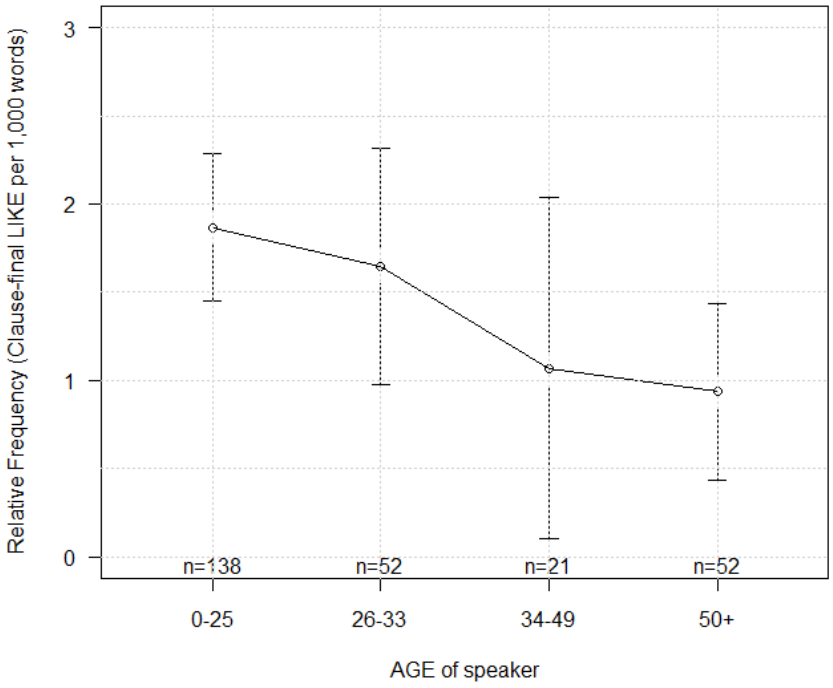

Figure 39: Age distribution of clause-final LIKE in IrE

Given the substantial numbers of clause-final LIKE in the speech of younger cohorts, it seems implausible that it is becoming obsolete, as in CanE (D'Arcy 2005:5).

The high frequency of clause-final LIKE raises the question of whether clause-medial LIKE is similarly distributed across the age cohorts or whether they occur in complementary distribution. In other words, clause-medial and clause-final LIKE may be functional rivals, or they may co-exist as functionally distinct elements.

Figure 40 corroborates the latter statement, as clause-medial uses are similarly stratified across age groups and thus age-graded, as are uses of clause-final LIKE. In contrast to the near-linear recess visible in Figure 39, clause-medial uses of LIKE peak among speakers aged 26 to 33.

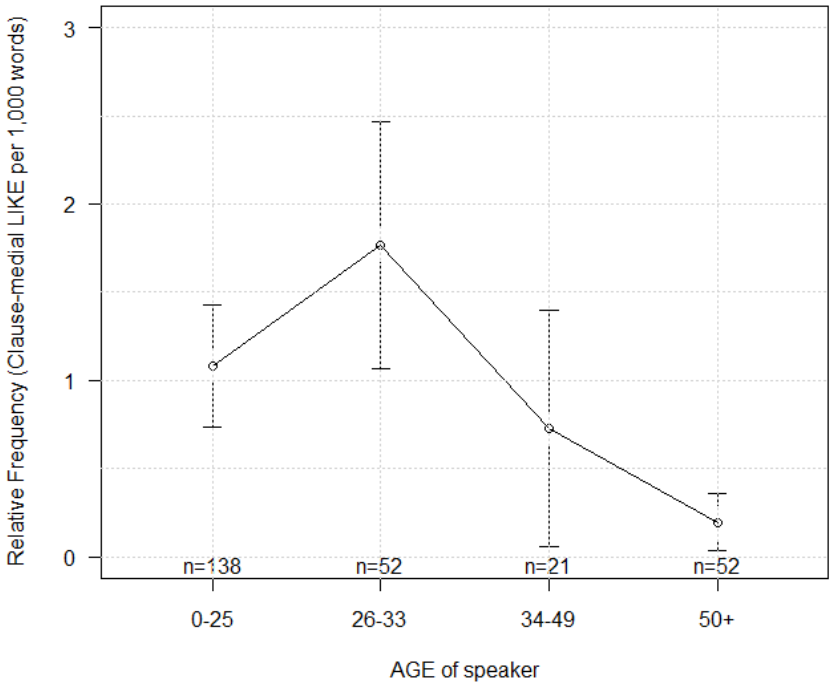

Figure 40: Age distribution of clause-medial LIKE in IrE

7.4.1 Statistical analysis of LIKE usage in Irish English

As the following statistical analyses are more powerful than mere visual inspections, they aim to uncover correlations which previously escaped detection. In addition, they serve to evaluate the intuitions derived from the graphical displays.

As before, (Quasi-) Poisson regressions are able to deal with extremely heterogeneous variance distributions caused by small numbers of instances in a given interval (in this case 1,000 words). By applying this protocol, it is possible to accurately study the effect of age (A1 = 0-25 years, A2 = 26 to 33, A3 = 34 to 49, A4 = 50 years and older), gender, education, and occupation (OCC) on the frequency of LIKE use. The initial saturated model is fitted in a step-wise procedure and arrives at a minimal adequate model that contains only significant predictors. The same procedure is applied to study the effect of

age, gender, and occupation on the frequency of clause-initial LIKE, clause-medial LIKE, and syntactically unbound LIKE without scope LIKE (non-clausal LIKE).

Table 61: Results of the multivariate regression[63] for LIKE use in IrE

ALL	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	1.58	0.093	16.99	<0.001***
A2	0.12	0.152	0.85	.39
A3	-0.59	0.316	-1.86	<.1
A4	-1.28	0.208	-6.176	<0.001***

The regression model predicting the frequency of LIKE per 1,000 words in IrE based on the age of speakers is statistically significant (χ^2 = 31.370, df=3, p<.001***). In addition, the regression output reports that age is the only significant variable as it confirms a negative correlation between the frequency of LIKE used by speakers aged 25 and younger (the reference group) and the frequency of LIKE used by speakers aged 50 and older. Furthermore, the order of the age coefficients indicates a curvilinear distribution of LIKE use. While the coefficients indicate a steady increase in LIKE use from old to young among speakers aged 26 to 50 and older, the positive coefficient of age group 2 signifies that LIKE use peaks among speakers who are in their mid-twenties to mid-thirties. In other words, the regression indicates the existence of an implicational hierarchy of the form A2>A3>A4.

To confirm this implicature indicated by the regression analysis, non-parametric one-tailed t-tests are applied to evaluate its robustness. As the highest use of LIKE is by speakers in their late twenties and early thirties, age group 2 will serve as the reference category (cf. Table 62).

63 During the following statistical analyses, Poisson regression models serve as default statistics, while non-parametric t-tests are employed whenever model assumptions are not met, i.e. when the number of observations is too low to allow for a meaningful interpretation of the regression model.

Table 62: LIKE with respect to AGE

	A2-A1	**A2-A3**	**A2-A4**
All	n.s.	2,65**	5,20***

The results of the t-test confirm the existence of an implicational hierarchy of the form A1≤A2>A3; A2>A4.

7.4.1.1 Clause-initial LIKE

The following section concerns itself with clause-initial uses of LIKE as in (92).

(92) a. What else is sold at that market <#> *Like* I've never been at it. (ICE Ireland:S1A-010$B)

b. And *like* she'd seen him about two days before that. (ICE Ireland:S1A-055$B)

c. Grainne used to always get to do the dishes and *like* I hated doing the dishes (ICE Ireland:S1A-063$B)

Functionally and positionally clause-initial LIKE is indistinguishable from similar instances in other varieties. Instances of clause-initial LIKE typically occur either utterance initially without being preceded by a connector, as in (92a), or between the connector and the following utterance, as in (92b), and (92c). All instances in (92) introduce explanatory specifications, which concurs with Miller's (2009:30-32) analysis of clause-initial uses of LIKE in NZE and AusE.

The following section will focus on the age and gender distribution of clause-initial LIKE and provide an advanced statistical analysis of its sociolinguistic profile.

Table 63: Age and gender distribution of clause-initial LIKE in IrE

AGE	MALE		FEMALE			
	LIKE (N)	LIKE (mean)	LIKE (N)	LIKE (mean)	SUM (N)	ALL (mean)
1 (0–25)	25	0.99	111	1.33	136	1.25
2 (26–33)	4	0.68	59	1.51	63	1.41
3 (34-49)	4	0.77	3	0.42	7	0.54
4 (50+)	1	0.13	13	0.36	14	0.30
SUM	34	0.74	186	1.12	220	1.04

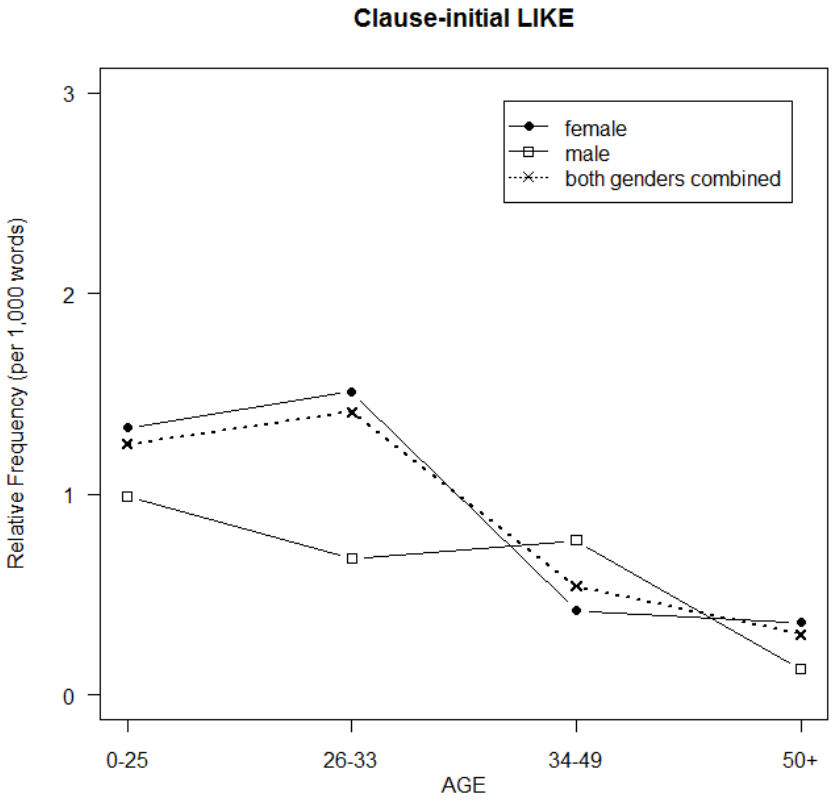

Figure 41: Age distribution of clause-initial LIKE in IrE

Table 63 and Figure 41 indicate that the use of clause-initial LIKE decreases as age increases. In addition, the distribution indicates the existence of a gender difference in the use of clause-initial LIKE among speakers in their late twenties to early thirties.

Table 64: Results of the multivariate regression for clause-initial LIKE in IrE

INI	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.50	0.172	-2.93	<.01**
A2	0.02	0.218	0.12	.9
A3	-0.64	0.357	-1.79	<.1
A4	-1.18	0.334	-3.54	<.001***
PAI	0.04	0.006	6.16	<.001***

The regression output reports two correlations: a positive correlation between the frequency of clause-initial LIKE and the PAI index and a negative correlation between the age of speakers and the use of clause-initial LIKE. The increasing values of the age coefficients again suggest that LIKE use increases from old to young, with a peak between ages 26 and 33. With respect to the PAI index, the regression model reports a minute but highly significant correlation between the overall frequency of LIKE in the respective conversation and the occurrence of the clause-initial discourse link. The effect is, however, so minute that it lacks any substantial impact. Again, neigther the occupation of speakers nor region correlate with clause-medial LIKE use, indicating that clause-initial LIKE is common among all social strata and in regions in Ireland.

To confirm the implicature indicated by the order of age coefficients, non-parametric one-tailed t-tests are applied to evaluate its robustness. As the highest use of LIKE is by speakers in their late twenties and early thirties, age group 2 will serve as the reference category (cf. Table 65).

Table 65: LIKE with respect to AGE

	A2-A1	A2-A3	A2-A4
INI	n.s.	2,33*	2,92**

The results of the t-test confirm the existence of an implicational hierarchy of the form A1≤A2>A3; A2>A4.

7.4.1.2 Clause-medial LIKE

The next section deals with clause-medial LIKEas in (93).

(93) a. They have they take over *like* a suite just for themselves so … (ICE Ireland:S1A-012$A)

b. [T]hey 've got the the best screens in the window because they 're just plain wooden frames with *like* a white gauze on them (ICE Ireland:S1A-013$A)

c. [N]o there is ones <,> there is ones actually that uh <,> you know they 're set up so you can like learn the Jewish language and all this here (ICE Ireland:S1A-014$B)

The clause-medial variant of LIKE in IrE is not only functionally equivalent to clause-medial LIKE in AmE, CanE, and EngE, but also syntactically. In each instance, LIKE precedes the element it modifies to its right. Despite its functional and syntactic similarity, clause-medial LIKE's sociolinguistic profile suggests regional variation and a rather distinct age and gender distribution.

Table 66: Age and gender distribution of clause-medial LIKE in IrE

	MALE		**FEMALE**			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (0–25)	24	1.01	93	1.09	117	1.08
2 (26–33)	4	1.18	60	1.83	64	1.76
3 (34-49)	8	1.24	5	0.46	13	0.72
4 (50+)	0	0.00	7	0.25	7	0.19
SUM	36	0.83	165	1.12	201	1.01

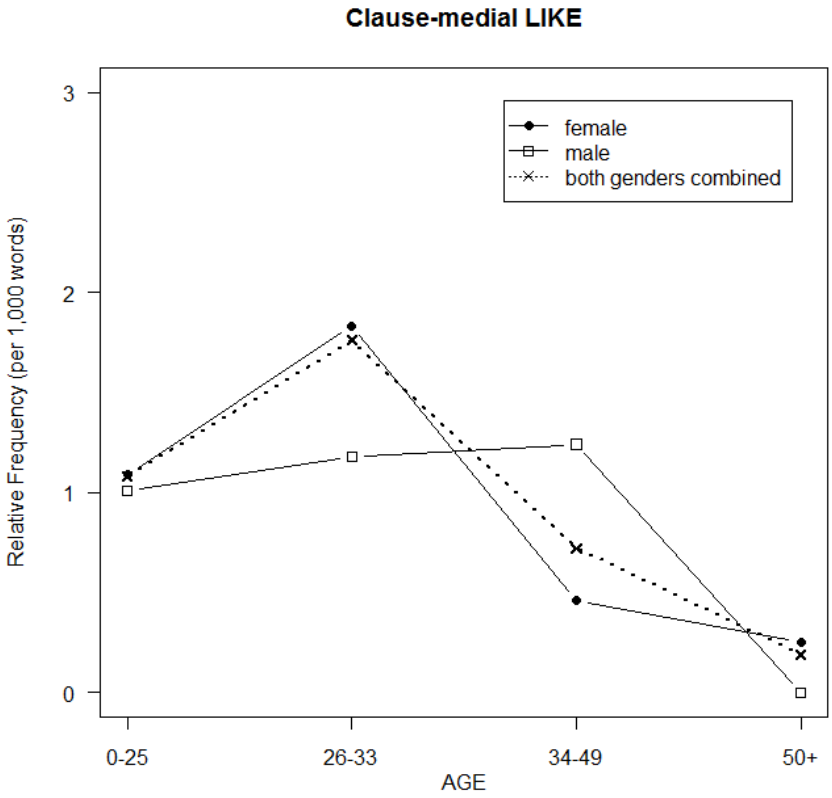

Figure 42: Age distribution of clause-medial LIKE in IrE

Figure 42 suggests that male and female speakers differ substantially in their overall rate. The two genders also seem to follow quite distinct gender-specific trajectories. Among males, age seems to lack a distinct effect, as the mean frequencies follow a straight line until its frequency plummets among speakers over 50 years of age. Among females, the use is clearly age-graded with a peak among female speakers in their late twenties to mid-thirties.

Table 67: Results of the multivariate regression for clause-medial LIKE in IrE

MED	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.35	0.230	2.30	<0.05*
A2	0.06	0.289	0.230	.81
A3	-0.58	0.488	-1.19	.23
A4	-1.19	0.441	-2.71	<.01**
PAI	0.04	0.009	5.25	<.001***
REG: South	-0.50	0.270	-1.88	<.1
A4*SEX: Male	-15.09	0.636	-23.71	<.001***

The final regression model reports various correlations between social stratification and the use of clause-medial LIKE. The results of the statistical analysis do not confirm the gender and age distribution visible in Figure 42. Instead, the multivariate analysis indicates that the gender difference is caused by variables which have not been considered in Figure 42, such as geographical region or the PAI index.

Nonetheless, the increasing negative values of the coefficients for age groups 3 and 4 indicate that clause-medial LIKE use decreases with increasing age once speakers pass their mid-thirties. If gender differences are neglected, then the order of age coefficients again conforms to an implicature of the form A1≤A2>A3; A2>A4. To evaluate the robustness of this implicature, we apply non-parametric, one-tailed t-tests.

Table 68: Clause-medial LIKE with respect to AGE

	A2-A1	A2-A3	A2-A4
MED	1,79*	2,23*	4,63***

The results of the t-test confirm the existence of an implicational hierarchy of the form A1<A2>A3; A2>A4.

In addition to age stratification, the regression analysis reports a significant, albeit weak, correlation between clause-medial LIKE and the PAI index, suggesting that clause-medial LIKEis mildly affected by priming and accommodation effects. A rather unexpected finding is the negative coefficient

of region, which indicates that speakers of southern IrE use this variant less than speakers in NIE (cf. Table 69).

Table 69: Clause-medial LIKE in IrE with respect to region

	Mean (LIKE per 1,000 words)
Northern Ireland (NIE)	1.21
Republic of Ireland (Southern IrE)	0.84

Finally, the model reports that in contrast to females age 50 and above, clause-medial LIKE is almost never used by their male peers. The fact that males do not use this vernacular form leads to a negative value of the interaction between clause-medial LIKE and age group 4. This age-specific gender difference in apparent-time is suggestive, as it indicates later adoption of clause-medial LIKEby males.

7.4.1.3 Clause-final LIKE

We now turn to traditional use of LIKE in clause-final position as in (94). In her extensive discourse-analytic study, Columbus (2009) asserts that utterance-final LIKE functions as an invariant tag which is functionally versatile, but commonly marks focus. Hence, her analysis corroborates Harris' (1993) assertion that LIKE in IrE commonly serves to "focus on a constituent or command the listener's attention" (Harris 1993:176).

(94) a. He was really a bit of a boring prat *like* he was (ICE Ireland:S1A-019$C)
b. You know big Jane *like*. (ICE Ireland:S1A-024$C)
c. I don't mind getting defeated once *like*. (ICE Ireland:S1A-027$A)

The representative examples of clause-final uses of LIKE in (94) validate previous qualitative analyses as, for example, Columbus' detailed discourse-analytic investigation. Indeed, there appears to be no notable difference between clause-final uses of LIKE in IrE and its use in other regional varieties such as CanE. In the following, the focus is placed on the sociolinguistic profile

of this form to determine which subsections of the Irish population make over- or under-proportional use of this form.

Table 70: Age and gender distribution of clause-final LIKE in IrE

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (0–25)	49	2.24	128	1.75	177	1.86
2 (26–33)	5	1.22	67	1.70	72	1.64
3 (34-49)	7	1.65	4	0.77	11	1.06
4 (50+)	12	2.18	13	0.51	25	0.93
SUM	73	2.05	212	1.43	285	1.57

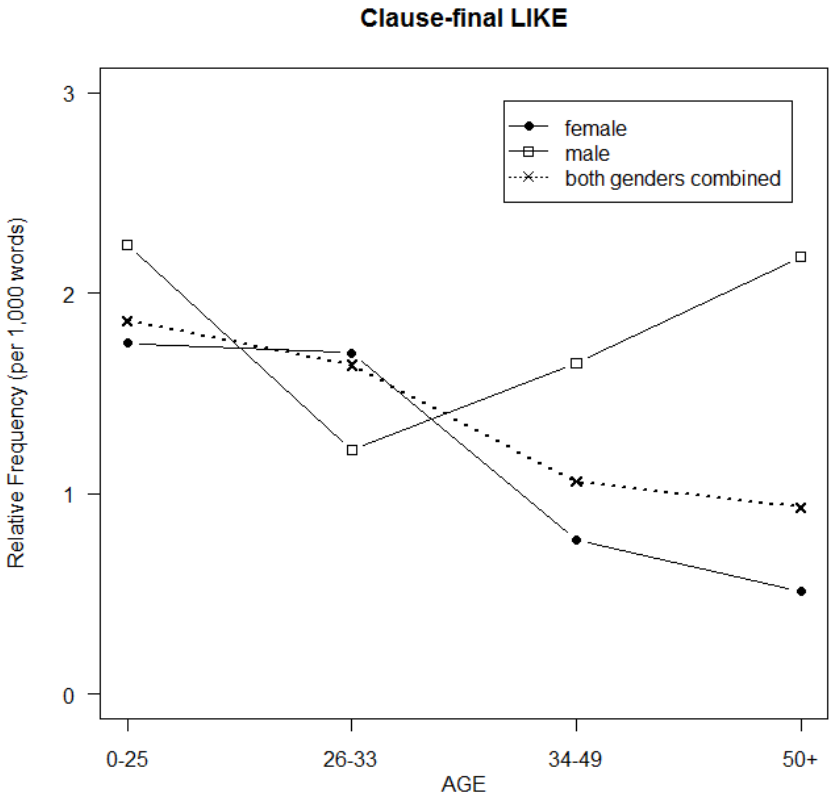

Figure 43: Age distribution of Clause-final LIKE in IrE

The patterning shown in Figure 43 is not particularly striking with respect to females, but exceptional with respect to males. Female speakers exhibit the common age-graded pattern in apparent-time with a steady decrease in the use

of clause-final LIKE use as speakers grow older. The age stratification of clause-final LIKE among males, on the other hand, is distinctly different: for males older than 26 years of age, there is a positive correlation between age and the use of clause-final LIKE. In other words, the use of this non-standard form increases as age increases. This antithetical patterning of male and female rates does not conform to the patterns observed so far. Among males, young and old speakers use it equally often, while speakers who are in their twenties seem to disprefer this vernacular feature. This patterning indicates that clause-final LIKE has been a typical male feature in IrE. This male preference has waned, however, as female rates have increased.

The following multivariate analysis is, therefore, necessary to provide a more fine-grained analysis of this atypical distribution.

Table 71: Results of the multivariate regression for clause-final LIKE in IrE

FIN	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.21	0.151	-1.41	.15
A2	-0.12	0.214	-0.59	.54
A3	-0.44	0.361	-1.23	.21
A4	-0.51	0.228	-2.25	<.05*
SEX: Male	0.28	0.182	1.56	.11
PAI	0.04	0.004	8.60	<.001***

The results reported by the regression model evaluating the social stratification of clause-final LIKE show that neither occupation nor region significantly influence its use, while age and gender do. In contrast to what might be expected if it were assumed that this traditional variant is becoming obsolete, it is younger speakers who use this form most. In fact, the progression of the negative coefficients for age shows a steady and almost linear correlation of clause-final LIKE use and age. However, Figure 43 reports that age coefficients decline only among females, while LIKE use among male speakers is more or less constant, with a distinct decline among males in their late twenties to early thirties.

Figure 43 indicates that the use of clause-final LIKE is extremely gender sensitive, while the multifactorial analysis does not report a gender difference that is independent of the speaker's age. Indeed, the only significant gender difference detected by the regression model is restricted to speakers who are 50 years and older.

Table 72: Gender differences in the use of clause-final LIKE with respect to AGE

	A1	**A2**	**A3**	**A4**
FIN	n.s.	n.s.	n.s.	-2.044*

The t-tests confirm the results of the regression model, as they report that only age group 4 differs significantly form age group 1(cf. Table 72).

7.4.1.4 Non-clausal LIKE

So far, studies of LIKE in IrE have not particularly focused on instances which indicate planning difficulties or serve to introduce specifications as exemplified in (95).

(95) a. But you see technically then that doesn't count because in up in Derry there's *like* what there's *like* I think there's four or five (ICE Ireland:S1A-030$C)

b. Mmm I know <,> but *like* <,> but you wouldn't feel comfortable. (ICE Ireland:S1A-057$A)

c. Uhm <,> I don't know UCD *like* first of all well well UCC supposedly it's meant to be easier to get into second year Psychology. (ICE Ireland:S1A-048$B)

As in AmE, CanE and IndE, non-clausal uses of LIKE primarily serve as floor-holding devices and thus function predominantly as a processing time buyer in case the speaker experiences planning difficulty. We will now turn to the age and gender distribution and statistically evaluate the sociolinguistic profile of this non-standard use of LIKE.

Table 73: Age and gender distribution of non-clausal LIKE in IrE

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (0–25)	24	0.94	47	0.52	71	0.62
2 (26–33)	4	0.75	17	0.57	21	0.59
3 (34-49)	1	0.18	2	0.30	3	0.26
4 (50+)	1	0.13	6	0.18	7	0.17
SUM	30	0.65	72	0.45	102	0.50

The use of non-clausal LIKE has not received scholarly attention, which is unfortunate given that it differs markedly from the variants discussed above. In contrast to clausal uses of LIKE, this non-clausal form appears not at all or only mildly sensitive to age (Table 74). Nonetheless, its use is notably elevated among young speakers, indicating that its frequency decreases with increasing age. In addition, Figure 44 indicates that there is a weak effect of gender among young speakers: below the age of 25, females seem to use it more than their male peers.

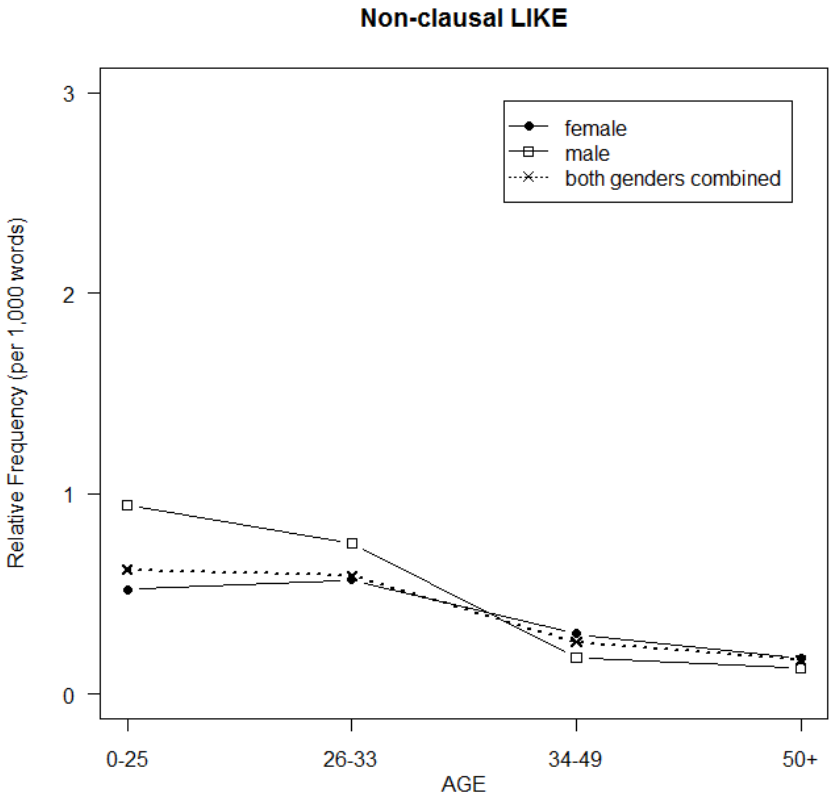

Figure 44: Age distribution of non-clausal LIKE in IrE

The age distribution in Figure 44 suggests moderate age-grading and a minor lead among male speakers younger than 33 years of age.

Table 74: Results of the multivariate regression for non-clausal LIKE in IrE

NON	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(intercept)	-2.16	0.456	-4.73	<.001***
ADC	1.06	0.434	2.44	<0.05*
PAI	0.03	0.013	2.72	<0.01**

The output of the regression model shows that non-clausal LIKE exhibits neither a significant effect of age nor sex. In fact, the only sociolinguistic correlation reported by the model is one between the occupation of speakers and their rate of syntactically unbound LIKE. This correlation implies that speakers with clerical and managerial occupations use non-clausal LIKE less than those pursuing academic careers and those working in the professions. This is particularly surprising given that LIKE is commonly portrayed as a sub-

standard or vernacular feature. It could be that LIKE use is taken to be typically Irish and, thereby, serves as marker of national identity, particularly among well-educated speakers. Though identity marking, i.e. the association with certain socio-cultural concepts is well attested (Dailey-O'Cain 2000), the use of LIKE as a marker of national identity is tentative. Confirming this association would require further research by means of e.g. attitudinal studies. In addition to socio-economic status, the PAI index has a significant but very weak effect. The obvious interpretation is that the more LIKE is used within a conversation, the higher the probability that non-clausal LIKE will occur as well. This is not surprising, as informal conversations are often marked by disruptions, incomplete clausal structures such as ellipses, and phrasal specifications (Miller & Weinert 1998).

Since regressions are systematically conservative, they are complemented with χ^2-statistics to confirm the impact of age. The reason for this procedure is that despite lacking a clearly visible trend among females, Figure 44 indicates moderate age-grading among males. This impression, however, is not confirmed statistically. The results of the χ^2-tests do not provide enough evidence to confidently reject the null-hypothesis, i.e. that the trend is merely incidental (Table 75).

Table 75: Results of the χ2-test evaluating the effect of age for non-clausal LIKE among male speakers in IrE[64]

NON	χ2-value	df	p-value	Cramér's φ
A2 –A3(26-49)	2.709	1	.09	0.009
A4 (50+)	1.268	1	.26	0.004

7.4.2 Evaluation of the apparent-time construct

As we have done with the CanE data, the following section will focus on differences linked to the date of data compilation to evaluate the apparent-time

[64] Age group 1 (0-25) serve as the reference group. Only speakers whose age was given are included and speakers of age group 2 (26-33) and age group 3 (34-49) are collapsed.

hypothesis. If the date of data compilation significantly affects the frequency of LIKE in this within group analysis, then this would shed light on the use of LIKE over time. From the apparent-time analysis, we would predict that LIKE was less frequent from 1990 to 1994 compared with 2002 to 2005. If this increase in LIKE over time was the case, it would support the assumption of the apparent-time construct that speaker age reflects different stages of the communal grammar.

The regression models predicting the overall frequency of clause-initial, clause-final, and non-clausal LIKE based on the date of data compilation failed to provide meaningful results. This suggests that the present data fails to substantiate the assumption that the observable age stratification in apparent-time reflects real-time changes. However, the frequencies of clause-medial LIKE do significantly correlate with the date of data compilation.

Table 76: Results of the multivariate regression for clause-medial LIKE use in IrE

MED	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.19	0.144	-1.34	.17
D3	0.53	0.217	2.47	<.05*

The positive coefficient indicates that in IrE, the frequency of clause-medial LIKE has increased in real-time. As Table 77 shows, the frequency has almost doubled, which is remarkable considering the short time-span covered by the real-time data.

Table 77: Clause-medial LIKE use with respect to the date of data compilation

	Mean (LIKE per 1,000 words)	Non-LIKE users (N)	LIKE users (N)	Ratio	LIKE users of all speakers (%)
1990-1994	0.82	150	59	0.39	28.2
2002-2005	1.40	59	41	0.69	41.0

In contrast to the CanE data, the IrE data supports the apparent-time hypothesis with respect to clause-medial LIKE. The result suggests that in addition to age-grading, there is ongoing change on the communal level in real-time. This hypothesis is corroborated by the dramatic increase in speakers

using clause-medial LIKE: while only 39 percent of speakers used this form in the data collected from 1990 to 1994, nearly 70 percent of speakers have used it in the data collected from 2002 to 2005. That it is only clause-medial LIKE, which correlates positively with the date of data compilation, further indicates that all other variants have stabilized. This, on the other hand, challenges the notion that the dialectal, traditional clause-final variant is generally becoming obsolete, which is obviously not the case in IrE.

The increase in the use of clause-medial LIKEseems to corroborate the assumption that it has only recently begun to spread among speakers of IrE and is, thus, still diffusing through more conservative social strata.

The extent of the increase in the use of clause-medial LIKE, moreover, suggests that other uses may also have increased, but that the multifactorial design was either too coarse-grained to detect real-time changes, or that these changes are restricted to certain subsections of the speech community. Hence, non-parametric t-tests are applied in order to test for weak effects of the date of data compilation.

Table 78: LIKE use with respect to AGE and the date of data compilation

	A1 (0-25)	A2 (26-33)	A3 (34-49)	A4 (50+)
ALL	-1.36·	-3.13**	n.s.	n.s.
INI	-1.60·	n.s.	n.s.	n.s.
MED	-1.48·	-3.00**	-1.50·	n.s.
FIN	1.39·	-3.22**	n.s.	n.s.
NON	-2.29*	n.s.	n.s.	n.s.

Table 78 provides interesting insights about the boundednesss of ongoing change in real-time: the t-statistics confirm that the change is confined to younger cohorts, and the negative values indicate that the change has progressed in the expected direction. In other words, LIKE increased between 1990 and 2005 in all cases but one – the use of clause-final LIKE has decreased among speakers younger than 25 years of age. The fact that significant values are reported predominantly for younger speakers strongly suggests that this ongoing change has affected predominantly younger speakers, while the use of

LIKE among older speakers has not notably changed over time. The most dramatic rise is observable among speakers aged 26 to 33, as the overall frequency of LIKE as well as the frequencies of clause-medial and clause-final LIKE, have significantly increased. In summary, the real-time analysis confirms that the apparent-time approach was justified. The use of LIKE has notably increased over time, particularly among younger speakers.

7.4.3 Summary and discussion of LIKE use in Irish English

The use of discourse marker LIKE in IrE has turned out to be markedly different from its use in the regional varieties discussed so far. Uses of this non-standard feature in all its forms are abundant in IrE and represent a salient feature of contemporary IrE vernacular: two thirds of speakers (219 of 329) in the Irish ICE-based data used this form at least once.

An outstanding aspect of LIKE use in IrE is the high frequency of the so-called "traditional" variant of LIKE with backward scope. Although the fact that this archaic form is present in IrE is not surprising in itself – particularly given the number of previous accounts of this form (Columbus 2009; Harris 1993; Siemund, Maier & Schweinberger 2009) – it is its frequency and its distribution that are noteworthy. In contrast to its use in CanE, where clause-final LIKE has been predicted to become obsolete (D'Arcy 2005:4), it is clearly not in IrE.

Though thriving with respect to frequency, clause-final LIKE exhibits signs of overt stigmatization, particularly among older speakers. The elevated use of this variant by older males suggests that it is perceived as substandard or rural within the older population. Indeed, its sociolinguistic profile is congruent with a re-occurring pattern observable for stigmatized elements: non-standard forms are more common among males, while females avoid overtly stigmatized forms as they display more conservative behavior in linguistically stable settings (Labov 2001:266). This principle of gendered language use is, nonetheless, compatible with the increase in the use of other variants of LIKE: the real-time change shows that the stigma associated with clause-final LIKE

waned when other uses of LIKE became increasingly associated with AmE during the latter half of the twentieth century. It thus seems plausible that the change of attitudes associated with LIKE paved the way for its rapid and ongoing diffusion through the IrE speech community. In other words, the shift with respect to the connotations associated with clause-final LIKE is caused by the increase in the use of other LIKE variants carrying quite distinct associations.

While clause-final uses of LIKE which are still pervasive, particularly among older males were once regarded as a substandard feature marking "Irishness", they are no longer today. The use of linguistic variants to mark local identity is one of the cornerstones of sociolinguistics. The concept of local identity as a driving force of ongoing change was, for instance, explored in detail in Labov's (1963) study of linguistic behavior on Martha's Vineyard. Nowadays, the association of vernacular LIKE with local identity has waned. Indeed, younger speakers of IrE quite possibly perceive LIKE to be a modern US American element rather than a marker of national identity. The change of LIKE from a marker of nationality to a social identity marker conforms to a process described by D'Arcy (2007:397):

> It is possible, for example, that vernacular uses of *like* were recycled as a Valley Girl phenomenon once their initial association with the counterculture groups waned among subsequent generations of speakers.

The increase in innovative American uses of LIKE (Andersen 2001:221) among adolescents might not only have led to a weakening of the status of clause-final LIKE as a typically Irish particle, but it might also have triggered the rise of clause-medial LIKEas an identity marker among female IrE speakers. It is likely that this increase in the use of clause-medial LIKE among female adolescents subsequently caused the decrease in the gender difference in overall LIKE use. As a consequence, LIKE has been losing its status as a marker of both rural-ness and male-ness.

The driving force behind the increase in cause-medial LIKE use was, however, not primarily its functional versatility, but its association with a socio-cultural reference group (Labov 2010:368). Therefore, it was not primarily the discourse-pragmatic functionality of LIKE, but its status as a fashionable prestige variant which lead to its spread in IrE. This argument rests on Labov's analysis of driving forces underlying linguistic change (Labov 2007, 2010). As a matter of fact, Labov concludes that it is precisely the attachment of variants with socio-cultural associations which trigger and drive linguistic change:

> In one form or another, they [continuation, acceleration or completion of change] involve the association of social attributes with the more advanced forms of a change in progress: local identity, membership in communities of practice, social class, age or gender. (Labov 2010:368)

This tentative analysis is corroborated by the dramatic increase in clause-medial LIKE use in IrE over the past two decades: while only 28.2 of speakers used this from in the data compiled between 1990 and 1994, 41 percent of speakers used it in data compiled only about ten years later. The growing number of speakers using LIKE is accompanied by a dramatic increase in frequency: staring with 0.82 instances per 1,000 words in the early 1990s its frequency rose to 1.4 instances in data compiled between 2001 and 2005. In addition, this trend is not limited to clause-medial LIKE, but affects all variants of LIKE. The dramatic change in real-time is, nonetheless, restricted to younger cohorts. In fact, the most notable increase is observable among speakers between the ages of 26 and 33, whose frequencies of clause-medial and clause-final LIKE have increased very significantly. In addition, among speakers younger than 25 years of age, all variants of LIKE – except for clause-final LIKE – have notably gained ground. The effects of the date of data compilation are particularly noteworthy, as this real-time analysis covers merely a brief, ten- to fifteen-year period.

Moreover, clause-medial LIKEoffers an additional, unexpected insight: while the use of this variant of LIKE peaks in the speech of females in their late twenties and early thirties, the rates among males follow a slightly different trajectory. Among males, clause-medial LIKEis used at similar rates by all speakers below the age of 50. Above 50, the frequency of clause-medial LIKEplummets dramatically, irrespective of gender. In fact, it is almost non-existent in the speech of the older population.

The different rates for males and females suggest an initial male lead. Among speakers who are in their mid-thirties to approximately the age of 50, male speakers clearly outperform females in terms of LIKE use. It thus seems that males have adopted this feature earlier than females – probably due to its similarity to the clause-final variant. Based on the apparent-time trajectories we can infer that this initial male lead waned rather quickly. Indeed, the direction of the gender lead is already reversed among speakers in their late twenties to mid-thirties. This reversal corroborates the assertion that females have adopted clause-medial LIKElater than males. However, once they had implemented it into their grammar, they used it more frequently and with more versatility than their male peers. Once this feature became increasingly associated with female speech, it decreased steadily among males.

In contrast to what would be expected, clause-medial LIKE, as with clause-initial LIKE, are not most frequent among the youngest female cohort, but rather women aged 26 to 33. Hence, the assumption that LIKE use is typically and uniformly a feature of adolescents does not apply to the Irish setting. Indeed, the age stratification among females is suggestive as it may relate to the beginning of the rapid spread of LIKE which is still affecting the IrE population today. The elevated frequencies show the point in time when clause-initial and clause-medial LIKE became increasingly popular among adolescents. Taking the apparent-time distribution as a base, the peak in use among females between the ages 26 and 33 indicates that both variants of LIKE started to gain ground among adolescents somewhere between the mid-eighties and mid-nineties. In this view, it is the female speakers, who were

adolescents at the time, who were the leaders of this change. The incrementation of LIKE in IrE required that the females, who were the initial leaders, retained the use of LIKE as a functional element in their individual grammar. As these females grew older, they continued to use LIKE, which to them is a fully functional element as a result of its pragmatic versatility. As they matured, they introduced this vernacular feature into formerly constrained registers and to yet older speakers. This means that, once these former leaders of the change had outgrown adolescence, it became possible for older cohorts to adopt this feature, since only then did it cease to be perceived as a marker of youth. Adding weight to the hypothesis that the increase in the use of clause-medial LIKEis a relatively recent phenomenon is the fact that it is the only variant of LIKE which is affected by region. The multifactorial analysis confirmed that it is distributed very unevenly between the southern and northern Ireland. Although this scenario is hypothetical, it is based on and compatible with the present data, as well as the timeline of previous accounts of LIKE use (e.g. Andersen 2001; D'Arcy 2005, 2007, 2008), and well documented mechanisms of diffusion and incrementation (Labov 2001, 2010).

7.5 LIKE in Indian English[65]

With the exception of IrE and SctE, the use of LIKE in non-standard and learner varieties of English has mostly escaped scholarly attention. This is unfortunate, as non-standard varieties, in particular, offer opportunities for testing the cross-cultural stability of mechanisms of language change. However, the systematic analysis of mechanisms underlying the transmission, diffusion, and incrementation of innovative linguistic variants has so far focused mostly on standard varieties of English, i.e. EngE and AmE. The present approach, therefore, offers new insights about the interplay of socio-cultural norms and mechanisms of language change.

While the Labovian strand of sociolinguistics assumes that mechanisms of change are rather stable and, hence, somewhat independent of cultural practice (cf. Labov 2001:280 and the sources cited therein), studies focusing particularly on dialect contact, multilingualism and creoles which challenge this assertion (cf. Thomason & Kaufman 1991). Thomason and Kaufman, for instance, point out that "they [language-internal factors for change] are easily overridden when social factors push in another direction" (1991:4). Labov, on the other hand, finds that the effect of socio-cultural variation is well within the bounds of what was expected by non-cultural factors alone. One case in point is the hypothesis that the

> curvilinear pattern of a sound change from an originating group to a neighboring group may then be the simple product of frequencies of interaction. *The account based on covert attitudes is redundant* [italics M.S.] to the extent that the network of daily interaction brings people into contact with the new forms in proportion to their distance from the originating group. (Labov 2001:192).

65 For a complete overview of the final data set of IndE used in this study cf. Table 142 in the Appendix.

The validity of the assumption that variation is to a large extent independent from culture-specific practices and attitudes is explored in the section to come.

Previous research on IndE reported surprising results with respect to the use of the discourse marker LIKE. Siemund, Maier & Schweinberger (2009), for example, found that the positional distribution of LIKE is remarkably similar to the IrE distribution. According to Siemund, Maier & Schweinberger (2009), IndE speakers strongly prefer LIKE in clause-marginal position and, interestingly, show elevated LIKE use in clause-final position. Another study explicitly treating LIKE in IndE has been conducted by Valentine (1991). With respect to LIKE, Valentine notes that "[a]lthough *like* does not appear as extensively nor perform as many functions as in AmE, […] instances of this element occur frequently and freely in the Indian English discourse" (1991:332). On a functional level Valentine (1991) notes that LIKE conveys a "sense of superficiality, […] functions as a place filler, and […] introduces an example" (1991:332). Valentine concludes that the relation and emergence of LIKE in IndE has not yet been extensively studied and that future research with a focus on comparisons with other regional varieties of English would greatly benefit our understanding of culture-dependent, variety-specific use of pragmatic markers (Valentine 1991:333-334). In addition, Valentine (1991:333) stresses that the

> examination of discourse markers in Indian English represents a part of a growing interest in the pragmatics of the Englishes of the world-an area important for gaining insights into discourse strategies and cultural expectations in the world Englishes. […] I found that spoken Indian English exhibits typical features of established English discourse, but the interaction is clearly Indian, not only in terms of its syntactic, lexical, and stylistic features, but in terms of its pragmatics: discourse markers. Moreover, the cultural identity of the speaker, her choice of topic, her nature of development, and her tone and style of talk contribute to Indianizing the English text. That these Indian English discourse are identified by such

features is only a partial attempt toward a true description of the forms and functions of the Indian variety of English.

Before analyzing the IndE data in detail, the basic statistics of the data are reviewed to get a first impression of its structure. Of the 236 speakers present in the data, 107 make use of LIKE at least once. In other words, nearly half of IndE speakers used this non-standard form. This quantity is rather large for an L2 variety and suggests that LIKE is indeed a common pragmatic marker in IndE, as suggested by Valentine (1991). The high proportion of LIKE users is particularly noteworthy, since vernacular LIKE has hardly been studied in this regional variety.

Table 79: Overview - LIKE use in IndE

	N (total)	**Mean (per 1,000 words)**	**Median (per 1,000 words)**	**Standard deviation (s)**
LIKE	331	1.509	0.000	2.649

Table 79 reports that LIKE's mean use is 1.509 times per 1,000 words. The difference between mean and median indicates that LIKE use is not distributed homogenously among all IndE speakers: the majority of speakers have not used this pragmatic marker, while a substantial subsection of IndE speakers has used it rather frequently.

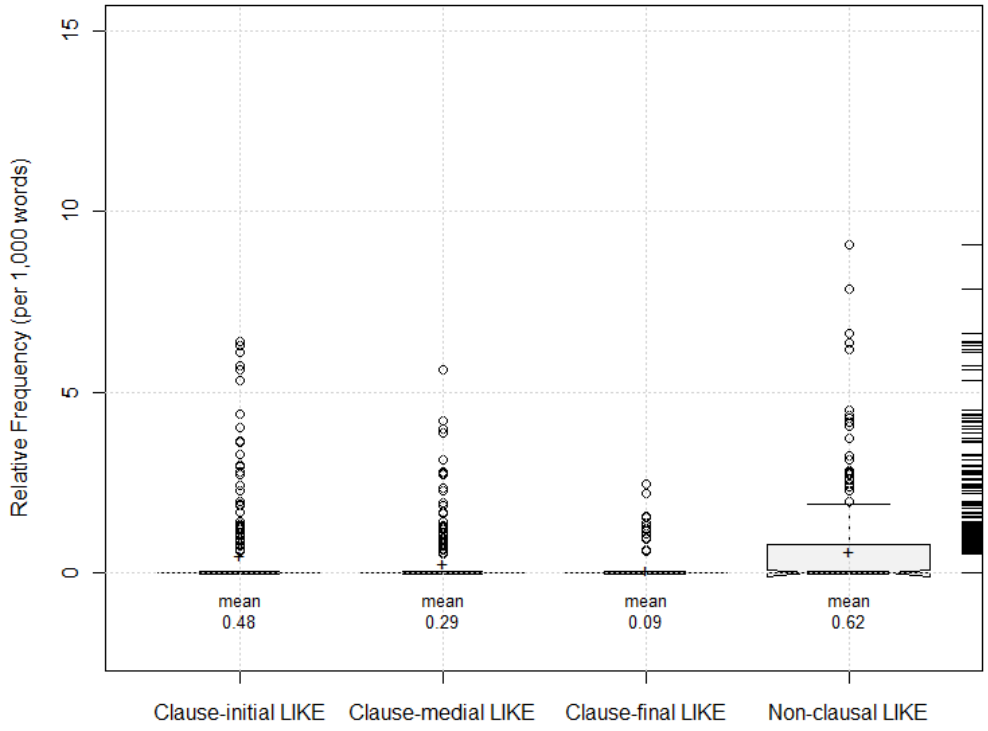

Figure 45: Rate of LIKE variants in IndE

Figure 45 shows that the positional distribution of LIKE in IndE is quite distinct from the profiles of other regional varieties. Thus, the distinctness of the Indian patterns illustrates the variety-specific diversity in LIKE use across the English-speaking world. Speakers of IndE strongly prefer non-clausal LIKE, while other uses of LIKE are negligible. Hence, LIKE in IndE is commonly used as a floor-holding device, to buy processing time, or to introduce phrasal specifications. Given this idiosyncracy, the age and gender distribution is especially intriguing. It is not implausible to hypothesize that in addition to the variety-specific usage pattern, IndE may possess specific mechanisms of diffusion and incrementation, which would surface in an unusual age and gender distribution.

Table 80: Age and gender distribution of LIKE in IndE[66]

			MALE		FEMALE			
AGE	**Speakers**	**Words**	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (18 – 25)	66	56,824	65	2.63	135	3.26	200	3.09
2 (26 – 33)	56	52,140	13	0.76	45	1.77	58	1.32
3 (34 – 41)	42	40,485	13	0.51	22	1.06	35	0.77
4 (42+)	64	57,840	24	0.53	15	0.81	35	0.60
NA	8	4,357	--	--	--	--	3	0.94
SUM	236	211,646	117	0.92	275	1.5	331	1.51

In contrast to the anomalous distribution of functional variants, the age and gender distribution is not particularly noteworthy: Table 80 reports an increase in LIKE in apparent-time and a minute female lead across all age groups. A graphical display may, however, provide a more refined understanding of this distribution (cf. Figure 46). While the graphs in Figure 46 corroborate a change in apparent-time, they strongly suggest that the gender difference is insignificant, as the 95 percent confidence intervals overlap.

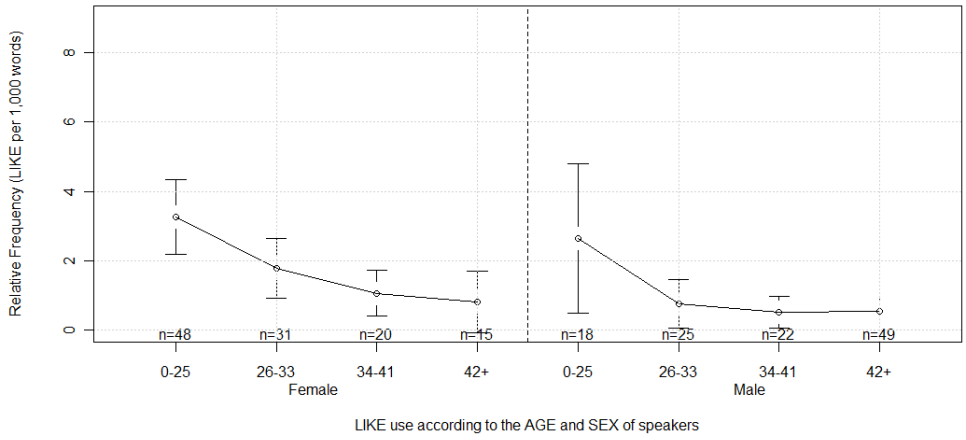

Figure 46: Age and gender distribution of LIKE in IndE

Since Figure 46 indicates that gender is insignificant, it is appropriate to collapse the genders into a single cohort for each age group. The increase in

66 Speakers for whom age and gender information is not available are classified as NA.

speakers within each age cohort enables a more accurate display of the apparent-time distribution (cf. Figure 47).

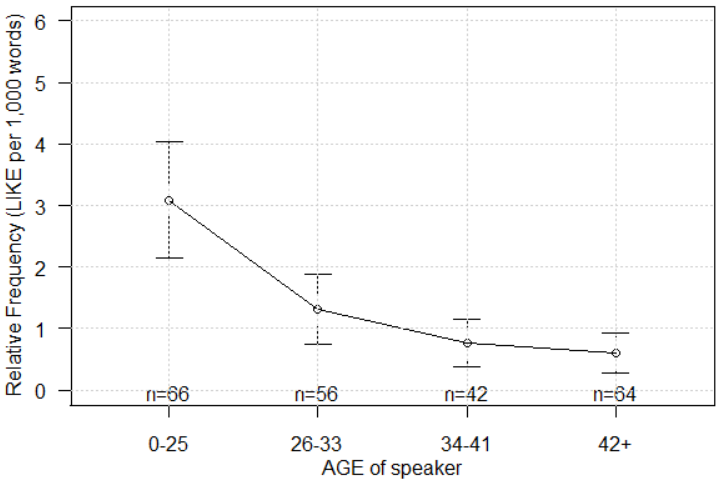

Figure 47: Age distribution of LIKE in IndE

Collapsing the genders has indeed improved the display, which confirms the inference that LIKE use in IndE is significantly stratified across age groups. In fact, the distribution is highly suggestive of change in apparent-time, as it indexes age grading. However, this distribution could also be the result of generational change. Hence, the apparent-time analysis requires real-time confirmation in order to differentiate the former from the latter type of change.

7.5.1 Statistical analysis of LIKE in Indian English

The following statistical analysis evaluates the inferences derived from the summaries and graphical displays. The independent variables included in the multivariate analyses are age (A1 = age 18 to25, A2 = age 26 to 33, A3 = age 34 to 41, A4 = 42 years and older), gender, first language, number of participants, and the current occupation of speakers. As before, the initial saturated model is fitted in a step-wise procedure. The final minimal adequate model contains

only predictors with significant effect sizes. This procedure is applied to all functionally distinct variants of LIKE.

Table 81: Results of the multivariate regression for LIKE in IndE

ALL	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	1.22	0.149	8.15	<0.001***
SEX: Male	-0.42	0.274	-1.53	.12
A2	-0.81	0.269	-3.03	<0.01**
A3	-1.26	0.293	-4.31	<0.001***
A4	-1.41	0.346	-4.07	<0.001***

The regression model predicting the overall frequency of LIKE per 1,000 words in IndE based on the age and sex of speakers performs significantly better than a base-line model (χ^2= 47.223, df=4, p<.001***). The sequence of increasing negative coefficients corroborates the impression derived from Figure 46, according to which LIKE use recesses with age. In addition, the steady increase in age coefficients suggests that this negative correlation between age and LIKE use is near-linear and it indicates the existence of an implicational hierarchy of the form A1>A2>A3>A4, which is validated by additional t-tests (cf. Table 82).

Table 82: T-Test results for LIKE use with respect to AGE

	A1-A2	A1-A3	A1-A4
ALL	3.296***	4.480***	4.914***

Furthermore, the model output reports a substantial gender difference. The negative coefficient implies that males use LIKE less often than females. Although the effect size of gender is insignificant, it remained in the model as the coefficient is substantial, with a value exceeding .3. That gender does indeed have a significant effect is validated by non-parametric t-tests (cf. Table 83).

Table 83: T-Test results for LIKE use with respect to AGE and GENDER

	A1	A2	A3	A4
ALL	0.555 (n.s.)	1.901* (<.05)	1.077 (n.s.)	0.566 (n.s.)

The t-tests detect a significant gender difference among speakers aged between 26 and 33. However, the rather low t-value suggests that the effect of gender is moderate.

Since the effect size of the present occupation of speakers neither reach the necessary level of significance nor returned a substantial coefficient, it was excluded from the multivariate analysis. Accordingly, the regression output fails to confirm that LIKE use is socially stratified, although the gender difference might reflect the existence of a female-dominated change and thus systematic social heterogeneity.

7.5.1.1 Clause-initial LIKE

We now turn to clause-initial uses of LIKE as in (96), which establishes textual coherence by linking a preceding element to a subsequent part of discourse. In addition, clause-initial LIKE may introduce specifications which may either take the form of explanations, or a list of examples. Although Valentine (1991) asserts that LIKE in IndE does not perform the whole range of functions it fulfills in other regional varieties, she explicitly notes that LIKE in IndE frequently introduces examples (1991:332).

(96) a. *Like* once you are intro <,> introduced to that system <,,> so it becomes easy for you to grasp. (ICE India:S1A-073$A)

b. *Like* we leave this tourist <,> only if you return that uh <,> terrorist <,> which you've uh arrested. (ICE India:S1A-054$A)

c. *Like* they can literally predict your every move. (ICE India:S1A-056$C)

The examples in (96) indicate that the use of this variant does not functionally differ from its use in the inner circle varieties discussed so far. All instances introduce specifications or explanations of previously stated propositions or parts thereof.

The following subsection will focus on the gender and age distribution of this variant to allow for a more detailed analysis of its use.

Table 84: Age and gender distribution of clause-initial LIKE in IndE

AGE	MALE		FEMALE			
	LIKE (N)	LIKE (mean)	LIKE (N)	LIKE (mean)	SUM (N)	ALL (mean)
1 (0–25)	24	0.99	44	1.02	68	1.01
2 (26–33)	8	0.54	13	0.47	21	0.50
3 (34-49)	1	0.05	8	0.37	9	0.20
4 (50+)	5	0.10	2	0.11	7	0.10
SUM	38	0.33	67	0.63	105	0.48

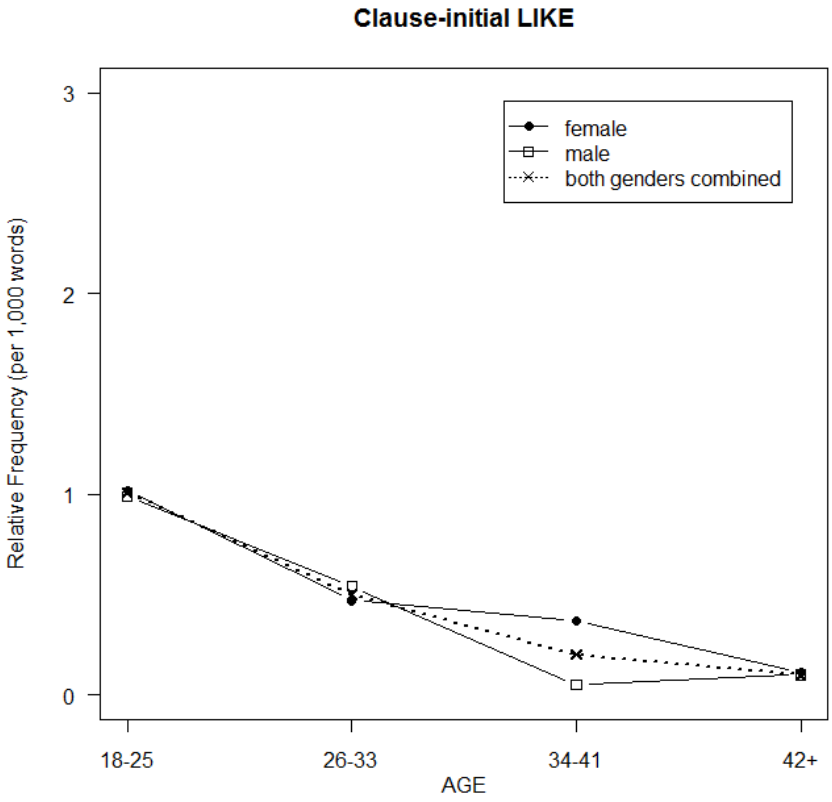

Figure 48: Age distribution of clause-initial LIKE in IndE

Figure 48 suggests age stratification and the absence of gender differences, although the slight female lead among speakers aged 34 to 41 might be interpreted as implying a female lead during the incipient phase of LIKE's introduction to IndE. In other words, females might have adopted this form prior to males, who seem to have lagged behind about ten to fifteen years but

caught up rather quickly, as the rates are almost identical among speakers aged 26 to 33.

Table 85: Results of the multivariate regression for clause-initial LIKE in IndE

INI	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.95	0.272	-3.50	<.001***
SEX: Male	0.22	0.335	0.67	.50
A2	-0.31	0.405	-0.78	.43
A3	-0.38	0.528	-0.72	.46
A4	-1.55	0.477	-3.26	<.01**
PAI	0.08	0.010	8.24	<0.001***
SEX: Male*A3	-2.23	1.134	-1.97	<.05*

The multifactorial analysis reports a near-linear negative correlation between the frequency of clause-initial LIKE and the age of speakers. In fact, the steadily increasing negative values of the age coefficients corroborate the existence of an implicational hierarchy of the form A1>A2>A3>A4.

The absence of gender differences is confirmed by non-parametric t-tests (cf. Table 86).

Table 86: Clause-initial LIKE with respect to AGE and GENDER

	A1	A2	A3	A4
INI	-0.064 (n.s.)	-0,195 (n.s.)	1.875* (<.05)	0.310 (n.s.)

The t-tests detect marginally significant gender difference among speakers aged 34 to 41. The low t-value suggests, nonetheless, that the effect of gender is very moderate. In addition, the t-statistics confirm the systematic effect of age and corroborate an implicational hierarchy of the form A1>A2>A3>A4.

Table 87: Clause-initial LIKE with respect to AGE

	A1-A2	A1-A3	A1-A4
INI	2.017* (<.05)	3.468*** (<.001)	4.196***(<.001)

The PAI index is a significant predictor in the regression model but it has a very small effect size. In contrast to previous models, the value of the PAI index here is notably higher. This elevated value might be taken to mean that the use of clause-initial LIKE clusters more heavily in certain conversations than it did

in other varieties. Though the gender of speakers is not significant in itself, it is part of a significant interaction. This interaction confirms that among speakers between the ages of 34 and 41, males are significantly less likely to use this type of LIKE than females, corroborating the impression derived from Figure 48.

7.5.1.2 Clause-medial LIKE

The following section is dedicated to the use of clause-medial LIKE, as in examples (97) to (100). With respect to this variant of LIKE, Valentine (1991:332) reports that it is rather limited with respect to its textual functions and suggests that clause-medial uses of LIKE function primarily as "place fillers". Examples drawn from the Indian ICE component challenge these hypotheses. For instance, the uses of LIKE in (97) to (99) focus the following element while the instances in (100) are best described as hedges. In other words, clause-medial LIKEis functionally versatile on a discourse-pragmatic level.

(97) a. But then nobody *like* motivated me. (ICE India:S1A-021$B)

b. And to purchase a ticket you have to *like* fight with persons just to get a ticket. (ICE India:S1A-061$B)

(98) c. Hinglish <,> once I came across *like* Hinjali <,> isn't it. (ICE India:S1A-028$B)

d. And at the same <,> no that's what she told when uh <,> a question was raised *like* this last time. (ICE India:S1A-064$B)

(99) e. But I never found Gulbarga *like* <,,> very appealing to me. (ICE India:S1A-061$B)

f. [Y]ou get disturb *like* emotionally. (ICE India:S1A-069B)

(100) g. He is also using *like* something about his political powers and all that Chandraswami then again. (ICE India:S1A-012$B)

h. Friends are *like* <,,> more closer than parents in hostel. (ICE India:S1A-054$A)

In addition to being pragmatically versatile, the use of clause-medial LIKE appears to be systematic with respect to sentence structure. Despite being regarded as syntactically unconstrained (e.g. Siegel 2002), it is confined to fixed syntactic slots. Adding weight to this view is the fact that the instances of LIKE

in (97) have scope of the following VP, those in (98) over a following NP, and those in (99) over an AP. In all cases, LIKE precedes the lexically heavy elements and occurs exclusively at the boundaries of phrases, or right after prepositions in PPs. These syntactic environments exactly match the well circumscribed contexts of clause-medial LIKEin CanE (cf. D'Arcy 2005, 2007, 2008; Tagliamonte 2005). Therefore, focus LIKE in IndE is clearly allowed in quite distinct syntactic environments, which points to the fact that it is syntactically bound.

We will now turn to the sociolinguistic profile of clause-medial LIKE.

Table 88: Age and gender distribution of clause-medial LIKE in IndE

	MALE		**FEMALE**			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (0–25)	10	0.54	24	0.55	34	0.55
2 (26–33)	2	0.06	9	0.29	11	0.19
3 (34-49)	4	0.15	5	0.23	9	0.19
4 (50+)	4	0.06	5	0.36	9	0.13
SUM	20	0.15	43	0.40	63	0.28

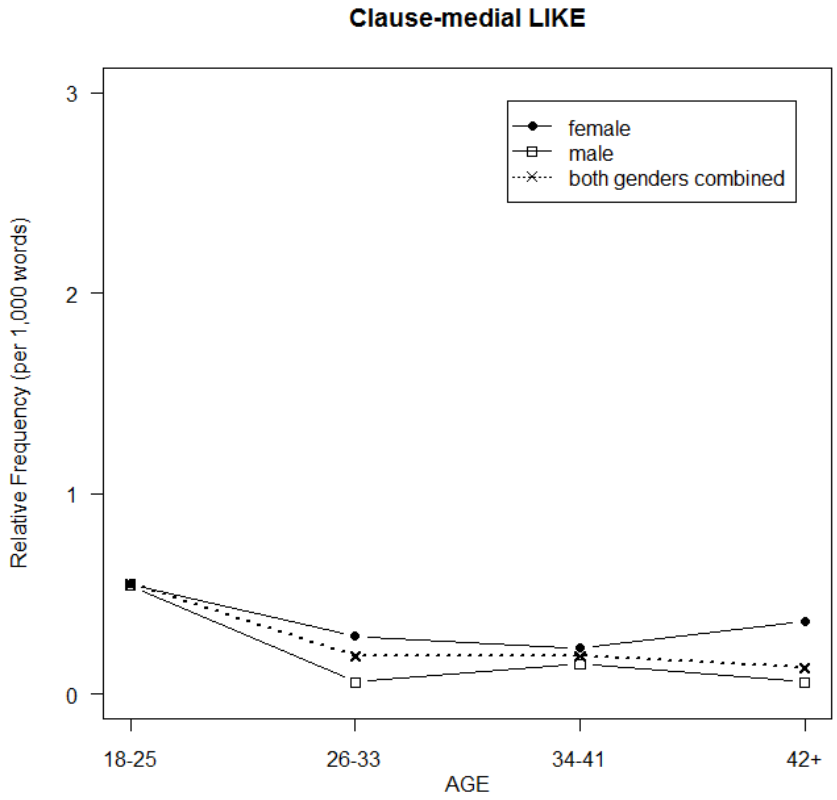

Figure 49: Age distribution of clause-medial LIKE in IndE

Surprisingly, Figure 49 neither indicates substantial age stratification nor a sizable gender difference. It does, however, suggest that clause-medial LIKE is used rather infrequently but homogenously across cohorts. Nonetheless, young adults appear to make more use of LIKE than do older speakers – although this tendency is rather tentative, as it is not as pronounced as in other displays. In addition, Figure 49 suggests a minor though consistent gender bias, as females use clause-medial LIKE on average more than males in all but the youngest age cohort.

Table 89: Results of the multivariate regression for clause-medial LIKE in IndE

MED	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.46	0.264	-1.77	<.1
SEX: Male	-.63	0.422	-1.49	.13
A2	-0.83	0.457	-1.82	<.1
A3	-0.68	0.441	-1.54	.12
A4	-1.04	0.650	-1.61	.10

The output of the multivariate design does not report social stratification. In fact, the analysis shows that none of the coefficients of the variables reaches the five percent level of significance. This implies that age fails to be statistically significant. The order of age coefficients corroborates a change in apparent-time. Although the age coefficients indicate that speakers up to the age of 25 years use clause-medial LIKE more than speakers older than 25, none of these coefficients exceeds the level of marginal significance. Beyond age, the gender of speakers seems to have an effect on LIKE use, but as in the case of age, the coefficient does not exceed the .05 level of significance. If we over-interpret the results, we might say that males and older speakers use clause-medial LIKE less often than females and younger speakers, but these inferences are tentative, not conclusive.

We validate the absence of gender and age differences by performing additional non-parametric t-tests.

Table 90: Clause-medial LIKE with respect to AGE and GENDER

	A1	A2	A3	A4
MED	0.215 (n.s.)	1.670 (<.1)	1.426 (<.1)	0.922 (<.1)

The t-tests detect marginally significant gender differences among speakers aged between 26 and 41. The rather low t-values imply, nonetheless, that the effect of gender – if it exists – is moderate at best.

Table 91: Clause-medial LIKE with respect to AGE

	A1-A2	A1-A3	A1-A4
MED	2.037*	1.875*	2.559**

The t-statistics for age confirm a significant difference between age group 1 and all other age groups. Thus, the t-tests lead to an improved understanding of the relation between age and LIKE use: A1>A2; A1>A3; A1>A4; A1>A2>A4.

In summary, the only claims one can assert with confidence based on the present analysis are that young speakers use clause-medial LIKE more than older speakers and that among speakers who are older than 26 years of age, females use clause-medial LIKE more than their male peers.

7.5.1.3 Clause-final LIKE

Clause-final LIKE as in (101) is comparatively frequent in IndE (cf Siemund, Maier & Schweinberger 2009). So far, however, there is no adequate explanation for this, as both substrate influence and fossilization are neither sufficient nor satisfactory to explain this phenomenon. Explanations favoring fossilization as an explanation for the elevated frequencies of clause-marginal uses of LIKE reported by Siemund, Maier & Schweinberger (2009) are problematic, because they fail to account for the low overall frequency of LIKE as well as the marginal quantity of clause-initial instances. Substrate influence, on the other hand, is unsatisfactory as it would predict regional clustering of LIKE use as a result of systematic correlations with substrate languages. However, the present analysis fails to confirm both regional clusters and significant interactions with L1s other than English.

(101) a. I am not trying for any exam *like*. (ICE India:S1A-039$B)
b. Ah <,> I thought you have bought some ordinary thing *like*. (ICE India:S1A-040$A)
c. Uhn this is the best time I tease him *like* you know (ICE Ireland:S1A-041$A)

Before going into more detail, the use of clause-final LIKE among speakers of IndE is shown in Table 92 to provide a preliminary understanding of its sociolinguistic distribution.

Table 92: Age and gender distribution of clause-final LIKE in IndE

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (0–25)	2	0.05	11	0.26	13	0.20
2 (26–33)	0	0.00	2	0.07	2	0.03
3 (34-49)	4	0.16	1	0.03	5	0.10
4 (50+)	1	0.02	0	0.00	1	0.01
SUM	7	0.05	14	0.13	21	0.09

Table 92 Indicates a moderate recess of clause-final LIKE among female speakers, while the overall distribution and the male rates do not follow a clearly visible and coherent trend. Only speakers younger than 25 seem to use this type of LIKE more than older speakers, although this tendency is tentative due to the relatively small quantity of instances.

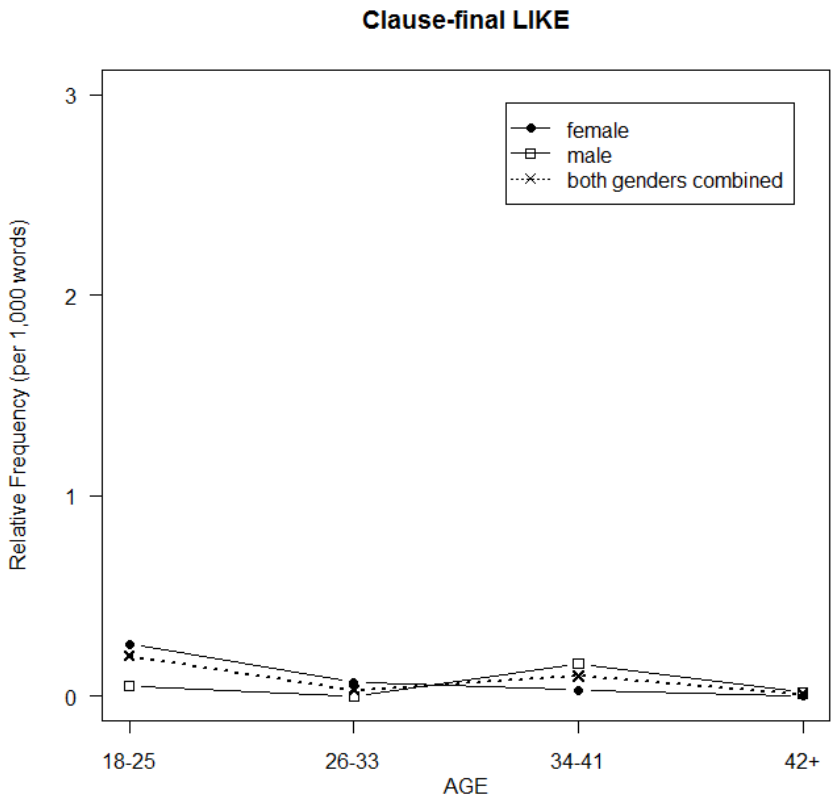

Figure 50: Age distribution of clause-final LIKE in IndE

Figure 50 indicates that the use of clause-final LIKE is neither gendered nor notably present among, for example, adolescents. Similar to clause-medial LIKE, it seems to be distributed rather homogenously among IndE speakers, as no distinct trend emerges.

Table 93: Results of the multivariate regression for clause-final LIKE in IndE

FIN	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.49	0.482	-5.163	<0.001***
SEX: Male	-1.23	0.455	-2.71	<.01**
A2	-1.11	0.772	-1.43	.15
A3	-1.01	1.055	-0.96	.33
A4	-1.05	0.960	-1.09	.27
PAI	0.09	1.390	1.71	<.001***
SEX: Male*A3	2.43	1.245	1.95	<.10

The regression model reports change in apparent-time, i.e. age-grading, indicated by the negative coefficients of the respective age groups. In addition, the analysis detects a gender difference. The regression output confirms that male speakers use clause-final LIKE less often than female speakers. The strongest predictor for clause-final LIKE in the present model is an interaction between the gender and the age of speakers. This interaction implies that when clause-final LIKE is used by males, it is used almost exclusively by males aged 34 to 41.

In addition, the multivariate statistics report a highly significant effect of the PAI index, which indicates that clause-final LIKE clusters in certain conversations. Considering the combined effects of age, gender and the PAI index, the results indicate that the use of clause-final LIKE is most common among younger women and middle-aged males. This finding is validated by non-parametric t-tests.

Table 94: Clause-final LIKE with respect to AGE and GENDER

	A1	A2	A3	A4
FIN	2.158* (<.05)	1.438 (<.1)	-0.774 (n.s.)	-1.000 (n.s.)

The t-tests detect a significant gender difference among speakers of the youngest age group and a marginally significant effect of gender among speakers aged between 26 and 33. The rather small t-values imply that the effect of gender is very moderate. This result is unexpected, since the interaction between gender and age reported by the regression model did not exceed the five percent level of significance in the more robust t-test. This suggests that the regression model over-estimated the gender effect as a result of the difference between the expected value for males of this group and their actual performance. In other words, it is not gender which is significant, but the fact that males aged 34 to 41 use clause-final LIKE more any other male cohort.

Table 95: Clause-final LIKE with respect to AGE

	A1-A2	**A1-A3**	**A1-A4**
FIN	2.602** (<.01)	1.374 (<.1)	3.026** (<.01)

With respect to age, the t-tests confirm a significant difference between age group 1 and all other age groups. The fact that the t-values do not show a steady increase from young to older indicates that the use of clause-final LIKE is not systematically stratified and, therefore, does not comply with an implicational hierarchy.

In summary, based on the present analysis, it can concluded that young IndE speakers exhibit significantly higher rates of clause-final LIKE than older speakers and that among young speakers, it is females who use clause-final LIKE most.

7.5.1.4 Non-clausal LIKE

In IndE, non-clausal LIKE as in (102) is by far the most frequent form of LIKE. This over- proportionate rate of syntactically unbound LIKE validates the claim that LIKE use in IndE is rather distinct compared to other regional varieties (cf. Siemund, Maier & Schweinberger 2009). In addition, the high rate of non-clausal LIKE substantiates Valentine's (1991) assertions that LIKE in IndE is used predominantly as a "place filler" or processing time buyer. The majority

of examples drawn from the Indian component of the ICE are indeed congruent with this assertion, as they frequently seem to function as floor-holders and indicators of planning difficulty as in (102).

(102) a. Yours is *like* <,>… (ICE India:S1A-073$B)

b. [W]e have got seven different uh <„> *like* uh <,> departments. (ICE India:S1A-090$B)

c. In human physiology we take the samples of the blood and then we test the blood and then we count <,> R B C count <„> then W B C count then <„> *like* uhm <„> after that we see <„> the hemoglobin. (ICE India:S1A-090$B)

The sociolinguistic distribution of this form is particularly interesting because of its frequency. The analyses of the other variants of LIKE in IndE did not confirm that they were particularly stratified. This null result may have been, however, at least partly caused by their low frequency. In contrast, non-clausal LIKE exceeds all other variants in terms of frequency and may, therefore, offer more intriguing and more reliable insights.

Table 96: Age and gender distribution of non-clausal LIKE in IndE

	MALE		**FEMALE**			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (0–25)	27	0.93	53	1.37	80	1.25
2 (26–33)	3	0.15	20	0.89	23	0.56
3 (34-49)	3	0.09	8	0.41	11	0.25
4 (50+)	14	0.33	4	0.33	18	0.33
SUM	47	0.34	85	0.93	132	0.64

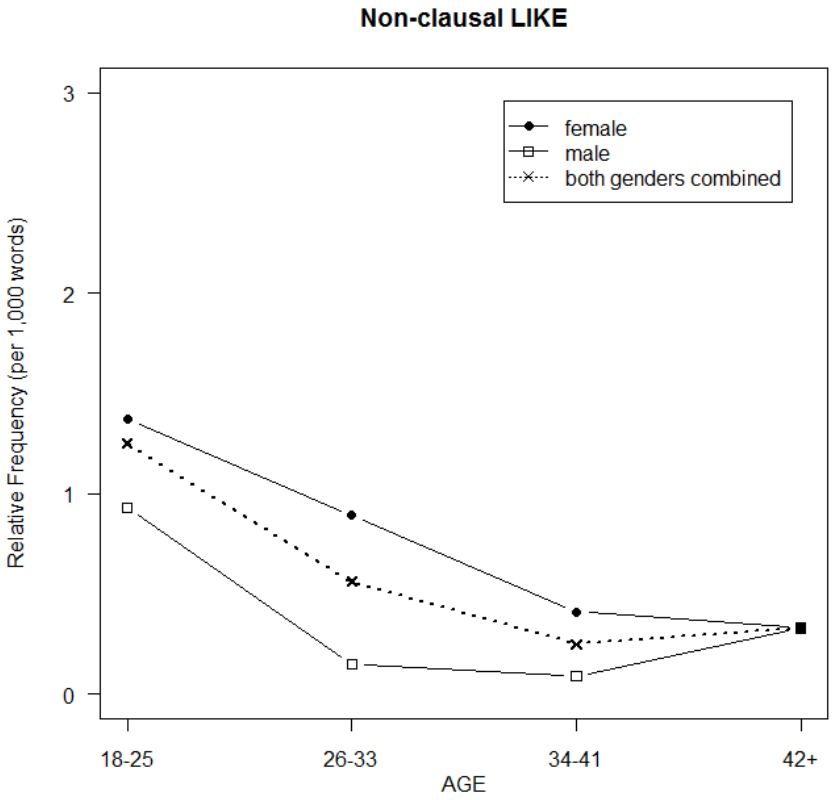

Figure 51: Age distribution of non-clausal LIKE in IndE

According to Figure 51, the use of non-clausal LIKE follows the common trajectory for LIKE use: The distribution of non-clausal LIKE confirms age grading in apparent-time and thus age-grading. The age distribution is near-linear among females, but clearly not among males. Indeed, males seem to lag behind about one generation. The fact that only the youngest males use LIKE could suggest that these males have acquired its use from their mothers, which not only explains the lag, but also conforms to Labov's (2001) argument concerning lower rates of males in female-dominated changes in general.

In addition, Figure 51 indicates that the use of non-clausal LIKE in IndE is highly gender-sensitive, as females use LIKE notably more than men in all but the youngest age group.

Table 97: Results of the multivariate regression for non-clausal LIKE in IndE

NON	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.79	0.216	-3.65	<.001***
SEX: Male	-0.46	0.228	-2.03	<.05*
A2	-0.21	0.280	-0.75	.45
A3	-0.76	0.350	2.18	<.05*
A4	-0.09	0.411	-0.23	.81
PAI	0.09	0.009	10.16	<.001***

The regression output confirms both the gender difference and age-grading. Among males, it not the oldest speakers who use it least, but the middle-aged ones. The pattern emerging from Figure 51 and the regression model indicates that LIKE use is extremely gendered in IndE. Again, non-parametric t-tests are applied to confirm these findings.

Table 98: Non-clausal LIKE with respect to AGE and GENDER

	A1	A2	A3	A4
NON	0.741(n.s.)	2.394*(<.05)	1.554 (<.1)	-0.052 (n.s.)

The t-tests detect a significant gender difference among speakers aged 26 to 33 and a marginally significant effect of gender among speakers between the ages of 34 and 41. However, gender has neither a significant effect among very young nor among speakers older than 42.

Table 99: Non-clausal LIKE with respect to AGE

	A1-A2	A1-A3	A1-A4
NON	2.412** (<.01)	3.906*** (<.001)	3,361*** (<.001)

In addition, the t-tests confirm that the youngest speakers in the data use non-clausal LIKE significantly more than older speakers, which confirms the regression analysis. However, speakers older than 26 years of age do not substantially differ in their use of this type of LIKE (cf. Table 100).

Table 100: Non-clausal LIKE with respect to AGE

	A2-A3	A2-A4	A3-A4
NON	1.575 (n.s.)	1.009 (n.s.)	-0.539 (n.s.)

The fact that none of the older groups differ significantly implies that with respect to LIKE use, IndE speakers can be divided into frequent LIKE users, which encompasses younger speakers and moderate LIKE users, which includes all speakers over the age of 26.

7.5.2 Evaluation of the apparent-time construct

The following section aims to unearth differences in the data correlated with the date of data compilation. In other words, the apparent-time hypothesis is evaluated by means of a real-time analysis. In contrast to the real-time analysis performed above, this analysis is confined to an even shorter time period, as the Indian data were compiled before 2001. The short distance between the two points in time that are compared renders a real-time analysis problematic and reduces the probability of finding actual differences.

As predicted, the regression models for all types of LIKE returned insignificant results and failed to outperform a model based solely on the intercept. It follows that the present real-time analysis fails to substantiate ongoing real-time change with respect to LIKE use. The validity of this null-result is confirmed by both χ^2-tests and non-parametric t-tests. However, to conclude that the real-time analysis confirms the absence of ongoing change would be premature, given the brief period covered by the data.

In fact, the age differences and the fact that each variant is gender sensitive is highly suggestive of ongoing real-time change in an early or mid-range stage – despite the real-time results.

7.5.3 Summary: LIKE use in Indian English

The use of LIKE in IndE, i.e. a second-language variety, differs notably from the use of LIKE in other varieties of English and provided rather unusual though interesting results. The Indian setting in particular supports interpretations stating that globally available resources undergo re-interpretation and

adaptation (e.g. Buchstaller 2008; Buchstaller & D'Arcy 2009; Kachru 1992; Meyerhoff & Niedzielski 2003). The support for this behavior of linguistic innovations stems from the remarkably high rate of non-clausal LIKE: while clause-medial LIKE is almost unanimously the most prominent form of LIKE across varieties of English, this clearly does not apply to IndE. The dominance of non-clausal LIKE in IndE is indicative of adaptational processes which have affected LIKE's syntactic properties during its implementation.

The idiosyncratic profile apparently corroborates Meyerhoff and Niedzielski's (2003:538) finding that in cases where quality face-to-face contact is lacking, only fairly superficial aspects of innovative forms are transferred. In this view, the idiosyncratic distribution of LIKE in IndE is caused by a lack of exposure to vernaculars in which LIKE has been integrated and subsequently stabilized. In turn, the lack of exposure granted a higher degree of flexibility with respect to the functionality, positioning, and association attached to LIKE. However, the similarity between CanE and IndE with respect to the constraints of clause-medial LIKE suggest that this systematicity in structural constraints is caused by language-internal or psychological factors.

A possible weakness of the results of LIKE use in IndE relates to the database of the present study, and to the adequacy and representativeness of the IndE data and whether it actually reflects the socio-cultural and linguistic diversity of the Indian population. With respect to the social status of speakers, it has to be noted that the ICE data is biased towards the upper end of the social hierarchy, because the ICE components aim to represent the national standard varieties of English (cf. Greenbaum 1988). This implies, of course, that speakers towards the lower end of the hierarchy are underrepresented. This problem is unavoidable, given that English is considered prestigious and not readily available to lower class speakers as a means of communication.

With respect to the representation of regional sub-varieties of IndE, the Indian component of the ICE provides rather balanced data; interviews have been recorded mostly in the western and urban parts of the Indian

subcontinent, however, leaving eastern varieties in particular underrepresented.

Table 101: Location of interviews in the S1A files of the Indian component

Region	Location	Number of interviews	Speakers
West	Bombay (Mumbai), Kolhapur, Nipani, Pune, Sankeshwar , Shivaji	42	91
North/ North-East	Chandigarh, New Delhi, Patna	14	30
South/ South-West	Dharwad, Madras (Chennai), Mysore	31	85
Centre	Hyderabad	9	22
--	NA	4	15

In terms of overall frequency, the analysis of the Indian data shows that LIKE as a vernacular feature is salient, since approximately forty -five percent of speakers in the data have used it at least once. The average use of LIKE in AmE, CanE and IrE, however, vastly exceeds the frequency observed in IndE. Nonetheless, non-clausal LIKE is used over proportionately compared with its use in other varieties while the clausal variants of LIKE are marginal at best. Indeed, non-clausal LIKE is highly frequent in comparison to other locales. The overall distribution of functionally distinct uses of LIKE thus challenges Valentine's assertion, in so far as she proposes that LIKE in IndE is similar to the colloquial American use of this marker (1991:332).

Moreover, the over-propotionate use of non-clausal LIKE clearly substantiates Valentine's (1991) hypothesis that LIKE is not performing the whole range of functions as it does in other regional varieties (1991:332-333). Hence, the evidence at least partially corroborates Valentine's (1991) discourse pragmatic analysis in that LIKE predominantly serves as a floor-holding device, to introduce phrasal specifications, or to buy processing time. However, the examples of clause-medial LIKE provided here illustrate that LIKE performs functions other than mere filling or introducing examples. In fact, clause-medial LIKE clearly functions as a hedge or focusing device similar to its use in prototypical inner circle varieties.

With respect to clause-initial and non-clausal uses of LIKE, the apparent-time distribution follows the common age stratification indicative of age-grading which, in case of non-clausal LIKE, even co- occurs with a female lead. Age-grading paired with a consistent female lead suggests ongoing change. However, the real-time analysis failed to corroborate change over time but it has to be kept in mind that the real-time results are tentative at best, as the validity of the real-time analysis suffers from the brevity of the time span covered by the data. This shortcoming rendered it unlikely to detect any tendencies in real-time and results have to remain preliminary until further research probes deeper into the matter.

The final question to be addressed here concerns the linguistic diversity available in the multilingual Indian setting. It is remarkable, given the number of first languages spoken in India, that the L1 of speakers did not significantly correlate with the use of LIKE. This is particularly noteworthy, as previous studies on LIKE use have suggested that the L1 can significantly impact the manner and rate of LIKE use (cf. Kastronic 2011; Müller 2005; Sankoff et al. 1997). In particular, Sankoff et al. hypothesized that an increase in the use of *comme* as a punctor, i.e. a modifying particle, was "influenced by their LI use of *like.* […] Transfer to the use of *comme* in French is thus a highly likely explanation for the patterns we observed." (1997:208). The study by Sankoff et al. (1997) shows that the use of certain pragmatic markers in one's L1 can clearly interfere notably with the use of similarly functioning elements in a speaker's L2 (Sankoff et al. 1997:231). In other words, the use of LIKE by Indian speakers may be affected by their use of similar markers in their L1. It can thus be argued that the high rate of non-causal LIKE is a product of substrate interference or shining-through effects of, for instance, Hindi, Kannada, or Punjabi. Considering the findings of the present analysis, this account is not implausible and offers a testable hypothesis for further research. In the present analysis, the number of speakers with different mother tongues is too small to detect L1 interference by applying the statistical models used here.

Sankoff et al. (1997) offer an alternative view on the matter. Consider Sankoff et al. (1997:214-215):

> Paradoxically, our results show that, as far as discourse markers are concerned, the educational system alone cannot supply children with these resources. As a vernacular feature, the appropriate use of discourse markers requires exposure to the vernacular. If French bilingualism indeed becomes normal for Anglophone Montrealers, we can expect vernacular competence to increase.

Applied to the present data, this means that speakers of IndE may not have been sufficiently exposed to LIKE use to have acquired the full set of its pragmatic functions. This is not to say that LIKE will necessarily approximate the distributions observed in prototypical inner circle varieties. For example, its use might be constrained by substrate interference, or it might fail to outperform rivaling forms. Nonetheless, the present analysis of the Indian data draws attention to supra-varietal patterns: the age stratification emerging from its distribution mirrors the patterning observable across varieties.

7.6 LIKE in Jamaican English[67]

To date, the discourse marker LIKE in JamE has remained largely unnoticed and has escaped the attention of scholars of this regional variety. Although it is listed as a focusing device in Jamaican Creole by Kortmann et al. (2004), it is neither illustrated nor explicitly discussed. Nonetheless, studying vernacular uses of LIKE in JamE is promising, as the target of teaching is still EngE while the regional Creole-like varieties have gained ground since WWII, particularly as a result of a growing sense of nationalism (cf. Schneider 2007:234). The rejection of EngE due to the increasing prestige of local identity paired with the geographic proximity to AmE predicts that JamE speakers assimilate to the AmE or CanE distribution rather than rejecting LIKE as EngE speakers do. Hence, JamE offers intriguing insights about the impact of the mass media as a means of diffusion particularly in regard to the use of clause-medial LIKE. If the global spread of (clause-medial) LIKE in the twentieth century indeed has its epicenter in the US, then this would predict that clause-medial LIKE exhibits a substantial degree of age stratification as a result of its relatively recent introduction into the speech community. Furthermore, if LIKE was introduced via the mass media, LIKE use should follow the distributional pattern of LIKE use in AmE or CanE rather than the pattern of EngE or IrE because the former varieties are more often represented in the mass media.

Figure 52 indicates that the distribution of clause-medial LIKE follows the AmE and CanE pattern rather than the EngE one. Although the age stratification strongly suggests that clause-medial LIKE has only relatively recently spread among the speakers of this regional variety, the quantity of LIKE examples among speakers aged 26 to 45 strongly suggests that this spread significantly predates the spread among British (pre-)adolescents. Furthermore, the somewhat steeper decline of LIKE use with increasing age among female

[67] For a complete overview of the final data set of JamE used in this study cf. Table 143 in the Appendix.

speakers suggests a female lead and thus ongoing change in real-time. The interpretation of the data is, however, only tentative as a result of a notable gender and age bias in the data.

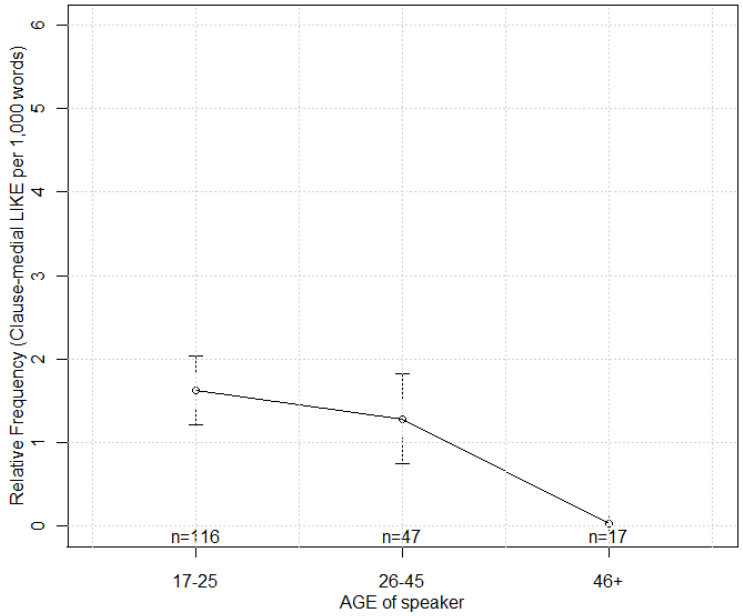

Figure 52: Age distribution of clause-medial LIKE in JamE

Before analyzing the data in detail, the basic characteristics of the data are given to provide a first impression of its structure. Of the 228 speakers present in the data, 121 make use of LIKE at least once. In other words, over half of JamE speakers used one or another variant of LIKE. This rather high quantity of LIKE identifies it as a salient feature of JamE vernacular despite not having received scholarly attention yet. Table 102 provides a first rough impression of LIKE use in this regional variety.

Table 102: Overview - LIKE use in JamE

	N (total)	**Mean (per 1,000 words)**	**Median (per 1,000 words)**	**Standard deviation (s)**
LIKE	577	2.447	0.630	3.780

Table 102 reports that LIKE is used with a mean frequency of 2.447 times per 1,000 words (mean) and that the median is at 0.630 instances per 1,000 words. The difference between mean and median indicates that LIKE use clusters and is thus not homogenous among JamE speakers, but exhibits an orderly heterogeneity.

Before turning to a more fine-grained analysis of the sociolinguistic profile of LIKE use in JamE, the frequencies of LIKE variants are displayed to illustrate the overall distribution of positionally and functionally distinct uses of LIKE.

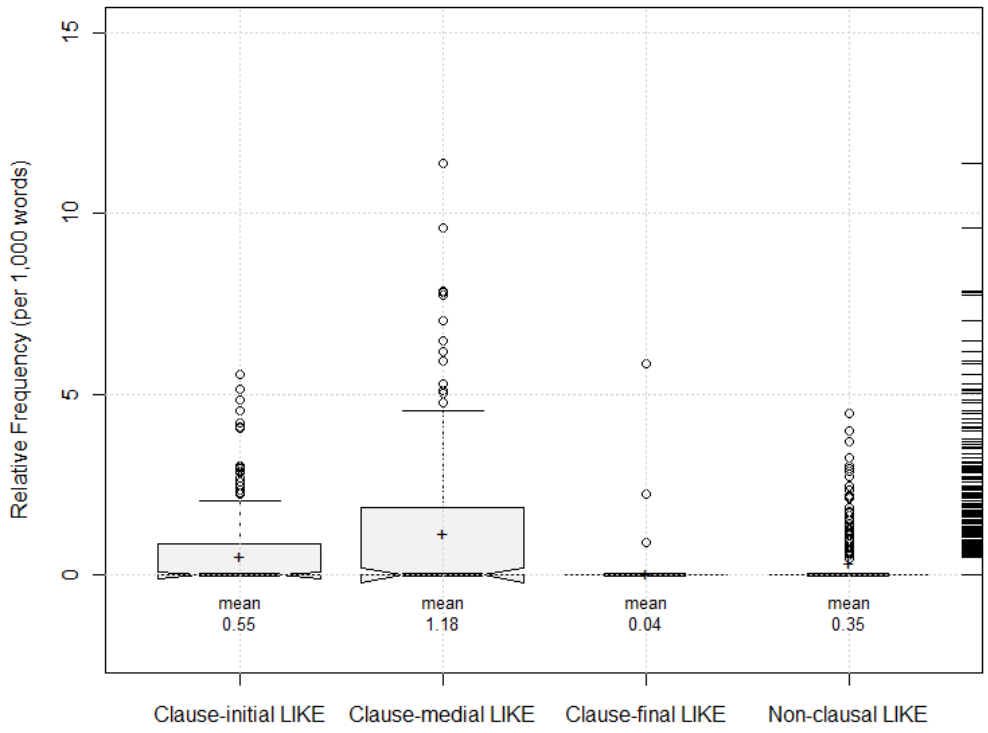

Figure 53: Rate of LIKE variants in JamE

Figure 53 immediately draws attention to the fact that JamE mimics the US American and Canadian distributions of LIKE variants as clause-medial LIKE is by far the most frequent variant.

Table 103: Age and gender distribution of LIKE in JamE[68]

			MALE		FEMALE			
AGE	**Speakers**	**Words**	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (17 – 25)	119	112,198	46	2.00	323	3.35	369	3.19
2 (26 – 45)	50	54,370	69	2.53	107	2.50	176	2.51
3 (46 – 65)	14	15,629	1	0.10	0	0.00	1	0.06
NA	49	28,131	15	0.10	30	1.84	31	1.27
SUM	232	210,328	131	1.18	460	1.92	577	1.76

Table 103 reports that LIKE is particularly common among younger females, and to a somewhat lesser degree among younger males, while speakers above age 46 seem not make use of LIKE at all. This pattern reflects the supra-local tendency for age-grading and confirms change in apparent-time. The difference between the mean frequencies male, female and all speakers in the youngest cohorts indicates a strong gender and age bias of the data, as female speakers age 17 to 25 seem to be over-represented while male and older speakers appear to be under-represented in this ICE component.

7.6.1 Statistical analysis of LIKE usage in Jamaican English

The following statistical analysis evaluates the intuitions derived from the summaries and graphical displays and seeks to evaluate general tendencies of LIKE use in JamE. Since statistical analyses are more powerful than visual inspections of the data, they aim at uncovering correlations which have previously escaped detection.

The independent variables included in the multivariate analyses are age (A1 = 17-25 years, A2 = 26 to 45, A3 = 46 years and older), gender, first language, education, and occupation (OCC) of speakers. The initial saturated model is fitted in a step-wise procedure and arrives at a minimal adequate

[68] Speakers for whom age and gender information is not available are classified as NA.

model that contains only main predictors. The same procedure is applied to study all other variants.

Table 104: Results of the multivariate regression for LIKE in JamE

ALL	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	0.93	0.326	2.86	<.01**
A2	0.79	0.298	2.65	<.01**
A3	-18.23	0.480	-37.91	<.001***
L1: ENG+	-1.01	0.273	-2.85	<0.01**
L1: PAT	-1.01	0.385	-2.63	<.01**
SML	0.86	0.335	2.58	<.01**

The final minimal adequate model used to analyze the overall frequency of LIKE per 1,000 words significantly out-performs an intercept-based baseline model (χ^2= 1835.9, df=5, p<.001***). Moreover, the regression output reports a rather complex setting of influences. The values of the age coefficients indicate that - when other variables are taken into account – middle-aged speakers use LIKE even more than younger speakers while the older speakers of JamE, i.e. speakers above the age of 46, do not use it at all. The total absence of LIKE from the vernacular of older speakers causes the extreme value of their coefficient. A second significant influence on LIKE use is the first language of speakers. Speakers whose first language is English only use LIKE significantly more than either speakers with Patois as L1 or bilingual speakers. It is tempting to attribute this effect to an underlying unknown variable, for example a socio-cultural factor, which differs among speakers of distinct L1 settings. However, as social class is accounted for in this multivariate analysis, the result does indeed indicate an interference effect, for example, incomplete acquisition.

The positive coefficient of (un-)skilled manual labor, furthermore, indexes that LIKE use is significantly socially stratified. Speakers pursuing academic and professional careers as well as speakers with clerical and managerial occupations use LIKE substantially less than (un-)skilled manual labors.

7.6.1.1 Clause-initial LIKE

So far, uses of the clause-initial discourse marker have not been analyzed in this regional variety. The examples in (103) suggest that clause-initial LIKE is functionally equivalent to its use in other varieties, as there discourse subsequent to these uses of LIKE clearly specify and elaborate on previous discourse.

(103) a. Yeah I mean if if *like* you're a little person then it could you know be daunting. (ICE Jamaica:S1A-009$B)

b. I didn't talk to them and then after a while we got through to each other and *like* the whole year I kind of freaked them out (ICE Jamaica:S1A-057$A)

c. Uh *like* I know I wonder who these strange people are that call him up and say all these nice things (ICE Jamaica:S1A-031$A)

As has already been shown, the use of LIKE in JamE is similar to the American pattern and not the British one. With respect to LIKE, the intriguing question is whether the sociolinguistic patterning reflects the American or Canadian distribution.

Table 105: Age and gender distribution of clause-initial LIKE in JamE

	MALE		**FEMALE**			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (17–25)	12	0.54	66	0.69	78	0.67
2 (26–45)	20	0.62	32	0.83	52	0.75
3 (46+)	1	0.11	0	0.00	1	0.05
SUM	33	0.49	98	0.67	131	0.63

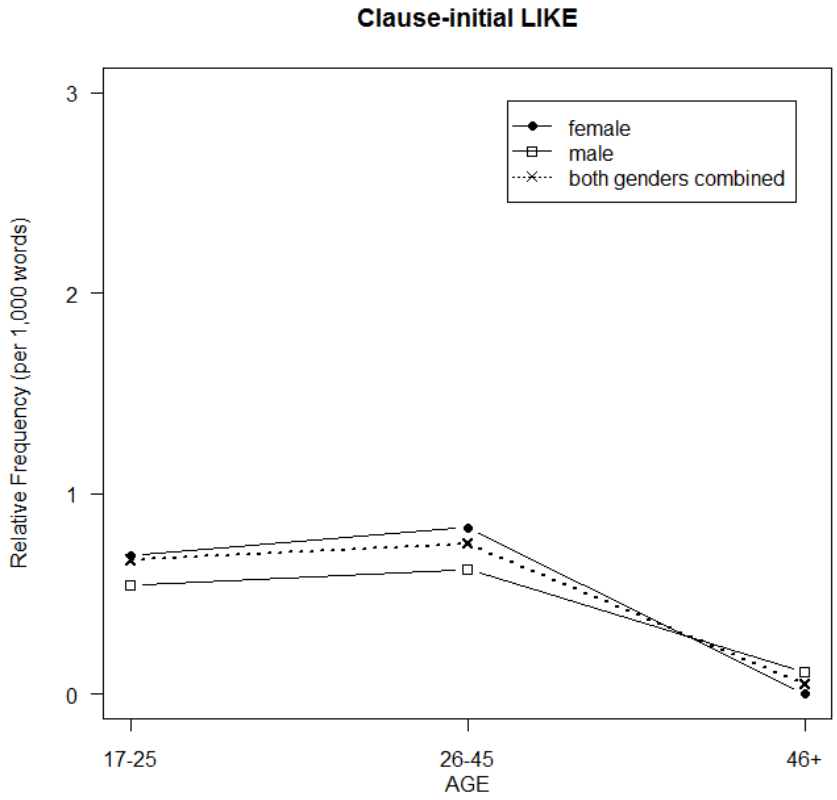

Figure 54: Age and gender distribution of clause-initial LIKE in JamE

Figure 54 indicates the absence of a uniform correlation of LIKE use with age, suggesting that clause-initial LIKE has stabilized at a moderate level of use without further increase among younger speakers. Accordingly, speakers from post-adolescence up to the mid-forties do not seem to differ with any socially motivated systematicity.

Table 106: Results of the multivariate regression for clause-initial LIKE in JamE

INI	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.98	0.214	-4.60	<.001***
A2	0.14	0.265	0.53	.59
A3	-1.89	0.956	-1.98	<.05*
PAI	0.08	0.018	4.45	<.001***

The regression output confirms pronounced age stratification in the form of a drastic difference in LIKE use between the two younger cohorts and the oldest speakers, which corroborates the impression derived from Figure 54.

While the model reports significant predictive power of the PAI index indicating priming or accommodation effects in conversations, it fails to confirm any additional correlations between social stratification and linguistic variation.

7.6.1.2 Clause-medial LIKE

Similar to clause-initial LIKE, clause-medial LIKE in JamE has escaped scholarly attention. As illustrated by the examples in (104), clause-medial LIKE conforms functionally with its use in other regional varieties. The instances provided below are representative of typical uses of clause-medial LIKE, which modify the subsequent constituent. With respect to their discourse-pragmatic function, (104a) and (104b) focus the following NP, while LIKE in (104c) hedges the subsequent PP. The hedging use of LIKE in (104c) communicates a certain degree of vagueness, which allows the listener to qualify or to indicate that the "for a gift" is a specification.

(104) a. I hear they mark *like* history lectures> and I don't do history cos it's hard (ICE Jamaica:S1A-009$B)

b. Yeah I find that understand patois I know I know people in *like* Upper St Andrew who just absolutely cannot fathom what it means. (ICE Jamaica:S1A-014$A)

c. Me not get it yet but me have it in mind *like* for a gift. (ICE Jamaica:S1A-025$B)

As in the case of clause-initial LIKE, the question arises whether the distributional pattern reflects ongoing change in apparent-time and where the individuals promoting LIKE are located.

Table 107: Age and gender distribution of clause-medial LIKE in JamE

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (17–25)	29	1.24	161	1.67	190	1.62
2 (26–45)	34	1.17	32	1.36	83	1.28
3 (46+)	0	0.00	1	0.07	1	0.03
SUM	63	0.96	211	1.50	274	1.38

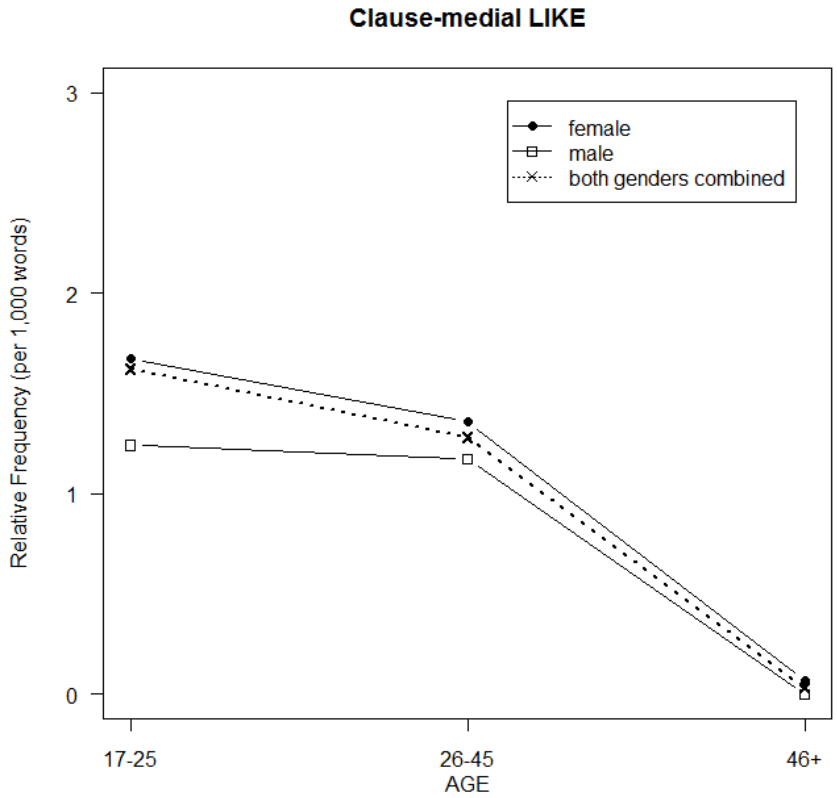

Figure 55: Age distribution of clause-medial LIKE in JamE

Figure 55 displays a similar distribution as in the case of clause-initial LIKE. The use of clause-medial LIKE is extremely rare among the oldest speakers, while speakers between 26 and 45 years of age use it almost as often as speakers between 17 and 25 do. This is particularly true of males, among whom the frequency of clause-medial LIKE is almost identical among middle-aged and young speakers. Women, on the other, hand show a more linear recess of LIKE use with increasing age. This recess reflects stereotypical age-grading or change in apparent-time.

Table 108: Results of the multivariate regression for clause-medial LIKE in JamE

MED	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	0.18	0.249	0.72	.46
A2	-0.06	0.170	-0.38	.70
A3	-17.34	0.38	-45.53	<.001***
PAI	0.09	0.020	4.76	<.001***
L1: ENG+	-0.56	0.212	-2.67	<.01**

The Poisson regression employed to analyze the use of clause-medial LIKE significantly out-performs the intercept-based base-line model (MED: χ^2=2577.0, df= 5, p<.001***). The output of the regression model reports a negative correlation between the age of speakers and the frequency of clause-medial LIKE. The extreme value of the coefficient for the oldest age group implies that clause-medial LIKE is not part of the grammar of speakers above age 42. The absence of this non-standard feature suggests either that it has only recently entered this speech community or that its use is covertly stigmatized among older speakers. However, the latter interpretation is not persuasive given the effect of the date of data compilation shown in Figure 57. Furthermore, fluent bilinguals use this form significantly less often than speaker who have been brought up as monolinguals with L1 English. The significant PAI index suggests clustering of LIKE use and thus that speakers who have acquired LIKE tend to use it freely, flexibly and frequently.

7.6.1.3 Clause-final LIKE

The low number of instances renders it inappropriate to using a regression model for clause-final uses of LIKE as in (105).

(105) a. I'm just come *like*. (ICE Jamaica:S1A-061#B)
b. That's wrong *like* (ICE Jamaica:S1A-031#B)
c. There's nothing here that really comes out *like* you know <,>that I really like to do. (ICE Jamaica:S1A-063#NA)

There are nine instances of clause-final LIKE in the Jamaican ICE data, eight of which occur in female discourse. None of these nine instances is uttered by speakers younger than 26 years of age. The context in which these instances

are used is indistinguishable from equivalent uses in IrE. Based on this similarity, it is reasonable to assume that clause-final instances of LIKE fulfill the same or very similar discourse pragmatic functions across varieties of English, which is to either focus or indicate a "minor non-equivalence between what is said and what is in mind" (Schourup 198:31). In other words, it appears as if clause final LIKE modifies preceding statements or elements thereof almost irrespective of geographical embeddedness of the discourse.

7.6.1.4 Non-clausal LIKE

In JamE, non-clausal LIKE as in (106) is substantially more often used than clause-final LIKE, but significantly less often used than clause-initial LIKE or clause-medial LIKE.

(106) a. You have some of them more *like like* for example one of my friend my gosh she just *like*<,> she's terrible uh terrible. (ICE Jamaica:S1A-051$A)

b. The mathematical part *like* uhm<,> how<,,> alright<,> Okay for instance<,> we learn<,> the... (ICE Jamaica:S1A-046$B)

c. Yes she's uh *like like* one of the normal<,> on it now. (ICE Jamaica:S1A-062$C)

Table 109: Age and gender distribution of non-clausal LIKE in JamE

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (17–25)	3	0.11	39	0.39	42	0.36
2 (26–45)	14	0.54	24	0.61	38	0.58
3 (46+)	0	0.00	0	0.00	0	0.00
SUM	17	0.29	63	0.41	80	0.38

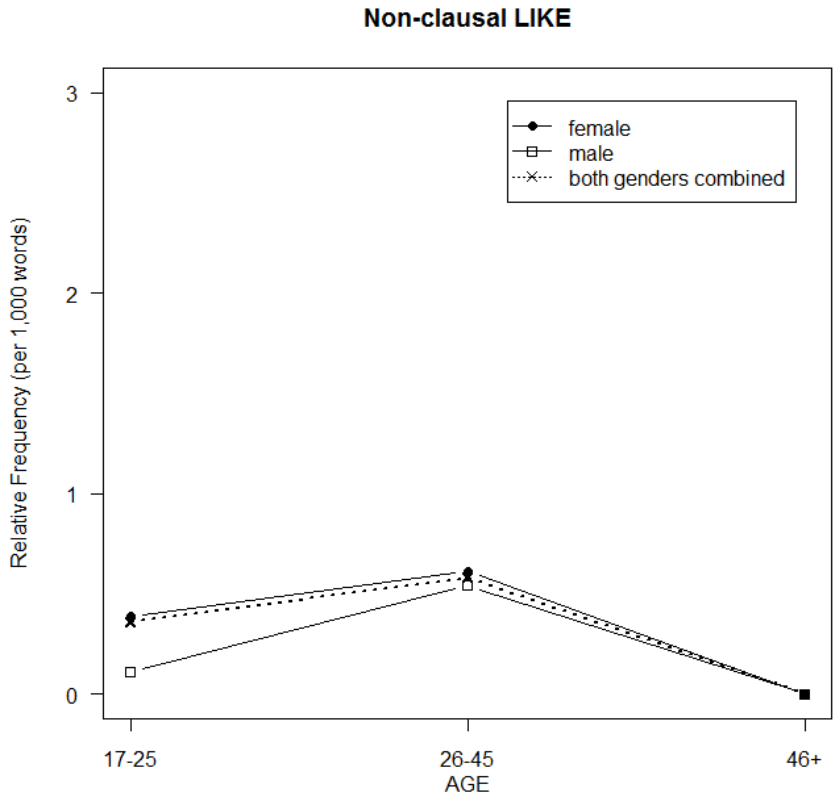

Figure 56: Age distribution of non-clausal LIKE in JamE

The pattern emerging from Figure 56 is unusual, as it suggests that non-clausal LIKE is used almost exclusively by speakers in their late twenties and early thirties, while it appears to be marginal among both the young as well as the old speakers.

Table 110: Results of the multivariate regression for non-clausal LIKE in JamE

NON	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.76	0.251	-7.00	<.001***
A2	0.52	0.293	1.80	<.1
A3	-16.58	0.345	-48.00	<.001***
PAI	0.09	0.020	4.69	<.001***

The regression output for the analysis of non-clausal LIKE mostly corroborates the impressions derived from Figure 56: the old speakers make almost no use of this form, while the middle aged speakers exhibit the highest rate when all variables are taken into account. The significant coefficient of the PAI index

informs us that the overall use of LIKE is a significant predictor for the use of non-clausal LIKE, which is not surprising considering effects of accommodation and priming. Otherwise, the output provides no further insights into the correlation between social stratification and use of non-clausal LIKE.

7.6.2 Evaluation of the apparent-time construct

The following section focuses on correlations between the date of data compilation and LIKE use to evaluate the apparent-time findings from a real-time perspective. If the date of data compilation significantly affects the frequency of LIKE in this within group analysis, this would shed light on the use of LIKE over time. From the apparent-time analysis, we would predict that LIKE was less frequent from 1990 to 1994 compared with 2002 to 2005. If this increase in LIKE use over time was the case, this would support the assumption of the apparent-time construct that speaker age reflects different stages of the communal grammar.

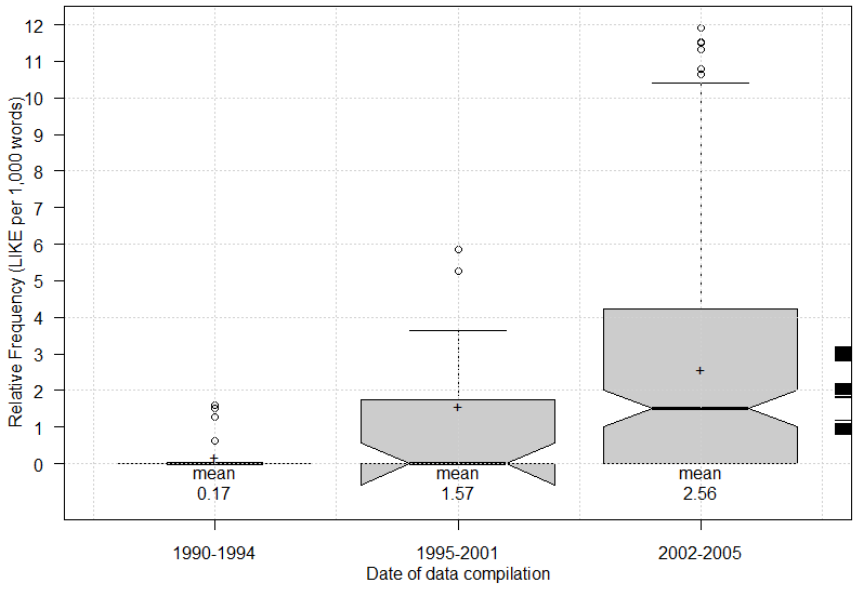

Figure 57: LIKE in JamE with respect to the date of data collection

Figure 57 clearly suggests an increase in LIKE use over time, i.e. it confirms the apparent-time hypothesis as it demonstrates that the age-grading in LIKE reflects ongoing changes in real-time.

Table 111: LIKE use per 1,000 words in JamE according to the date of data compilation

	ALL	INI	MED	FIN	NON
D1	0.17	0.05	0.12	0.00	0.00
D2	1.56	0.35	0.81	0.25	0.13
D3	2.56	0.65	1.40	0.01	0.42

With the exception of those for clause-final LIKE, all mean values show a remarkable, steady increase over time. The existence of this trend is endorsed statistically, as all return significant (ALL: χ^2= 9.058, df= 1, p<.01**; INI: χ^2= 6.122, df= 1, p<.05*;MED: χ^2= 9.108, df= 1, p<.01**). This confirms the prediction that the frequency of LIKE and its subtypes has increased significantly from the earliest period of data collection to the most recent.

In a next step, it is tested whether the proportion of speakers who used LIKE has significantly increased over time.

Table 112: Mean frequencies of clause-medial LIKE use; number of non-like users; and like users plus their ratio; and the percentage of LIKE users in JamE, according to the date of data compilation

	LIKE users	Non-LIKE users	Ratio	Percentage of LIKE users of all speakers
D1	4	25	0.16	13.7
D2	7	16	0.43	30.4
D3	110	66	1.66	62.5

The increasing values of the ratios of non-LIKE users and LIKE users, as well as the percentage of LIKE users within this community, clearly indicate that the proportion of LIKE users has dramatically increased over time, adding further evidence to the hypothesis that LIKE use has undergone a dramatic change in the recent history of JamE. Accordingly, the impression derived based on an apparent-time hypothesis – according to which LIKE has been undergoing change in JamE in the sense that it continues to spread within this speech community – is corroborated by this real-time analysis.

7.6.3 Summary: LIKE use in Jamaican English

Vernacular uses of LIKE have become a common phenomenon in JamE. With respect to the distributional pattern, we observe that JamE resembles AmE: both varieties show a pronounced preference for clause-medial and, to a lesser degree, clause-initial LIKE, while both clause-final and non-clausal LIKE are marginal. Functionally, uses of LIKE in JamE match the discourse-pragmatic analysis of other regional varieties, indicating sufficient exposure to other vernaculars – very likely AmE vernacular.

The diversity of the sociolinguistic profiles of LIKE in JamE is startling. This is particularly true for non-clausal LIKE, which not only lacks sloping age stratification, but its use seems to be more or less confined to middle-aged speakers. Though less pronounced, similar trajectories emerge for clause-initial and clause-medial LIKE. The patterning in JamE is thus distinct from the near linear negative correlation of LIKE use with increasing age observed in the majority of cases dealt with so far. Possibly, this divergence indexes a difference in LIKE's social embedding and, thereby, its social meaning. In other words, while LIKE is associated with friendliness, cheerfulness, attractiveness, and successfulness in AmE (Dailey-O'Cain 2000:73), it is reasonable to assume that it carries different connotations in JamE. As the social embedding of elements is crucial with respect to marking in-group membership, differences in the social embedding will be reflected in distinct trajectories of social dispersion. Whatever the truth may be, this patterning has yet to be explained by means of more elaborate analyses able to probe more deeply into the socio-cultural embedding of this form.

Strikingly, LIKE use in JamE has drastically increased over a fairly short period of time. The proportion of speakers using LIKE at least once during the latest period of data compilation (i.e., 62 percent), suggests that this dramatic change is continuing from mid-range to nearing completion (Nevalainen-Raumolin-Brunberg 2003:55). The rate at which LIKE has been spreading is extraordinary, as the proportion of speakers has increased more than fourfold.

Contrary to IndE, the analysis detected significant effects of language contact with respect to clause-medial LIKE and overall LIKE use. The effects of a speaker's first language can be interpreted in at least two ways: either these effects reflect social stratification, or they can be viewed as suggesting substrate interference. The former approach is consistent with previous findings which show that social stratification is typical of ongoing change, as new features do not intrude into all social strata at the same rate. The latter view is corroborated by the fact that pragmatic markers require substantial contact and a high proficiency in the target language to be used in similar rates and with the same functions. Müller (2005), for instance, found that L1 German speakers differ markedly form L1 AmE with respect to their use of LIKE and attributes this difference to insufficient exposure of L1 German speakers to L1 English vernacular (Müller 2005:239).

Clause-initial LIKE and non-clausal LIKE are not affected by a speaker's L1, but show a similar increase over real-time (cf. Table 111), thereby suggesting that these forms are also part of this process. It is however noteworthy that this obvious change is not accompanied by gender differentiation or social-stratification beyond age-grading. The rapid increase in LIKE use indicates that LIKE in JamE is undergoing not generational but communal change, in which where the individual and the community change together.

The profile of LIKE in JamE differs distinctly from its use in EngE, despite the fact that EngE is still the target variety in Jamaican schools (Schneider 2007:234). Hence, the present findings are in line with Mair's (2009) finding that "[o]n the whole, present-day JamE turns out to be rather different from EngE, its historical 'parent' variety" (2009:39). In fact, with respect to the frequency of clause-medial LIKE particularly vernacular JamE, i.e. mesolectal Jamaican creole, appears to be more heavily influenced by American than by EngE.

7.7 LIKE in New Zealand English[69]

The discourse marker LIKE in NZE and AusE has only recently been discovered as a subject for linguistic research. One study addressing in particular vernacular uses of LIKE in NZE and AusE is Miller's (2009) analysis of discourse markers based on the Australian and New Zealand components of the ICE and the ART corpus. While Miller's (2009) analysis of LIKE use provides an extensive overview of its functional and positional distribution in scripted and unscripted spoken discourse, it is rather superficial with respect to its sociolinguistic distribution. Based on his data, Miller (2009:317) asserts that LIKE is particularly common among teenagers who are, however, "not the sole users" (Miller 2009:324), as it is also attested in the speech of 50-year-olds (2009:317). Regarding social stratification, Miller (2009:327) confirms that it is widely used by speakers of diverse social backgrounds "including manual workers, skilled tradesmen, and various types of professionals" (Miller 2009:317). Moreover, D'Arcy (2007) shows that discourse marker uses of LIKE have a long history in NZE, as it is attested in radio transcripts from 1946 to 1948 (D'Arcy 2007:401). In fact, all instances occur in the speech of NZE speakers born between 1851 and 1919. Based on the fact that speakers using LIKE had immigrated from England, Scotland and Ireland, it is highly plausible that this discourse marker was imported from the British Isles rather than emerging from contact or as a parallel development.

The present analysis of LIKE use in NZE elaborates on Miller's study and offers a more fine-grained sociolinguistic analysis, as it covers a wide range of relevant extra-linguistic factors.

Before analyzing the data in detail, the the basic statistics of the data are summarized in Table 113 to get a first impression of its structure. Of the 227 speakers present in the data, 146 make use of LIKE at least once, which

69 For a complete overview of the final data set of NZE used in this study cf. Table 144 in the Appendix.

amounts to nearly two-thirds of NZE speakers. This high proportion of LIKE users is, however, not unexpected, as Miller (2009:232) reports a frequency of 2.78 instances of LIKE per 1,000 words in NZE and 3.41 instances per 1,000 words in AusE.[70]

Table 113 reports the basic statistics of the NZE data and provides a first rough impression of LIKE use in this regional variety.

Table 113: Overview - LIKE use in NZE

	N (total)	**Mean (per 1,000 words)**	**Median (per 1,000 words)**	**Standard deviation (s)**
LIKE	529	2.175	1.229	2.766

On average, LIKE is used slightly more than two times per 1,000 words (mean). The rather high median and the small standard deviation demonstrate that LIKE use is very common and rather homogenous in NZE. Before turning to age- and gender-specific usage patterns, we will have a brief look at the frequency of positionally and functionally distinct types of LIKE, which serves to evaluate cross varietal similarities, i.e. which varieties are most similar to NEW Zealand with respect to their LIKE use.

[70] In fact, Miller (2009: 323) reports a frequency of 27.8 and 34.2 instances of LIKE per 10,000 words, respectively, which have, however, been transformed into per 1,000 word frequencies to fit the other frequencies reported in the present analysis.

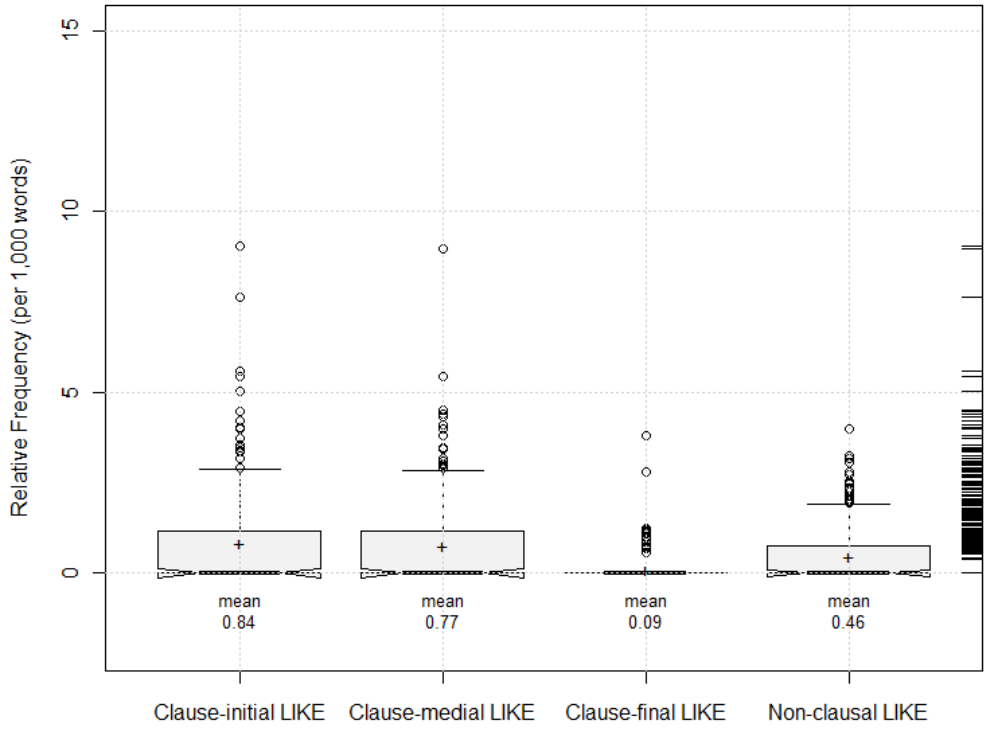

Figure 58: Rate of LIKE variants in NZE

The distribution of LIKE variants in NZE approximates the distributions of CanE and to a somewhat lesser extent, AmE. In contrast to CanE and AmE, LIKE use in NZE is less common, and it is not clause-medial LIKE which is the most frequently used variant, but clause-initial LIKE – this difference is, however, rather small and in all likelihood insignificant.

Table 114: Age and gender distribution of LIKE in NZE

			MALE		FEMALE			
AGE	**Speakers**	**Words**	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (16 – 19)	36	35,176	46	3.19	98	4.13	144	3.84
2 (20 – 29)	105	100,740	106	2.54	195	2.78	301	2.69
3 (30 – 39)	31	32,701	23	1.37	21	1.00	44	1.15
4 (40+)	55	60,576	16	0.72	24	0.66	40	0.69
SUM	227	229,193	191	1.99	338	2.29	529	2.175

Table 114 strongly suggests that LIKE use is significantly age-graded, as it appears to recess with age. In addition, it indicates an interaction between age and gender, since post-adolescent females are ahead of their male peers; the opposite is true of older cohorts, where males are in the lead. This difference is, however, rather subtle, so that the reality of this interaction may only be confidently verified by means of advanced statistical procedures.

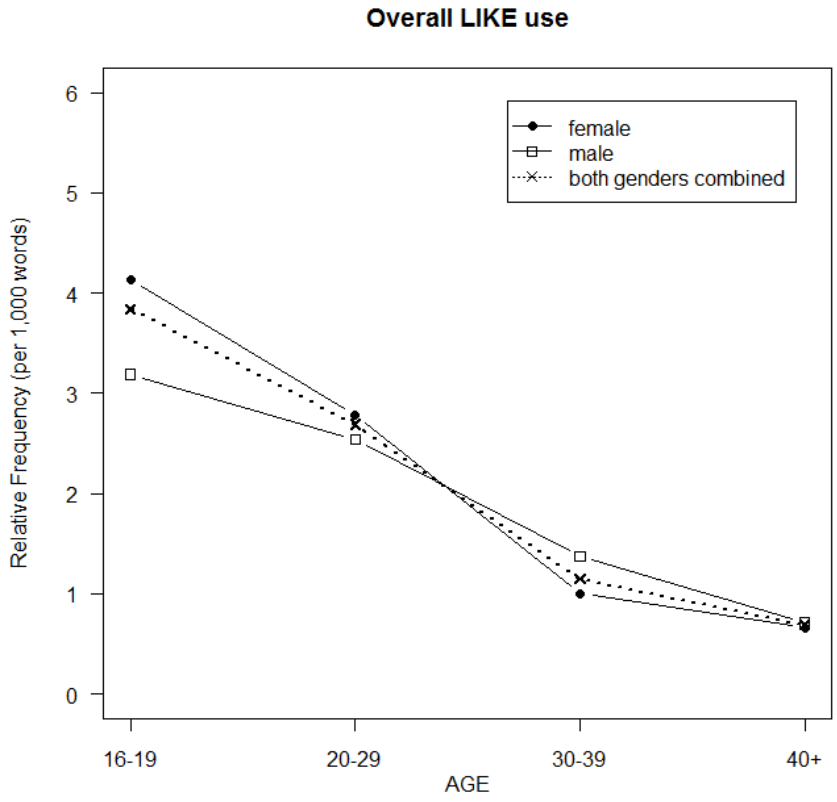

Figure 59: Age and gender distribution of LIKE in NZE

Figure 59 shows the common age-graded pattern, i.e. a decrease in LIKE use with increasing age. With respect to patterns of gender- and age-related LIKE use, Table 114 suggests a marginal female lead among speakers below age 30, while the proximity of the trend lines indicates an absence of significant gender effects. This similarity of trajectories indicates that if LIKE is undergoing change in real-time, this change is not accompanied by gender-specific usage patterns and thus is apparently not a female-dominated change.

The negative correlation between a speaker's age and LIKE use is ambivalent, as it may reflect either age-grading or ongoing change in real-time. Given the ambivalence of this distribution, it is useful to assess the effects of socio-economic status. In contrast to stable age-grading, real-time changes are accompanied by social stratification in the majority of cases (Labov 2002). Hence, real-time change would predict significant effects of social class possibly accompanied by age-grading, while the absence of social layering of LIKE use would corroborate findings that the age distribution reflects age-grading only.

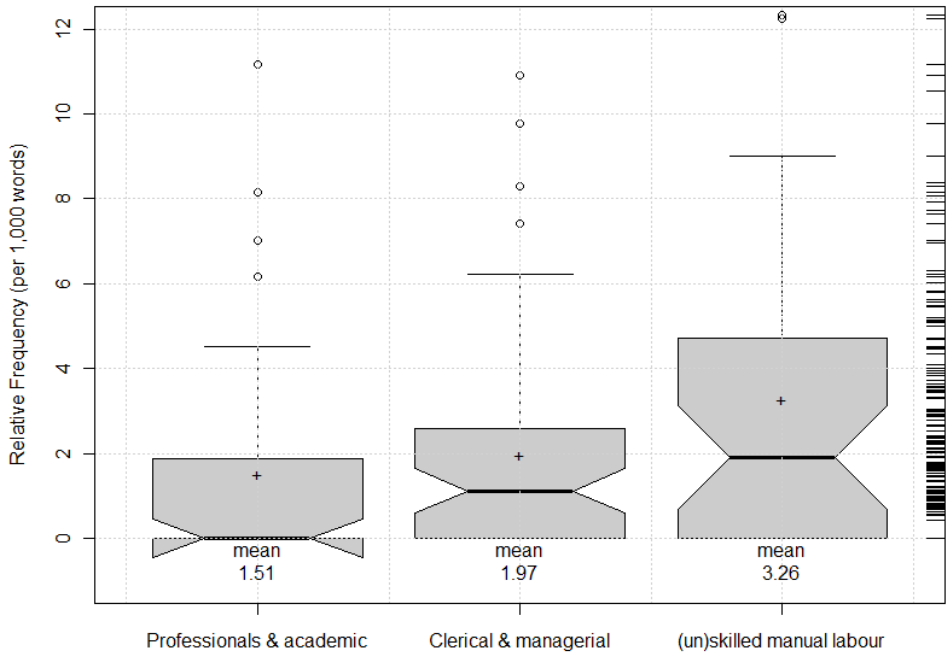

Figure 60: LIKE use in NZE with respect to the occupation of speakers

The distribution in Figure 60 suggests a significant correlation between social class and the use of LIKE as the rates appear to decrease towards the higher end of the socio-economic hierarchy. In other words, that speakers pursuing professional and academic careers use LIKE substantially less often than speakers in clerical and managerial occupations, and skilled and unskilled manual labor. Considering the age distribution, the notable social layering of

LIKE use corroborates ongoing change in real-time, which is counter-indicative of age-grading only, but suggestive of a conglomerate of simultaneous processes.

7.7.1 Statistical analysis of LIKE usage in New Zealand English

The following statistical analysis aims to evaluate the intuitions derived from the summaries and graphical displays of LIKE use in NZE. As this analysis is more powerful than mere visual inspections, it may uncover correlations which have so far escaped detection.

The multivariate analyses include age, gender, occupation, and the ethnicity of speakers as main predictors, as well as secondary interactions between age and gender. The saturated models are fitted in a step-wise procedure to arrive at a minimal adequate model containing only predictors whose coefficients are significant. The explanatory power of final minimal adequate models is tested by comparing them to base-line models using the intercept as sole predictor variable. The same procedure is applied to study clause-initial LIKE, clause-medial LIKE, and syntactically unbound LIKE without scope, i.e. non-clausal LIKE.

Table 115: Results of the multivariate regression for LIKE in NZE

ALL	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	1.35	0.134	10.08	<0.001***
A2	-0.37	0.172	-2.14	<0.05*
A3	-1.23	0.296	-4.17	<.001***
A4	-1.72	0.268	-6.44	<.001***

The final minimal adequate model is statistically significant (χ^2= 48.573, df= 3, p<.001***) and reports that speakers older than 20 use LIKE significantly less often than post-adolescents aged 16 to 19, i.e. the reference group. In addition, the multivariate model confirms a near-linear negative correlation between the frequency of LIKE and the age of speakers reflected in the steady increase in the age coefficients. The fact that neither any other main effects nor

interactions reach the five percent level of significance implies that none of the variables other than age is able to predict the frequency of LIKE better than chance alone. This finding suggests that LIKE use in general is neither confined to certain social strata nor socially layered. Accordingly, the effect of occupation indicated by Figure 60 appears to be true only of certain types of LIKE, which underpins the importance of complementing inferences based on visual inspections with statistical testing.

7.7.1.1 Clause-initial LIKE

After investigating factors which affect overall LIKE use, the following section will determine whether functionally distinct uses of LIKE exhibit similar or divergent social profiles. The first type of LIKE to be analyzed is the clause-initial discourse link, as in (107), which, according to Miller (2009:331), is used freely in both public and private dialog. Functionally, this linking use of LIKE highlights the upcoming clause (2009:331) and serves to "flag crucial information to the listener" (Miller 2009:332).

(107) a. [B]ut *like* i've mean for every year. (ICE New Zealand:S1A-015$A)

b. [A]nd *like* i didn't even get out of my netball skirt till about eight o clock (ICE New Zealand:S1A-022$B)

c. [*L*]*ike*]] now i've stopped buying nail polish and i'm using... (ICE New Zealand:S1A-022$H)

The examples in (107) corroborate Miller's (2009) discourse-pragmatic analysis, as they either introduce explanations of previous discourse, as in (107b) and (107c), or specify preceding utterances as in (107a). The discourse link in NZE thus coincides functionally with clause-initial uses in other regional varieties.

We will now turn to the sociolinguistic profile of clause-initial uses.

Table 116: Age and gender distribution of clause-initial LIKE in NZE

AGE	MALE		FEMALE			
	LIKE (N)	LIKE (mean)	LIKE (N)	LIKE (mean)	SUM (N)	ALL (mean)
1 (16–19)	15	0.94	46	1.99	61	1.67
2 (20–19)	38	0.85	76	1.02	114	0.95
3 (30-39)	5	0.27	12	0.60	17	0.46
4 (40+)	7	0.30	10	0.28	17	0.29
SUM	65	0.64	144	0.96	209	0.84

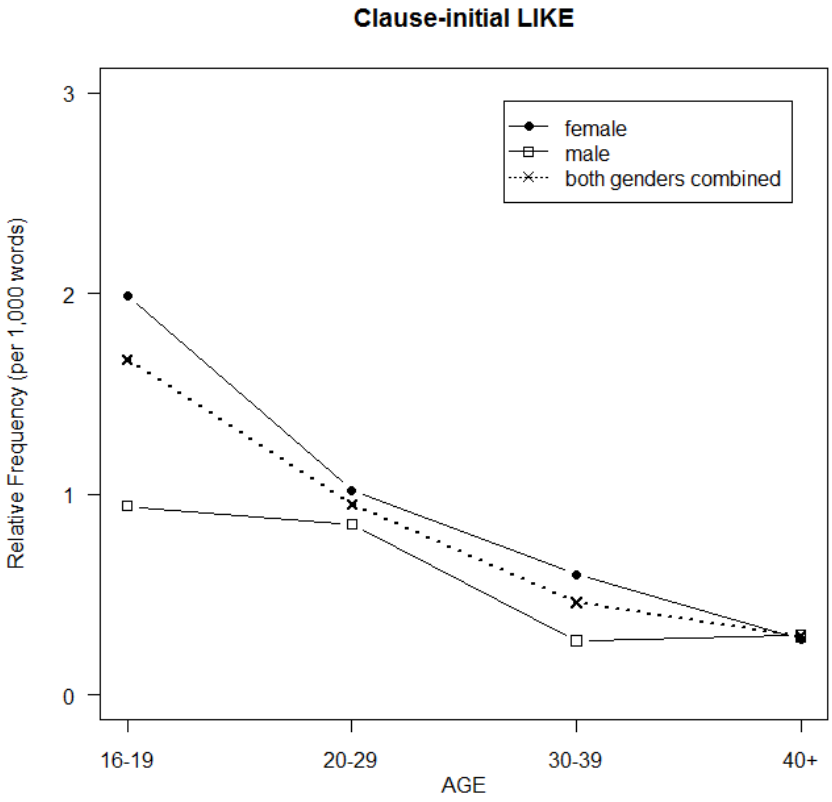

Figure 61: Age and gender distribution of clause-initial LIKE in NZE

According to Table 116 and Figure 61, clause-initial LIKE is age-graded and gender-sensitive, as the distribution suggests a coherent female lead. But while the difference appears insignificant among speakers older than 20, there is a substantial difference between the genders among post-adolescents. This interpretation is evaluated statistically by means of non-parametric t-tests

rather than a multivariate analysis as the regression model suffers from over-dispersion.

Table 117: Clause-initial LIKE with respect to GENDER and AGE

	A1	A2	A3	A4
INI	1.995* (<.05)	0.463 (n.s.)	1.339 (n.s.)	-0.459 (n.s.)

Table 118: Clause-initial LIKE with respect AGE

	A1-A2	A1-A3	A1-A4
INI	1.967* (<.05)	3.540*** (<.001)	4.153*** (<.001)

The t-statistics confirm a significant gender difference among post-adolescents younger than 20 years of age, and a negative correlation between the use of clause-initial LIKE and the age of speakers. The steady increase in the t-values from young to old suggests that this negative correlation is near-linear, and thus substantiates the impressions derived from Figure 61. Indeed, the ordering of t-values is suggestive of an implicational hierarchy of the form A1>A2>A3>A4. Additional t-statistics comparing A2 to A3 (t-statistic= 2.666**; p<.01) and A3 to A4 (t-statistic= 1.110; p=0.135) offer a more accurate understanding. Based on these additional tests, the hierarchy is more appropriately described as A1>A2>(A3+A4). This patterning indicates age-grading with a general trend towards lesser LIKE use among older speakers and higher frequencies of LIKE among the young.

7.7.1.2 Clause-medial LIKE

Based on the functionally oriented analysis of LIKE use in NZE and AusE, Miller (2009) concludes that clause-medial LIKE, as in (108), "is shown to put the spotlight on the following part of information and give it additional rhetorical and dramatic force" (2009:334). Although Miller (2009) validates the assumption that LIKE use is cross-culturally rather similar, he does not address issues related to systematic variation of LIKE in the New Zealand speech community. In order to complement Miller's pragmatic analysis, the following section will provide a statistical analysis of sociolinguistic factors determining its use.

(108) a. [T]hey're going to start a um a composting station which is for *like* green waste which is trees and stuff rather than food. (ICE New Zealand:S1A-014$F)

b. [A]nd there's like a paddock on each side and there's rows of pine trees down the middle you know (ICE New Zealand:S1A-033$A)

c. [H]e's *like* a real tall skinny little adolf <.>hitler. (ICE New Zealand:S1A-041$J)

d. [S]he's got hair to about here and it's curly and she's *like* dresses a bit different yeah (ICE New Zealand:S1A-022$H)

e. [A]re they going to *like* cut up the shirts (ICE New Zealand:S1A-042$J)

f. [H]e can't even talk to her about anything eh and she just <.>*like* quite mean really (ICE New Zealand:S1A-041$J)

g. [A]nd i think people are either a parent *like* permanently or else they're er a spouse a husband or wife (ICE New Zealand:S1A-040$A)

The NZE instances of clause-medial LIKE are functionally and syntactically equivalent to uses in other varieties. Adding weight to this view is the fact that the instances of LIKE in (108a), (108b) and (108c) have scope over a following NP; the instances in (108d) and (108e) have scope of the following VP; and those in (108f) and (108g) have scope over the subsequent AP. In all cases, LIKE precedes the lexically heavy elements and occurs either at the boundaries of phrases, or, in case of PPs, between the prepositions and the determiner, as in (108a). These syntactic environments match the circumscribed contexts of clause-medial LIKEin CanE (cf. D'Arcy 2005, 2007, 2008; Tagliamonte 2005). Therefore, focus LIKE in NZE is clearly allowed in quite distinct syntactic environments, which points to the fact that it is syntactically bound.

Table 119: Age and gender distribution of clause-medial LIKE in NZE

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (16–19)	21	1.43	28	1.15	49	1.24
2 (20–19)	42	1.01	60	0.94	102	0.97
3 (30-39)	14	0.88	4	0.16	18	0.46
4 (40+)	7	0.33	7	0.17	14	0.24
SUM	84	0.87	99	0.70	183	0.76

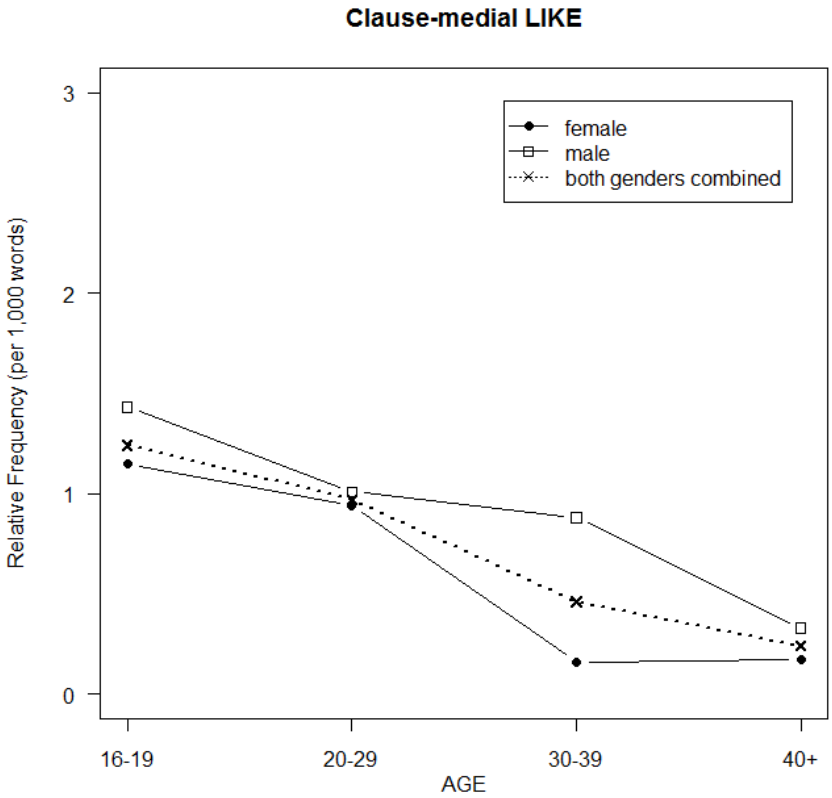

Figure 62: Age and gender distribution of clause-medial LIKE in NZE

The near linear decrease in clause-medial LIKE with increasing age in Figure 62 indicates age-grading and a notable gender difference among speakers aged 30 to 39. What is remarkable in the present case is that it is not female speakers who use clause-medial LIKE most, but males.

Table 120: Results of the multivariate regression for clause-medial LIKE in NZE

MED	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.13	0.433	-2.61	<.01**
SEX: Male	0.85	0.327	2.90	<.01**
A2	-0.08	0.402	2.16	<.05*
A3	-0.28	0.487	-0.59	.55
A4	-0.96	0.564	-1.70	<.1
PAI	0.16	0.023	7.02	<.001***
ADC	-073	0.277	-2.65	<.01**
SML	-0.26	0.247	-1.06	.28
SEX: Male: A2	-1.24	0.418	-2.97	<.01**

The final minimal model contains insignificant coefficients, which had to remain in the model; dropping them would have resulted in unreliable coefficients, as they would otherwise have been added to the reference category.

The regression model reports a gender difference in the expected direction. Given that Figure 62 indicated that the gender difference is confined to speakers in their thirties, it is somewhat surprising that gender breaches the level of significance. When taking a closer look at the effect of gender within each age cohort using χ^2-tests (cf. Table 121), the regression report is somewhat qualified. Indeed, χ^2-results support the pattern of Figure 62, i.e. that the gender difference is indeed confined to speakers aged 30 to 39. This result is, nonetheless, reflected in the regression model as it reports a significant interaction between gender and age group 3.

Table 121: χ2-test results: the effect of gender on clause-medial LIKE within age groups In NZE[71]

MED	χ2-value	df	p-value	Cramér's φ
A1 (16-19)	2.269	1	.13	0.008
A2 (20-29)	0.343	1	.55	0.001
A3 (30-39)	10.763	1	<.01**	0.018
A3 (40+)	0.806	1	.36	0.003

[71] Only speakers whose age was given are included.

The order of age coefficients suggests a monotonic decline, which typically suggests age-grading. However, the significant coefficients for social class confirm significant social layering implying that clause-medial LIKE is most common among speakers in clerical and managerial occupations, i.e. in the central stratum. The social stratification adds weight the assumption that age-grading alone is insufficient to account for the variation in LIKE use. Thus, the social stratification reported by the regression model confirms that LIKE is the subject of ongoing real-time change (Labov 2001:319-320).

7.7.1.3 Clause-final LIKE

With respect to the function of clause-final LIKE, Miller (2009) states that it "marks the clinching argument in an exploratory conversation" (Miller 2009. 335). On first sight, this may be taken to imply that clause-final LIKE in NZE differs functionally from equivalent uses in LIKE in IrE and EngE. With respect to IrE and EngE, Jespersen (1954) claims that clause-final LIKE is primarily used "parenthetically by inferiors addressing superiors to modify the whole of one's statement, a word or phrase, modestly indicating that one's choice of words was not, perhaps, quite feliticious" (Jespersen 1954:417). In NZE on the other hand, Miller's analysis suggests that it has a focusing function rather than a hedging one (cf. (109)).

(109) a. [D]on't make that face at me *like*. (ICE New Zealand:S1A-015$A)
b. [O]h we'll set it up *like*. (ICE New Zealand:S1A-088$A)
c. [H]e <,> this t yeah i just came up *like*. (ICE New Zealand:S1A-047$D)

In her extensive and detailed discourse-pragmatic analysis of clause-final LIKE in IrE, Columbus (2009) finds that Jespersen's (1954) analysis is inadequate to account for the majority of instances occurring in the Irish component of the ICE. In fact, Columbus (2009) concludes that clause-final LIKE functions as an emphatic focus marker, complementing Miller's (2009) analysis of similar uses in NZE and Miller and Weinert's (1995) account of clause-final LIKE in SctE. In other words, the implied difference between regionally bound discourse functions appears to be attributable to the

superficial study of this vernacular form by Jespersen (1954) rather than to an actual functional dissonance.

Table 122: Age and gender distribution of clause-final LIKE in NZE

AGE	MALE		FEMALE			
	LIKE (N)	LIKE (mean)	LIKE (N)	LIKE (mean)	SUM (N)	ALL (mean)
1 (16–19)	1	0.10	4	0.16	5	0.14
2 (20–19)	4	0.15	10	0.13	14	0.14
3 (30-39)	0	0.00	0	0.00	0	0.00
4 (40+)	1	0.03	0	0.00	1	0.01
SUM	6	0.09	14	0.09	20	0.09

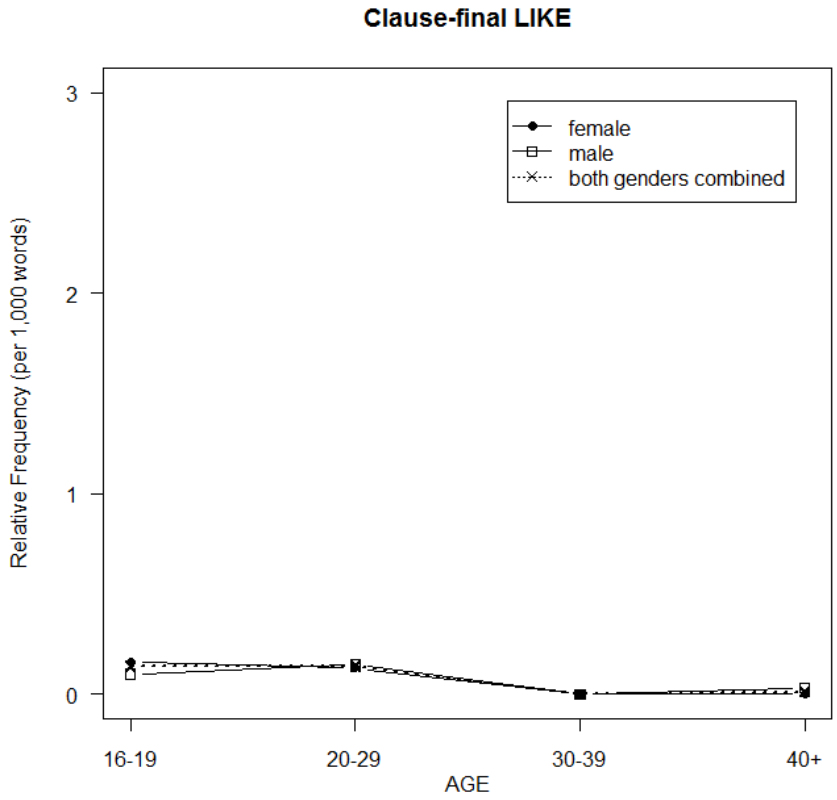

Figure 63: Age and gender distribution of clause-final LIKE in NZE

As clause-final LIKE is relatively scarce compared with clause-initial and clause-medial uses of LIKE (Miller 2009:334), it is not surprising that the regression model used in the analysis of cause-final LIKE failed to outperform

the baseline model. Hence, more robust χ^2-tests are used to evaluate meaningful correlations between the predictor variables and the frequency of clause-final LIKE. The χ^2-statistics confirm a significant correlation between the use of clause-final LIKE and age, showing that older speakers use clause-final LIKE less than expected. In other words, younger speakers use it more often than expected when age did not significantly affect the use of this type of LIKE (A1 v A3: χ^2= 9.79, df= 1, p<.001***, Cramér's φ=0.01, A2v A3: χ^2= 34.94, df= 1, p<.001***, Cramér's φ=0.01). Gender, on the other hand, does not have a significant impact (χ^2= 0.505, df= 1, p=.43).

7.7.1.4 Non-clausal LIKE

This section deals with the syntactically unbound, non-clausal uses of LIKE, as in (110), which were not included in Miller's (2009) analysis.

(110) a. [I] i i give credit for um *like* for for *like* the samoan debaters even though we came second. (ICE New Zealand:S1A-058$T)

b. [W]hat i was going to do was photocopy it on colour paper with *like* um… (ICE New Zealand:S1A-096$A)

c. [T]hey *like* well i mean they might not know too much about it. (ICE New Zealand:S1A-088$B)

As in other regional varieties, syntactically unbound instances of LIKE serve as floor-holding devices and, thus, function primarily as processing time buyers. The use of non-clausal LIKE thus conforms closely to functionally and syntactically similar instances in other regional varieties.

The next section dedicates itself to the sociolinguistic distribution of LIKE in NZE.

Table 123: Age and gender distribution of non-clausal LIKE in NZE

AGE	MALE		FEMALE			
	LIKE (N)	LIKE (mean)	LIKE (N)	LIKE (mean)	SUM (N)	ALL (mean)
1 (16–19)	9	0.71	19	0.77	28	0.75
2 (20–19)	22	0.51	49	0.67	71	0.61
3 (30-39)	3	0.15	5	0.22	8	0.19
4 (40+)	1	0.04	7	0.19	8	0.13
SUM	35	0.36	80	0.52	115	0.46

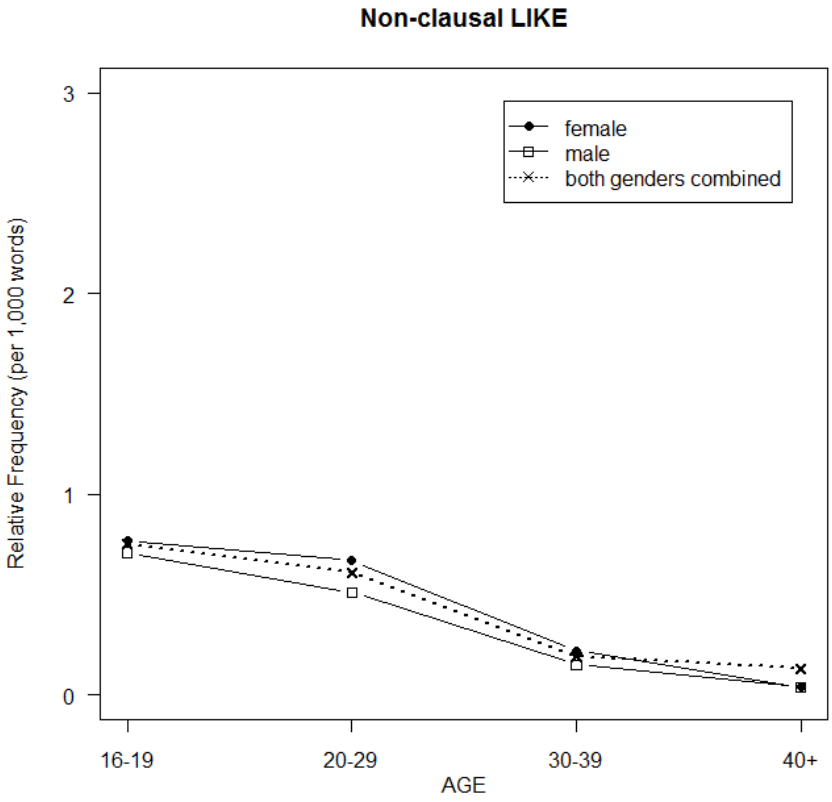

Figure 64: Age and gender distribution of non-clausal LIKE in NZE

As with both clause-initial and clause-medial LIKE, syntactically unbound uses of LIKE show a near-linear negative correlation with age in apparent-time. In contrast to clause-initial and clause-medial LIKE, the use of non-clausal LIKE is apparently not affected by gender, as the lines Figure 64 are almost identical. The following statistical analysis aims at confirming these tentative impressions.

Table 124: Results of the multivariate regression for non-clausal LIKE in NZE

NON	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.42	0.303	-4.70	<0.001***
A2	-0.12	0.288	1.45	.14
A3	-0.42	0.446	-0.95	.33
A4	-0.68	0.467	-1.46	.14
PAI	0.15	0.027	5.43	<.001***

The age stratification resembles the common monotonic decline, implying that the older a speakers is, the less likely it is that he or she uses non-clausal LIKE. The positive correlation between the use of non-clausal LIKE and the PAI index indicates that this variant of LIKE clusters in certain conversations. Moreover, the multivariate analysis confirms the impression derived from Figure 64, according to which the use of non-clausal LIKE is not gendered and indicates that it is not significantly socially stratified.

Unfortunately, the date of data compilation is not given in this ICE component, and the apparent-time analysis cannot be validated from an additional real-time perspective.

7.7.2 Summary: LIKE use in New Zealand English

The analysis of LIKE in NZE has shown that this non-standard feature is salient in contemporary NZE vernacular: nearly two thirds of the speakers in the ICE data used this non-standard feature at least once. Following Nevalainen & Raumolin-Brunberg (2003:55), this large quantity suggests that any changes LIKE may have undergone are nearing completion. The absence of social stratification in the case of clause-initial, clause-final, and non-clausal LIKE corroborates this interpretation. However, this depiction does not fully apply to clause-medial LIKE, as the evidence shows that clause-medial LIKE is both gendered and socially stratified. The complex social embedding of clause-medial LIKE implies ongoing change and that the use of this form has not yet fully diffused through all social strata. Interestingly, clause-medial LIKE is not more common in female but in male speech. A likely explanation for the male

lead in the present case is a combination of both overt stigmatization and a male dominated change.

This male lead calls into doubt claims which assert that, with regard to the spread of LIKE in "the New Zealand context, the pivotal role of women remains fundamental" (D'Arcy 2007:391). The crucial point here is why this male-dominated change does not disappear without a trace as in the majority of male-dominated changes (Labov 2001:462). One possibility relates to socio-cultural influx and the quantity of contact. According to Labov (2001:461-463), male-dominated changes typically vanish as the incoming form is not transmitted in sufficient quantities to children through the input of their primary care-takers. In the case of LIKE, this is probably not the most relevant factor. The ascendency of LIKE is more closely linked to associations of this form with its perceived reference group, i.e. US American adolescents (Buchstaller 2001). Hence, it is primarily not a matter of first language acquisition and thus generational change, but a matter of identity marking via linguistic means and thus a case of lexical diffusion best described as communal change. In light of these considerations, the male lead corroborates the assumption that during its implementation into the local system, LIKE itself – and very likely also the attitudes associated with it – have undergone re-interpretation, which triggered the emergence of distinct sociolinguistic profiles. This attitudinal-change hypothesis is appealing, as it can potentially account for male instead of female preferences or even the absence of systematic gender differences.

Taking this perspective also provides a handle on the stigmatization of this vernacular feature. The problem appears to be two-fold: younger speakers probably perceive this feature to be a progressive, anti-authoritarian, innovative feature of American vernacular English, while among older speakers, it is more likely to be recognized merely as a local non-standard feature on a par with clause-final LIKE. Adding weight to this view is that a similar development is described by Hickey (2003) with respect to the diffusion of certain phonological features in IrE. Hickey (2003:360) concludes

that the mass media, despite being commonly neglected as a medium of linguistic change, has substantially affected the dissemination of a late phonological change. Applying Hickey's (2003) argument to the spread of LIKE in NZE would mean that it is not the primary care-takers that matter most, but that it is sufficient exposure of the New Zealand youth to vernacular uses of clause-medial LIKE in AmE-dominated media that has played the decisive role. It is crucial to remember that the association of any linguistic form with a social reference group is initially arbitrary (Labov 2001:307), thus allowing for male-dominated changes. Only the cultural paradigm that primary care-takers are first and foremost females causes the vast majority of changes to be dominated by females rather than by males.

Whatever the real motivation for this change may be, the question remains as to why this change should be dominated by males and not by females. Figure 62 indicates that the most pronounced gender difference exists among speakers in their thirties. In fact, the female rate of LIKE use is almost negligible and mirrors the rate of speakers aged 40 and older. In contrast, males show a notable apparent-time increase in their use of clause-medial LIKE between the ages of 40 and 30 indexing that they were the first to adopt this incoming variant. Only later do the rates of males and females converge. The apparent-time distribution thus strongly suggests that the males adopted this form more readily than females, possibly due to its similarity with clause-final uses. Indeed, clause-final uses of LIKE have been attested in NZE for at least 60 or so years (D'Arcy 2007:401):

> [T]he marker is also attested in recordings of elderly speakers made by the New Zealand National Broadcasting Service in 1946–48. These data document the speech of native New Zealanders born in the period from 1851 to 1919. The vast majority of these speakers' parents had emigrated to New Zealand from England, Ireland, and Scotland; none had come from the United States.

The variety specificity of the NZE setting should, however, not be overstated because, despite the male lead, the distribution and functionality of LIKE conforms rather closely to the American usage pattern. What is remarkable here is that the similarity and stability of usage patterns challenge the assumption that exclusively quality face-to-face contact enables complete acquisition, while media exposure transmits only superficial aspects of the grammar of innovations. Indeed, the instances of clause-medial LIKE in New Zealand appear indistinguishable from equivalent uses in AmE. However, more fine-grained studies are required to confirm this functional similarity.

With respect to age, the present study unanimously confirms a negative, approximately linear correlation between the subtypes of LIKE and the age of speakers. This apparent-time distribution shows that younger speakers use LIKE significantly more than older speakers and implies a recess in LIKE use in apparent-time. Other than age, the extra-linguistic variables are almost negligible and fail to predict the frequency of overall LIKE use, as well as the clause-initial and non-clausal LIKE use better than chance.

Only clause-medial LIKE is sensitive to the socio-economic status of speakers, as those on the higher end of the socioeconomic hierarchy are less likely to use clause-medial LIKE than speakers on the lower end. There are at least two viable explanations for this finding. Firstly, this social stratification may index that clause-medial LIKE has not as yet diffused through all social strata. Secondly LIKE may have undergone a transformation which triggered its use as a social index. The first assumption is readily testable, as it predicts that the correlation with the social status of speakers will wane as its diffusion through society continues. The second hypothesis is less readily testable, but corroborated by the fact that the male lead is already suggestive of the fact that the associations attached to LIKE have undergone modification, allowing it to gain new, variety-specific connotations.

7.8 LIKE in Filipino English[72]

Filipino or Philippine English (PhiE) is an interesting variety to look at with respect to LIKE use, as it is "not a product of British but of American colonial expansion" (Schneider 2007:140). Although this colonial history is limited to the twentieth century, its linguistic impact has been intense "with the language having spread very rapidly" (Schneider 2007:140). Nevertheless, PhiE "tends to be monostylistic and close to the written norms" (Schneider 2007:141; Gonzalez 2004:12), while in informal settings speakers prefer Tagalog or the mixed variety called Taglish, i.e. a blend of *Tagalog* and *English* (Schneider 2007:141; Gonzalez 2004:12). The question at hand is whether its colonial history is reflected in the vernacular use of LIKE, i.e. whether LIKE use in PhiE follows the American pattern, which uses high frequencies of clause-initial and clause-medial LIKE. Despite the assertion that PhiE is rather close to the written standard and not the language of choice in intimate, informal conversation, the survey of LIKE use across varieties of English showed that LIKE use in PhiE follows the American pattern with remarkable similarity. The following analysis will probe more deeply into this assertion and evaluate the degree to which this analogy holds.

As before, we will have a look at the basic statistics in the data to get a first impression of its structure before analyzing it in greater detail. Of the 200 speakers present in the data, 121 make use of LIKE at least once, i.e. over sixty percent of the speakers of PhiE used this marker in the present data. The rather large proportion of speakers using this pragmatic marker shows that LIKE is common in PhiE.

The data summarized in Table 116 provide the basic statistics of the Philippine data and show a first rough impression of LIKE use in this regional variety.

[72] For a complete overview of the final data set of PhiE used in this study cf. Table 145 in the Appendix.

Table 125: Overview of LIKE use in PhiE

	N (total)	Mean (per 1,000 words)	Median (per 1,000 words)	Standard deviation (s)
LIKE	452	2.228	1.143	3.006

According to Table 125, LIKE is notably frequent in PhiE, with a mean slightly above two instances per 1,000 words and a median above one instance per 1,000 words. Although this difference is larger than the difference between mean and median in NZE, this rather moderate difference indicates that LIKE is not only widely used in this regional variety, but also that it is distributed rather homogenously among speakers of PhiE.

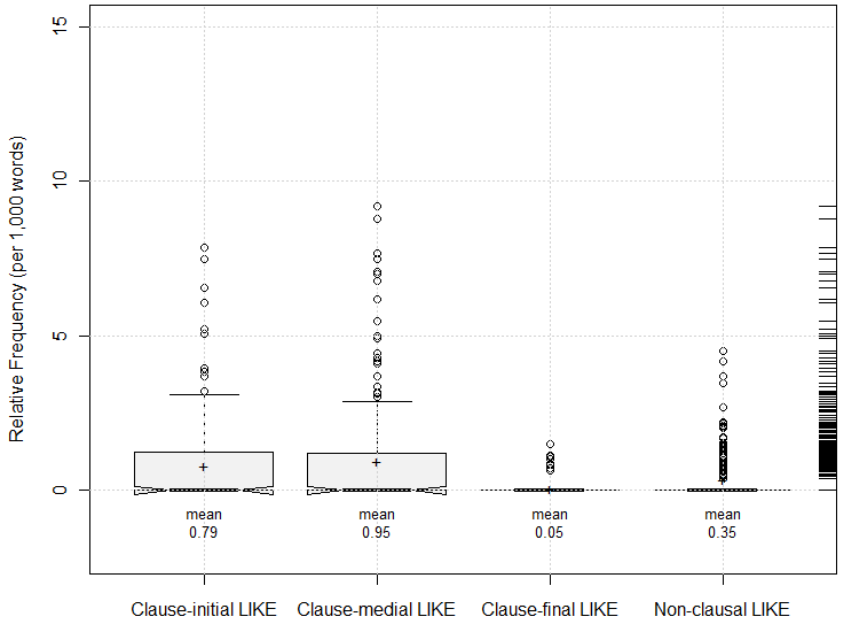

Figure 65: Rate of LIKE variants in PhiE

With the exception of a significantly lower frequency of non-clausal LIKE, the distribution pattern of LIKE variants in PhiE mirrors the distribution of NZE. The two most frequently used variants are doubtlessly clause-initial and clause-medial LIKE, while the use of clause-final and non-clausal LIKE is mostly negligible.

Table 126: Age and gender distribution of LIKE in PhiE[73]

			MALE		FEMALE			
AGE	**Speakers**	**Words**	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (16 – 20)	79	76,816	25	2.19	204	2.72	229	2.65
2 (21 – 30)	76	69,255	65	2.34	103	2.45	168	2.41
3 (31 – 40)	25	31,707	10	0.84	35	1.86	45	1.49
4 (41+)	15	11,897	0	0.00	2	0.41	2	0.22
NA	3	3,402	--	--	--	--	8	2.78
SUM	198	193,077	100	1.75	352	2.42	452	2.23

Table 126 reports a substantial female lead across all age groups and a negative correlation between age and LIKE use, as speakers above age 41 seem not to have acquired this pragmatic marker at all. This indicates that LIKE is highly sensitive to age in PhiE. To evaluate the female lead, Figure 66 shows LIKE use with respect to the age and gender of speakers.

[73] Speakers for whom age and gender information is not available are classified as NA.

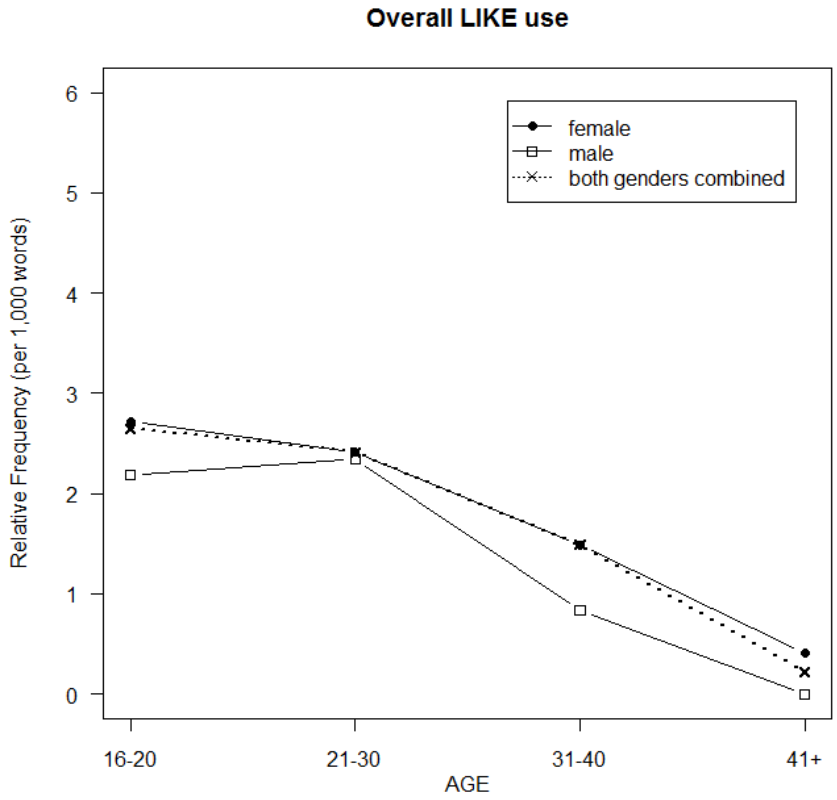

Figure 66: Age and gender distribution of LIKE in PhiE

According to Figure 66, the gender of speakers is probably insignificant, as all confidence intervals of parallel age groups overlap. Moreover, LIKE seems to have entered the speech of females during the 1980s and has stabilized in terms of frequency, as the curve in Figure 66 appears to level. Male speakers, on the other hand, followed an initial female lead, adopting LIKE slightly later. As Figure 66 indicates the absence of a real gender difference, male and female speakers are collapsed in Figure 67 to provide a maximally informative depiction of the age-grading in LIKE use.

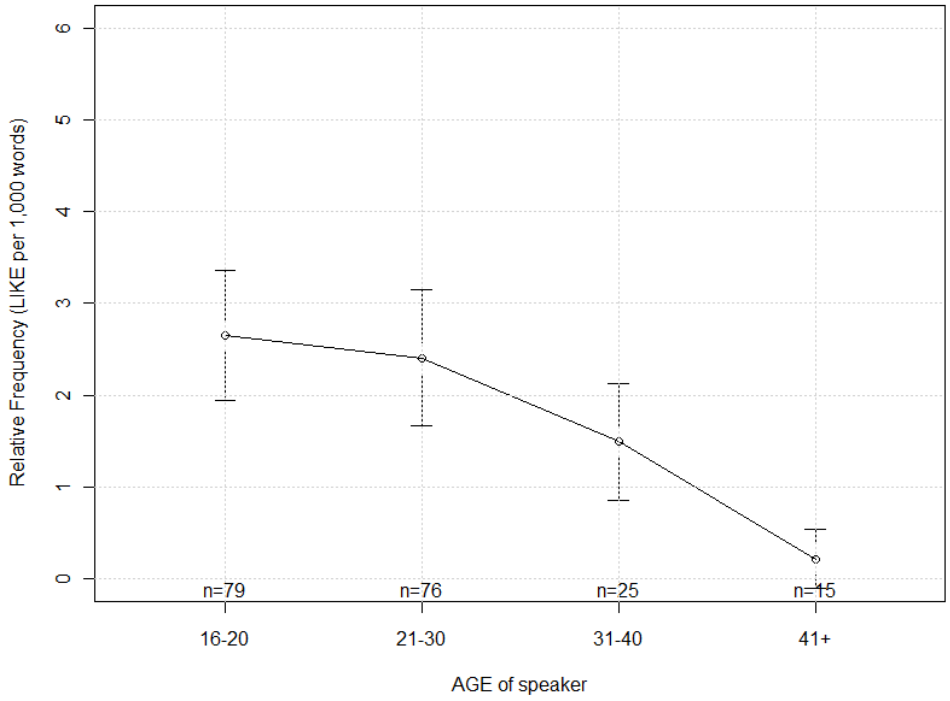

Figure 67: Age and gender distribution of LIKE in PhiE

The distribution of LIKE use across age groups in Figure 67 suggests that LIKE use may have rapidly droped by speakers twenty years ago, i.e. speakers now aged 31 to 40. In the more recent past, the speed with which this incoming form has been adopted decreased, leading to a shallower slope among the younger cohorts. Before proceeding with the analysis of other social factors, we will briefly look at the age distributions of the subtypes of LIKE to inform ourselves about whether distinct types of LIKE exhibit different sociolinguistic profiles with respect to age. According to the pattern emerging from Figure 67, it may be hypothesized that this trend in age-grading is not uniform across all variants of LIKE, but that clause-medial LIKE is more heavily age-graded than clause-initial LIKE.

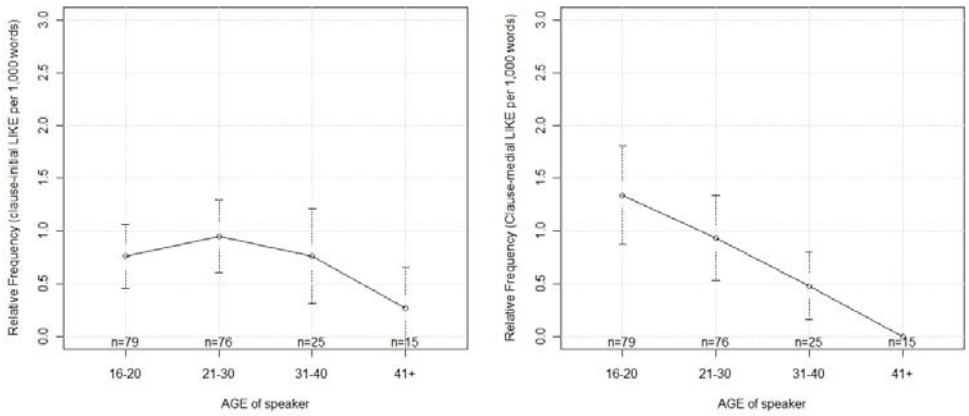

Figure 68: Age and gender distribution of clause-initial and clause-medial LIKE

Figure 68 corroborates the notion that the degree and pattern of age-grading is type-specific, as clause-initial LIKE is only mildly affected by the age of speakers, while clause-medial LIKE exhibits an almost linear negative correlation with increasing age. The fact that these two forms differ with respect to their degree and pattern of age-grading is suggestive of ongoing change, but again, this pattern may also reflect stable age-grading. In order to get a more detailed impression, the LIKE use of speakers with different occupations is plotted below to inspect possible correlations between social and linguistic stratification as additional indicators of ongoing change.

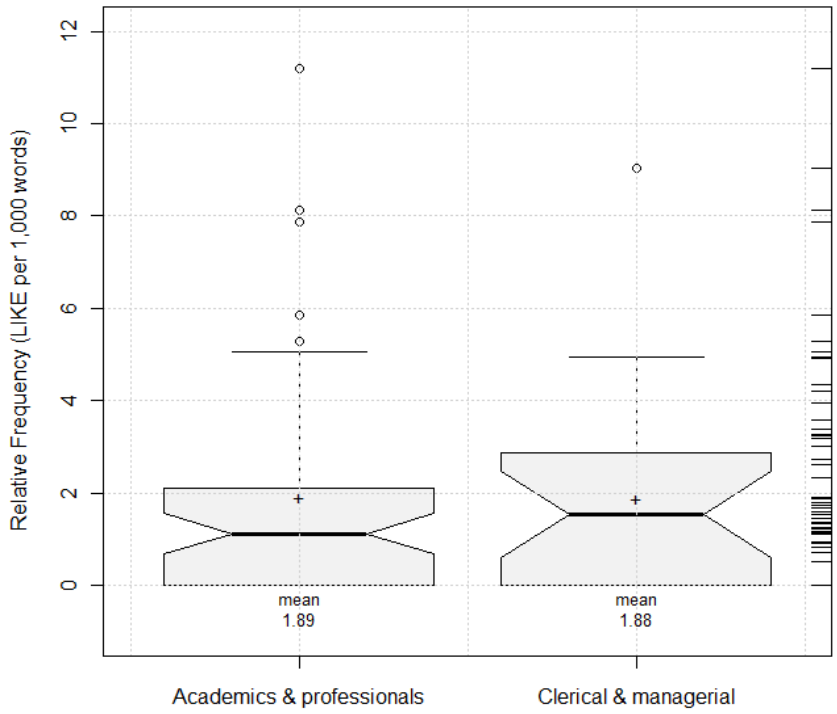

Figure 69: Frequency of LIKE in PhiE with respect to occupation

Unfortunately, the Philippine data does not contain speakers falling into the category "skilled and unskilled manual labor", which limits the validity of occupation as an explanatory parameter. Figure 69 alludes to an absence of social stratification of LIKE use.

7.8.1 Statistical analysis of LIKE usage in New Zealand English

The following statistical analysis evaluates the intuitions derived from the summaries and graphical displays, and seeks to evaluate general tendencies of LIKE use in PhiE. As this statistical analysis is more powerful than mere visual inspections, it may help to uncover correlations which have previously escaped detection.

Poisson regressions are used to detect possible correlations between the frequencies of LIKE and the age (A1 = 16 to 20, A2 = 21 to 30, A3 = 31 to 40, A4

= 41 years and older), gender (SEX), occupation (OCC) and the first language (L1) of speakers, as well as interactions between age and gender. As before, the initial saturated models are fitted in a step-wise procedure and subsequently arrive at a minimal adequate model that contains only predictors whose coefficients are significant. The explanatory power of final minimal adequate models is tested by comparing their predictive power to baseline models containing only the intercept as predictor variable. The same procedure is applied to study the effect of age, gender, and occupation on the frequency of clause-initial LIKE, clause-medial LIKE, and syntactically unbound LIKE without scope, i.e. non-clausal LIKE.

Table 127: Results of the multivariate regression for LIKE in PhiE

ALL	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	0.97	0.133	7.28	<.001***
A2	-0.09	0.203	-0.46	.64
A3	-0.55	0.238	-2.32	<.05*
A4	-2.29	0.671	-3.41	<.001***

The final minimal adequate model is statistically significant (χ^2= 14.376, df= 1, p<.01**), and reports negative correlations between the frequency of LIKE and age.

The output thus confirms the existence of age-grading and the increasing negative values of the coefficients imply a negative near-linear correlation between LIKE use and age. This implies not only the absence of interactions between the age and the sex of speakers; it also means that none of the predictor variables other than age are able to predict the overall frequency of LIKE use better than chance.

7.8.1.1 Clause-initial LIKE

Although clause-initial LIKE, as in (111), is common in PhiE, the multivariate analysis fails to detect meaningful correlations between the predictor variables and the frequencies of this subtype of LIKE.

(111) a. *Like* my uncle who's an American himself didn't like a beach in North Carolina I guess. (ICE Philippines:S1A-022$B)

b. *Like* how much is it? (ICE Philippines:S1A-041$B)

c. *Like* if you're a fourth year can you just like walk through a crowd of third years and everybody would like step aside. (ICE Philippines:S1A-029$B)

Table 128: Age and gender distribution of clausal-initial LIKE in PhiE[74]

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (16–20)	9	0.58	54	0.78	63	0.75
2 (21–30)	23	0.78	44	1.00	67	0.91
3 (31-40)	4	0.35	15	0.95	19	0.73
4 (41+)	0	0.00	2	0.40	2	0.21
SUM	36	0.57	115	0.85	151	0.77

The relative frequencies in Table 128 and Figure 70 clearly indicate age stratification, with a peak among speakers 21 to 30 and a consistent gender difference with a female lead. The age-stratification disagrees with the general trend of a monotonic declining pattern. Indeed, the distribution follows a curvilinear pattern, as the use of LIKE is most salient among speakers in their twenties and thirties while there is a notable recess among teenagers.

[74] Speakers for whom age and gender information is not available are not included.

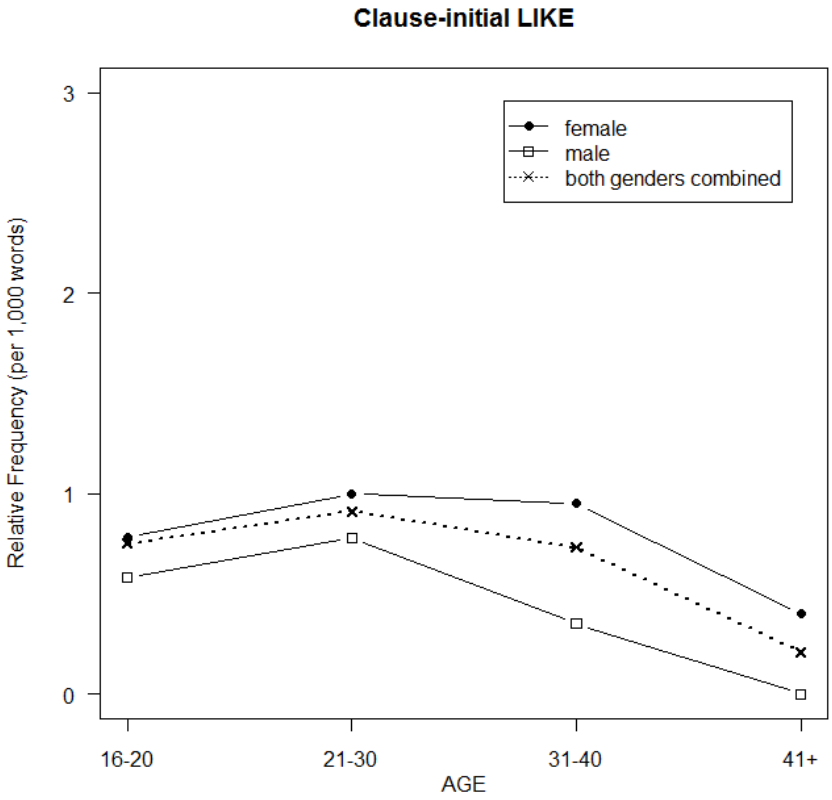

Figure 70: Age and gender distribution of clause-initial LIKE in PhiE

Although this variant of LIKE is widespread in PhiE, the multivariate analysis fails to detect meaningful correlations between the predictor variables and the frequencies of this subtype of LIKE. Hence, none of the extra-linguistic variables correlate significantly with the rate of LIKE use. Despite not being statistically validated, Figure 70 suggests a consistent female bias across all age groups. To probe more deeply into this aspect, this gender difference is re-evaluated using χ^2-tests.

Table 129: χ2-test results: the effect of gender on clause-medial LIKE within age groups in PhiE

AGE	χ^2	df	p-value	φ
1 (16 – 20)	0.065	1	.79	0.00
2 (21 – 30)	1.392	1	.23	0.00
3 (31 – 40)	2.020	1	.15	0.00
ALL	5.019	1	<.05*	0.00

Although the χ^2-tests fail to detect a gender difference within each age group, they are able to confirm an overall female lead, which nevertheless has only an extremely weak effect. Hence, clause-medial LIKE is used significantly more often by female speakers of PhiE. What is especially noteworthy is the curvilinear age stratification. This pattern suggests not an increase in LIKE use but, the opposite: it appears that LIKE use is receding significantly among speakers below the age of 20. The question arises as to why LIKE use is waning among these young speakers.

7.8.1.2 Clause-medial LIKE

In PhiE, clause-medial LIKE, as in (112), is the most frequent variant of LIKE with respect to its rate, but not with respect to the speakers who use it: of the 198 speakers in the present data, 73 used this form, while 79 used clause-initial LIKE.

(112) a. They were *like* so stupid and mindless. (ICE Philippines:S1A-041$B)
b. Uh uhm yeah unlike for us we have to *like* wait uh-huh. (ICE Philippines:S1A-013$A)
c. According to my sister the the scariest part of the movie was *like* in the beginning when the little boy entered the kitchen. (ICE Philippines:S1A-012$A)

Despite being equally common, their distributional profiles differ notably: while clause-initial LIKE appeared to wane among teenagers, this is certainly not the situation we observe for clause-medial LIKE.

Table 130: Age and gender distribution of clausal-medial LIKE in PhiE[75]

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (16–20)	10	0.99	111	1.40	121	1.35
2 (21–30)	23	0.72	41	1.06	64	0.93
3 (31-40)	3	0.27	10	0.48	13	0.40
4 (41+)	0	0.00	0	0.00	0	0.00
SUM	36	0.61	162	1.10	198	0.96

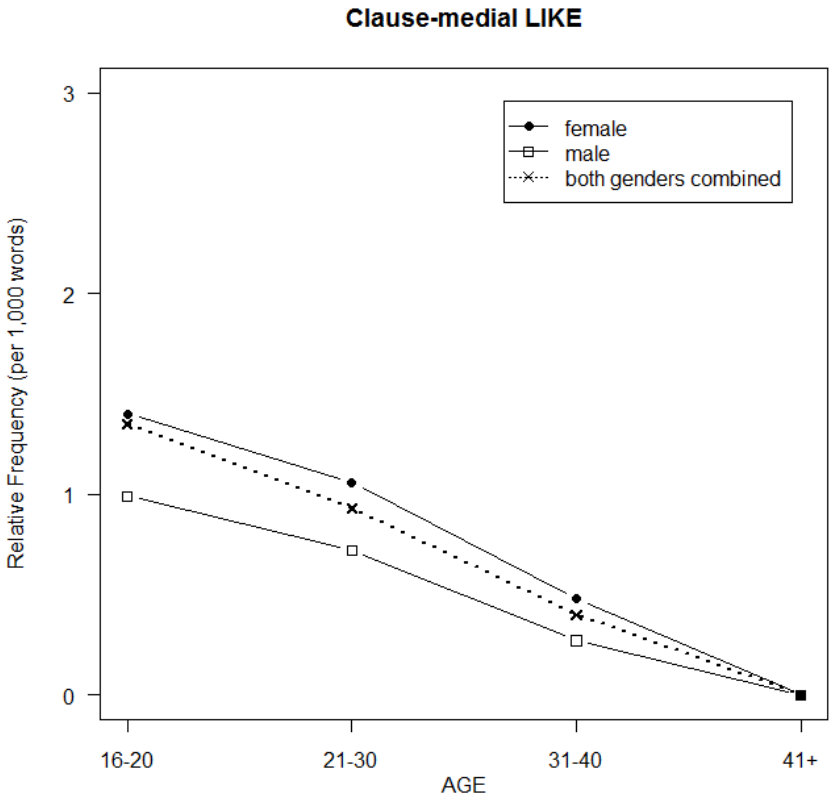

Figure 71: Age and gender distribution of clause-medial LIKE in PhiE

Both Table 130 and Figure 71 indicate the use of clause-medial LIKE complies with the typical recess of LIKE use with increasing age. Indeed, this recess is near linear and accompanied by an increasing gender difference, which is most

[75] Speakers for whom age and gender information is not available are not included.

pronounced among the youngest speakers and diminishes towards older cohorts.

Table 131: Results of the multivariate regression for clause-medial LIKE in PhiE

MED	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.81	0.209	-3.87	<.001***
SEX: Male	-0.29	0.263	-1.10	.26
A2	-0.27	0.239	-1.14	.25
A3	-0.45	0.320	-1.43	.15
A4	-16.42	0.331	-49.543	<.001***
PAI	0.14	0.013	10.414	<.001***

Although the multivariate analysis confirms a gender difference in the expected direction, its effect size is too weak to breach the common five percent level of significance. The age stratification is also confirmed by the increasing negative values of the coefficients. The extreme value reported for speakers above age 41, i.e. age group 4, results from the fact that clause-medial LIKE is practically non-existent among these speakers. What seems to have happened is that this variant has only recently entered this speech community so that speakers over age 42 have not been sufficiently exposed to this variant of LIKE and have, hence, not acquired its use.

Returning to age stratification, one can observe that clause-initial and clause-medial LIKE follow distinct patterns. As has been argued in the previous section, clause-initial LIKE is receding among teenagers while the opposite is true of the clause-medial form. The latter indeed is peaking among speakers between 16 to 20 years of age. First of all, the opposing trends show that LIKE is not generally waning in PhiE, but that the decrease is confined to the clause-initial form. Secondly, not only do the age patterns reflect opposing tendencies, but the degree of the gender difference does also: while steadily increasing with age in the case of clause-initial LIKE, the opposite holds true for clause-medial LIKE. Indeed, the gender difference is most salient among teenagers with respect to the latter form.

7.8.1.3 Clause-final LIKE

Due to the low number of instances, it is futile to run a regression model for clause-final uses of LIKE as in (113).

(113) a. There was one time that you told me you missed kids *like.* (ICE Philippines:S1A-066#B)
b. Do you have do you have uniforms there *like.* (ICE Philippines:S1A-029#A)
c. Yeah but I think it's more <,> it's more of a case to case basis *like.* (ICE Philippines:S1A-040#B)

There are a total of 10 instances of clause-final LIKE in the Philippines data, all uttered by different speakers, which indicates that the use of this variant is not an idiosyncracy of one or two individuals. Of the ten speakers who used the clause-final type of LIKE, eight are female and only two male, but all are below age 34. Nine of these ten speakers also used other variants.

χ^2-tests do, however, confirm a significant correlation between the use of clause-final LIKE and age, as older speakers use clause-final LIKE less than expected, i.e. younger speakers use it more often than expected when age did not significantly affect the use of this type of LIKE (A1 v A3: χ^2= 9.79, df= 1, p<.001***, Cramér's φ=0.01, A2v A3: χ^2= 34.94, df= 1, p<.001***, Cramér's φ=0.01).

Based on this distribution, it seems that a number of younger speakers of PhiE have adopted LIKE into their linguistic repertoire, but failed to differentiate between clause-medial LIKE and the clause-final form. Though this is speculative, the scenario is not altogether implausible given that the Philippines have socio-cultural influx from both the US and Australia. The linguistic input thus contains both variants in sufficient proportions to trigger their adoption in this second-language variety.

7.8.1.4 Non-clausal LIKE

Despite is low mean frequency, non-clausal LIKE (cf. (114)) is not uncommon in this variety: nearly one quarter of the speakers in the ICE data, i.e. 48 of 198, used it.

(114) a. In high school we *like* <,> we sit around watch T V second. (ICE Philippines:S1A-033$B)

b. Uhm *like* he would uh he would promise *like* he would you know get up on certain days…(ICE Philippines:S1A-010$B)

c. Agh <,> *like* <,> yeah don't you think it would be better if it was like that. (ICE Philippines:S1A-048$A)

Table 132: Age and gender distribution of non-clausal LIKE in PhiE[76]

	MALE		FEMALE			
AGE	**LIKE (N)**	**LIKE (mean)**	**LIKE (N)**	**LIKE (mean)**	**SUM (N)**	**ALL (mean)**
1 (16–20)	4	0.45	30	0.40	34	0.41
2 (21–30)	17	0.58	12	0.22	29	0.36
3 (31-40)	3	0.22	9	0.37	12	0.31
4 (41+)	0	0.00	0	0.00	0	0.00
SUM	24	0.42	51	0.31	75	0.34

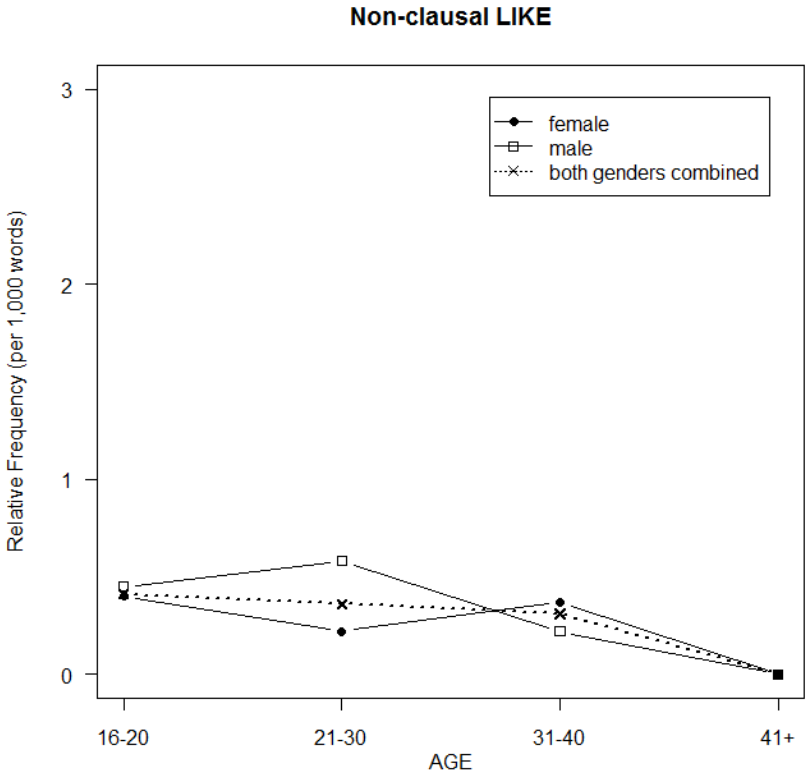

Figure 72: Age and gender distribution of non-clausal LIKE in PhiE

[76] Speakers for whom age and gender information is not available are not included.

Figure 72 suggests a slight decrease in LIKE use with increasing age. Nonetheless, there is a notable recess in LIKE use between speakers in their thirties and speakers older than 41. The gender of speakers does not systematically affect LIKE use, as the difference between males and females is not only minute but also lacks consistency.

Table 133: Results of the multivariate regression for non-clausal LIKE in PhiE

NON	Estimate (coefficient)	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.77	0.253	-7.01	<.001***
A2	0.23	0.335	-0.68	.49
A3	0.14	0.379	0.37	.70
A4	-15.59	0.357	-43.64	<.001***
PAI	0.12	0.020	6.11	<.001***

The multivariate analysis indicates a significant difference between the use of LIKE by teenagers and speakers above the age of 41. Moreover, the regression model does not report a monotonic recess of LIKE use with increasing age. Thus, the trend observable in Figure 72 is unreliable and may well be a result of chance rather than systematic differences among these age groups. Concerning gender, the statistical model corroborates the impression that there is no significant difference between the genders. With exception of the PAI index, none of the other predictor variables are significantly correlated with the frequencies of this subtype of LIKE.

7.8.2 Evaluation of the apparent-time construct

As in the preceding analyses, I will evaluate the apparent-time hypothesis by examining the use of LIKE in real-time. The most meaningful way to verify whether age-grading in apparent-time reflects ongoing change in real-time is, of course, to examine, whether the frequencies of LIKE differ with respect to the date of data collection (cf. Figure 73).

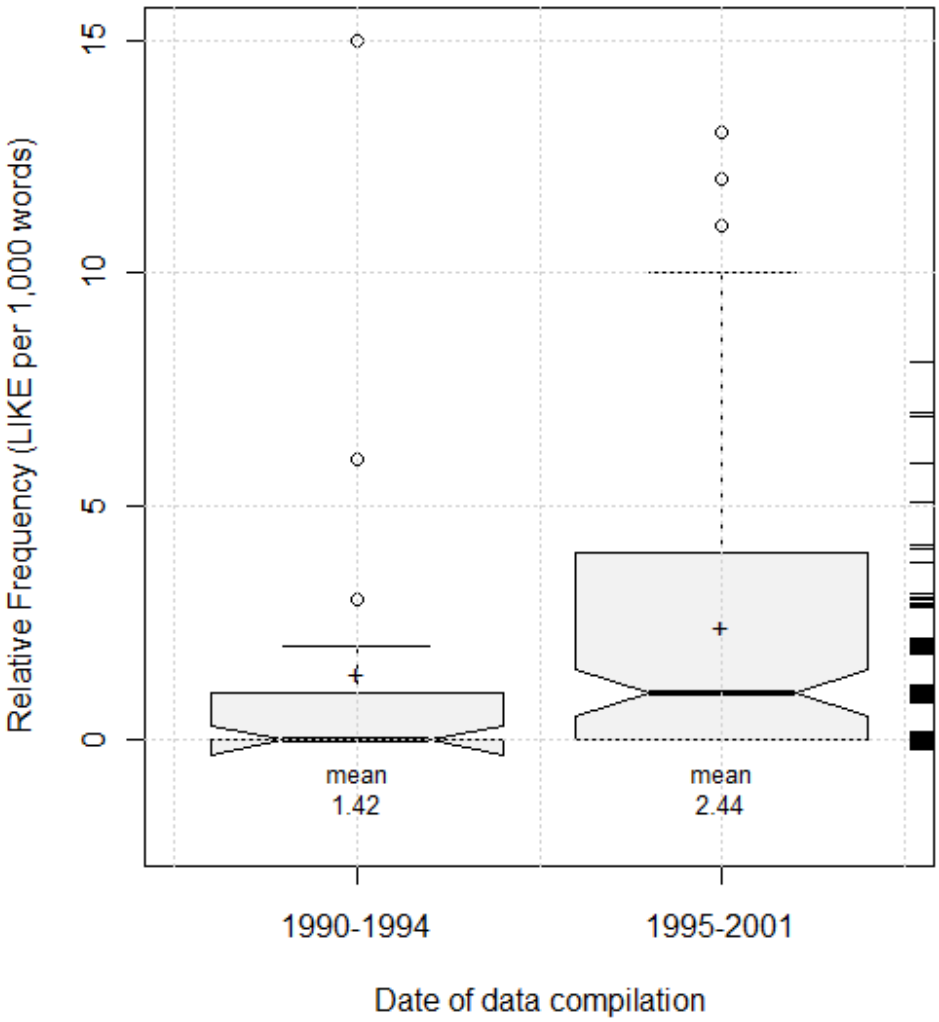

Figure 73: Frequency of LIKE in PhiE with respect to the date of data compilation

The results displayed in Figure 73 corroborate the assumption that LIKE is undergoing change in real-time even though the span of time is merely ten or so years. Unfortunately, the data compiled from 2002 to 2005 had to be excluded, since the small numbers of speakers rendered it unreliable. Indeed, this data compiled between 2002 and 2005 represents a sample of only nine speakers. Including the results of the latest period would skew the overall tendency and lead to understating the actual rate of change in LIKE use. To get an additional perspective on the data, the proportion of speakers who have

used LIKE during the different stages of data compilation are displayed in Table 134 below.

Table 134: Use of clause-medial LIKE with respect to the date of data compilation

	LIKE users	**Non-LIKE users**	**Ratio**	**LIKE users of all speakers (%)**
D1	10	16	0.62	38.4
D2	106	57	1.85	65.0

The pattern which emerges substantiates the adequacy of the apparent-time hypothesis as Table134 strongly suggests ongoing change. Thus, Table 134 corroborates the hypothesis that LIKE is increasing in frequency over time and thus implies that LIKE is currently spreading in PhiE. The problematic issue is whether this increase is a general trend which applies to all forms of LIKE or whether it is confined to certain uses.

Table 135: χ2-test results: the effect of the date of data compilation on the use of types of LIKE

Type	**χ2**	**df**	**p-value**
$INI_{D1vs\ D2}$	1.826	1	.17
$MED_{D1vs\ D2}$	21.812	1	<.001***
$NON_{D1vs\ D2}$	4.874	1	<.05*

The χ^2-statistics confirm a correlation between the date of data compilation and the occurrence of clause-medial and non-clausal LIKE. Clause-initial LIKE, however, does not significantly correlate with the date of data compilation. In other words, the most notable increase is observable with respect to clause-medial LIKE and to a lesser but still significant degree to non-clausal LIKE. The difference between functionally distinct types of LIKE strongly suggests that the increase is more or less confined to clause-medial LIKE. Given the fact that this increase in the use of this variant of LIKE appears to be a global trend, there is reason to belief that the slight increase in the rate of non-clausal LIKE is a subsequent effect resulting from the inability of speakers to differentiate between these two forms. Hence, one could maintain that the increase in non-clausal LIKE is dependent on the success of clause-medial LIKE and would not have taken place independently.

Since the increase may be confined to certain age cohorts as in IrE, we need to take age into consideration (cf. Table 136). The data are, hence, split into different age groups, and each age group is analyzed in isolation. Subsequently, the values of the data that have been compiled earlier are compared with the values of data compiled later. The values are tested for significance using non-parametric t-tests. The oldest age group had to be excluded, as they were distributed too unevenly to provide reliable results.

Table 136: LIKE use with respect to AGE and the date of data compilation[77]

	A1 (16-20)	**A2 (21-30)**	**A3 (31-40)**
All variants combined	-3.129***	n.s.	-3.037**
Clause-initial LIKE	n.s.	n.s.	-2.138*
Clause-medial LIKE	n.s.	n.s.	-2.449*
Clause-final LIKE	-2.304*	-2.180*	N.A.
Non-clausal LIKE	4.514***	-3.733***	n.s.

The fact that the t-tests report insignificant values for all age groups confirms change in real-time. In addition, the negative t-values confirm significant increases and thus corroborate that the trajectory of this change points in the expected direction. Only non-clausal LIKE has decreased among post-adolescents but shows a notable increase among speakers in their twenties. Contrary to the Irish setting, the real-time change is not confined to younger cohorts but affects all age groups. In fact, speakers in their thirties exhibit a notable increase in the use of both clause-initial and clause-medial LIKE while speakers in their twenties show an increase in clause-final and non-clausal LIKE use. Post-adolescents do not show a coherent trend, as they exhibit an increase in the use of non-clausal LIKE, but a decrease in clause-final LIKE while the rates of neither clause-initial nor clause final LIKE change significantly.

[77] N.A.= not available. The data of age group 2 has been compiled exclusively from 1995-2000.

These observations are startling as they indicate that speakers beyond adolescence in particular tend to accept LIKE into their grammar, while the rates of post-adolescents and young adults do not increase considerably. The acceptance of LIKE, particularly by speakers in their thirties, points towards a change in perception. While LIKE commonly indexes adolescence or youth, this association appears to be waning among PhiE speakers, suggesting variety-specific attitudinal re-interpretation.

7.8.3 Summary: LIKE use in Filipino English

The present analysis shows that LIKE has become a remarkably common feature in PhiE, as both its frequency and the proportion of speakers using it have nearly doubled in a short ten-year period. As in most varieties, the age of speakers is the strongest predictor for this non-standard feature, suggesting that LIKE use is most salient among adolescents and young adults. This profile is suggestive of age-grading, which would imply that speakers adapt their rate of LIKE as they mature. The age stratification of each form of LIKE varies to a certain extent, but younger speakers consistently show higher rates than older speakers. The patterning of the age distribution is not uniform across all variants, which implies that the consistency and the effect size of age is type-specific. Clause-initial LIKE, for instance, is only mildly affected by age and follows a curvilinear pattern, while clause-medial LIKE shows an almost linear negative correlation and thus a monotonic recess with increasing age.

With the exception of gender, which has a significant albeit small effect on the use of clause-initial LIKE, none of the extra-linguistic variables are significantly correlated with vernacular uses of LIKE. The lack of such correlations suggests an absence of social stratification and, hence, indicates a rather homogenous distribution. This is, however, not quite accurate. It should be remembered that clause-medial LIKE in particular has significantly increased over a very short ten-year period. While clause-medial LIKE follows a monotonic distribution which points to a recess in use with age, it has not

significantly increased among adolescents and post- adolescents in real-time. In fact, Table 136 draws attention to the fact that clause-medial LIKE has recently entered the speech of PhiE speakers who are in their thirties. The confinement of this increase to older cohorts suggests that social attitudes towards LIKE have been undergoing re-evaluation. During this re-evaluation, previous associations are overridden, allowing LIKE to percolate into formerly constrained environments. In view of these considerations, it is likely that the re-interpretation of social attitudes is a key factor in LIKE's recent increase. The resulting process of partial communal change, nonetheless, alone fails to fully account for the extent of the increase observable in the Philippine data. Hence, additional processes must be taken into account to interpret this rather dramatic change in real-time. The age distributions indicate that we are, in addition to communal change, dealing with a combination of age-grading and generational change. According to this view, the increase in LIKE use stems from an inadequate adaptation of younger speakers to the communal grammar of their elders. In other words, younger speakers, though adapting their rate of LIKE as they mature still maintain a level of use slightly higher than the rate of older peers. The higher rates of LIKE are subsequently dragged along as speakers mature.

In many, though not all aspects, the use of LIKE in PhiE is remarkably similar to its use in AmE. The most salient similarity between these two varieties is the bias towards clause-initial and clause-medial LIKE, while clause-final and non-clausal LIKE are rather marginal phenomena. This finding is remarkable, given that PhiE has been described as rather formal and close to the written standard (Schneider 2007:140). On a related note, the assertion that PhiE targets EngE as the prestige variety taught at school does not match the results of the present analysis. In fact, the frequencies obtained for LIKE suggest that vernacular uses of LIKE have ingressed PhiE significantly earlier than EngE. The result of this is that PhiE either lacks or has lost certain socio-cultural constraints which are decisive in EngE, or that the exposure to AmE vernacular is even stronger than commonly presumed. The real-time analysis

has provided compelling evidence for the former. Indeed, not only as PhiE speakers seem to view LIKE as being nativized (and thus non-foreign): they seem to regard it primarily as a functional element rather than a social index. Whatever the case, the present analysis shows PhiE has quickly and forcefully adopted this vernacular feature alongside its pragmatic functions.

8 Global synopsis and discussion

The analysis of vernacular uses of LIKE in eight geographically non-continuous varieties of English has provided valuable insights which offer a refined understanding of how borrowings diffuse from one speech community to another and, more importantly, how linguistic innovations diffuse through culturally distinct speech communities. Indeed, the analysis depicts a complex interplay of supra-locally stable tendencies, such as the monotonic age stratification and variety-specific idiosyncrasies of LIKE use as, for instance, the degree and direction of gender differences. Although these findings have partly been discussed in the respective sections, a general discussion taking a broader view of the findings has still to be provided. The following chapter therefore systematically reviews, discusses, and interprets the findings of the previous chapters and contextualizes them in light of contemporary sociolinguistic theory. The discussion consists of four parts. The first part serves as an introduction, while the second part discusses the findings for each positonal variant separately. The third part is a general discussion and addresses the hypotheses one by one. Finally, the chapter closes with a discussion of additional findings that the study has unearthed but which are not directly associated with the evaluation of the hypotheses.

8.1 Introductory remarks: the discourse marker LIKE in selected varieties of English

In contrast to Underhill's claim that LIKE "may even now be becoming archaic" (Underhill 1988:234), the results of the present study confirm that it has doubtlessly conquered the English-speaking world. Indeed, it represents one of the "dramatic 'new' discourse-pragmatic markers in the English language which have gained considerable high-profile attention in recent years" (Tagliamonte 2005:1897). Contrary to popular lore, the analysis has provided ample evidence confirming that LIKE is a salient and highly functional element

in contemporary English vernacular. Its ubiquity in informal discourse confutes assertions that it is confined to the speech of adolescents, that it is a typically female feature, as popular opinion suggests (cf. the discussion in D'Arcy 2007).

The systematic analysis of LIKE use demonstrates that the overall frequency of this non-standard feature varies substantially between regional varieties ranging from 0.49 in EngE to 4.38 instances per 1,000 words in CanE (cf. Table 18). The low frequency of occurrences of LIKE in EngE is startling particularly as it suggests a pronounced resistance to adopting this vernacular form. The British setting indicates that LIKE was flatly rejected by EngE speakers until fairly recently, which complies with Britain's (2002:618) second scenario in cases of contact between global innovations and the local norms of a given speech community: outright rejection (cf. Buchstaller & D'Arcy 2009:2929). The British rejection of LIKE is quite atypical, as quite different trends are observable in AmE, CanE, IrE, JamE, NZE, PhiE where LIKE use is frequent and a salient feature of the respective spoken varieties (cf. also D'Arcy 2005, 2007; Dailey-O'Cain 2000; Schweinberger 2012; Miller 2009; Siegel 2002; Tagliamonte 2005). In the latter varieties, LIKE appears to be primarily – although not exclusively – an identity marker, particularly among teenagers and young adults.

The intra-varietal analyses demonstrate that a superficial survey focusing merely on the overall frequency of LIKE insufficiently captures the variety-specific usage patterns. Probing more deeply into the functional employment of this form has corroborated the functionality of vernacular uses of LIKE, thereby confirming previous analyses which challenge the hypothesis that LIKE is merely a meaningless or functionally empty realization of a single and homogenous underlying form and thus corroborates D'Arcy (2007:395) on this issue:

> To suggest that *like* is no more than a linguistic crutch, signaling hesitancy and a lack of fluency or articulation (e.g., Siegel 2002, 47; see also citations in

> Diamond 2000, 2 and Levey 2003, 24), trivializes the complex juxtaposition of functions performed by this lexeme in the spoken language (cf. Levey 2003). In recognizing that numerous functions of *like* are operative in vernacular usage, the myth of meaninglessness is simultaneously demystified.

In fact, studies which focus exclusively on the overall frequency of LIKE are misleading and insufficient to account for the observable systematic differences between regional varieties, which can only be detected when more fine-grained functional analyses are applied. A case in point is the difference between IrE and CanE. Despite the fact that IrE and CanE show almost identical per 1,000 word rates of vernacular uses of LIKE, the varieties differ dramatically: instances of LIKE in CanE are almost exclusively confined to clause-initial and clause-medial uses of LIKE, while IrE owes its high overall rate to a substantial amount of the so-called traditional clause-final form (cf. Table 18). There is an underlying systematicity to the cross-varietal distribution of LIKE. In all varieties, except for IndE and IrE, clause-initial and clause-medial occurrences make up the main share of all uses of LIKE, and the patterning of LIKE use therefore resembles the AmE pattern of LIKE use. This patterning corroborates previous research (Siemund, Maier & Schweinberger 2009), which suggests that LIKE use follows two rather distinct distributional patterns resembling the distinction between IrE and IndE on the one hand, and all other varieties on the other hand. However, the present analysis has probed more deeply into the matter, indicating that – as for instance the intermediate distributional profiles of NZE suggest – there are not two distinct groups, but rather a continuum on which regional varieties can be located.

The continuum on which the varieties fall can be regarded as one representing the influence of AmE on the use of LIKE in other varieties of English: one end of the continuum represents varieties which are more strongly influenced by AmE, showing a pattern which approximates the positional profile of LIKE in AmE or CanE, i.e. instances of LIKE are most commonly cases of clause-initial or clause-medial LIKE as in PhiE and JamE.

The other end of the continuum represents varieties showing a preference for clause-final (IrE) or non-clausal LIKE (IndE), i.e. a non-AmE-like pattern. The latter varieties exhibit a wider variety of position profiles, because LIKE use is not affected only by contact, as in varieties which LIKE has only recently entered. For instance, the difference between the positional profiles of IrE and CanE can be explained as follows: In IrE LIKE already has a long history and clause-final LIKE is a fully functional element in IrE which is not heavily age-sensitive. The positional profile of LIKE in IrE is thus a mixture of the new pattern – dominance of clause-initial and clause-medial LIKE – and the traditional Irish pattern, with a high frequency of clause-final LIKE. The situation in Canada is quite different: LIKE only recently entered CanE, and LIKE use is strongly influenced by AmE: thus it follows the AmE pattern. The positional profile of LIKE in IndE is neither Irish nor American, because the target variety of InE is EngE in which LIKE is very infrequent. The over-proportional use of non-clausal LIKE shows that in IndE, speakers use LIKE not at all, or only partly according to the AmE pattern, but have established their own way of using LIKE. The differences in the varieties are therefore considered to spring from historical sources (as in IrE where LIKE has a long-standing history), contact (with AmE) as in CanE, which generally seems to follow the AmE pattern (despite some deviations form the AmE pattern), and mechanisms of change (adaptation) in which different social groups adopt LIKE as an identity marker (IndE).

To adequately capture the processes and mechanisms underlying the use of non-standard occurrences of LIKE, it is necessary to review the specific variants of LIKE in isolation. Thus, before turning to the discussion of the major findings and an evaluation of the hypotheses, the following section will review the regional differences of the respective types of LIKE.

8.1.1 Clause-initial LIKE

The analysis of instances of clause-initial LIKE has substantiated previous analyses, which found that clause-initial LIKE introduces clausal specifications or elaborations of previous propositions or elements thereof (Miller 2009:330). Such instances of LIKE are a common and salient discourse element in all regional varieties analyzed in the present study. The order of regional varieties based on the frequency of clause-initial LIKE is similar to the ranking based on overall LIKE use (cf. Figure 13) – with only AmE falling two levels. On average, clause-initial instances of LIKE constitute roughly a third (~31 percent) of all vernacular uses of LIKE (cf. Table 20). Indeed, their share ranges from a little over 23 percent of all occurrences in AmE to over 40 percent in CanE (cf. Table 20). In addition to other differences, this disparity between CanE and AmE draws attention to the fact that despite their geographical proximity and having been regarded as one supra-regional variety, i.e. North American English, (cf. e.g. D'Arcy 2007), they exhibit some noteworthy differences in their usage patterns of LIKE. For instance, CanE also exhibits a higher degree of social stratification of LIKE use than AmE.

8.1.2 Clause-medial LIKE

Clause-medial LIKE, the so-called innovative discourse particle (cf. Andersen 2001; D'Arcy 2005 2007), is probably the most intriguing variant among vernacular uses of LIKE. Indeed, it is this form that has received most scholarly attention and is "possibly the most familiar" (Miller 2009:332) and prototypical form of discourse marker LIKE. As has become apparent in this study, clause-medial LIKE has, and still is, undergoing a rather dramatic change. Indeed, it appears as if the changes we observe among other forms of LIKE are secondary phenomena. The increase in the use of other variants of LIKE is best portrayed as a consequence of the rapid diffusion of clause-medial LIKE, which caused not only the frequency of clause-medial LIKE itself to increase dramatically, but also the frequencies of other, positionally distinct

uses of LIKE. In essence, this is a consequence of speakers failing to differentiate between clause-medial LIKE and related though distinct forms of LIKE. In this view, the global diffusion of the virulent clause-medial variant during the latter half of the twentieth century has triggered an increase in the frequency of all variants of LIKE.

Clause-medial LIKE is well attested in all varieties investigated in this study and, with the exception of IndE, IrE, and NZE, the most frequently used variant. Remarkably, it is not AmE in which clause-medial LIKE is most frequent, but CanE. This is particularly striking for two reasons. Firstly this form has been described as an American borrowing (Andersen 2000:287); and secondly, the belief that LIKE is a typical American or more specifically a Californian feature is very widespread, as it is not only part of popular lore (D'Arcy 2007), but it has also been propagated in the media (cf. e.g. Johnson 1998; Mehren 1999; Osmers-Gordon 2008) and in scholarly linguistics (cf. Andersen 2000:216).

8.1.3 Clause-final LIKE

Despite being attested almost exclusively in Celtic varieties of English (cf. Miller 2009), clause-final LIKE is present in all regional varieties analysed in the present study. Although it is generally used rather infrequently, clause-final LIKE is extremely common in IrE which supports the hypothesis that it is a typically 'northern' phenomenon (Miller & Weinert 1995:368). The present analysis therefore substantiates previous claims according to which this traditional element has been associated mainly with the British Isles. The ICE data are not fully congruent with D'Arcy's (2007:413) finding that the use of clause-final LIKE is confined to the oldest members of the Toronto English speech community: the ICE data not only show that 25 of 245 Canadian speakers used this form at least once, but also that it is not confined to speakers above the age of 60, being distributed rather evenly among age cohorts.

Although being attested in AmE (Schoroup 1982:47), the use of clause-final LIKE is negligible in this variety. The American data contained only a single

instance which draws attention to another difference between AmE and CanE regarding LIKE use.

With respect to its functional employment, Jespersen's claim that clause-final LIKE is used parenthetically "to modify the whole of one's statement, a word or phrase" (Jespersen 1954:417) requires qualification, because the ICE data show that it is predominantly used to focus preceding utterances or constituents (cf. also Columbus 2009). This implies that the clause-final variant is functionally equivalent to clause-medial LIKE, despite a difference in its directionality of scope.

8.1.4 Non-clausal LIKE

Syntactically unbound instances of LIKE introduce phrasal specifications, indicate repairs, and buy processing time. Such instances of LIKE have been met with derision by prescriptivists and have been described as meaningless and considered symptomatic of careless speech (Newman 1975:15). Although syntactically unbound uses of LIKE are most adequately described as filler items, they are not merely parenthetical, but serve as floor-holding devices indicating that the speaker wants to continue the turn.

Non-clausal LIKE is common in all regional varieties, and their average frequency amounts to 0.45 instances per 1,000 words which translates into a fifth of all LIKE uses. With the exception of its high rate in IndE, syntactically unbound LIKE is rather inconspicuous. However, the high rate in IndE offers intriguing insights into the adaptive processes at work during localized implementation of globally available features (cf. section 8.3.2). Over 40 percent of all instances of LIKE in IndE are syntactically unbound which is about twice the quantity of the cross-varietal average. The high proportion of non-clausal LIKE shows that LIKE use in IndE differs markedly from its use in other varieties (cf. Siemund, Maier & Schweinberger 2009). This finding corroborates Valentine's (1991) observation that "[a]lthough *like* does not appear as extensively nor perform as many functions as in AmE, […] such

instances of this element occur frequently and freely in the Indian English discourse" (Valentine 1991:332). Contrary to Valentine's (1991) claim that LIKE use in IndE is functionally limited, the present analysis shows that this understanding is too simplistic. Considering the finding that basically all functionally distinct subtypes of LIKE are more or less part of each communal grammar investigated here, Valentine (1991:332) misinterprets a quantitative difference as a qualitative difference. The present analysis of the ICE data suggests that the specificity of the Indian profile is caused by differences in the relative frequencies of functionally distinct variants but not the absence of functional variants.

8.2 Discussion of the hypotheses

After having reviewed the most important findings from a perspective that focuses specifically on individual types of LIKE (section 8.1), the following sections evaluate the hypotheses this study set out to test.

8.2.1 Discussion of hypothesis 1: LIKE is a marker of teenage speech

The first hypothesis (cf. section 5.4), which proposes that *LIKE is most common among adolescents and young adults*, addresses LIKE's function as a marker of identity. The evidence supports hypothesis 1, because a monotonic recess of LIKE with age in apparent-time is the most consistent correlation between LIKE use and social factors emerging from the data is a monotonic recess of LIKE with age in apparent-time. Furthermore, the results of the study confirm that this trend is not locally confined, but corroborate D'Arcy's tentative assertion that "the association of *like* with younger speakers seems to hold across the English-speaking world" (D'Arcy 2007:391).

In all but four cases (cf. Table 137) the multivariate analyses have confirmed that age represents the best predictor for the use of this vernacular

feature. According to Labov (2001), the consistently declining rates with increasing age reflect either age-grading or generational change. In a case similar to the one analysed here, Tagliamonte (2005:1904) argues that

> [t]his trajectory resembles the classic pattern for age-grading — a change correlated with 'a particular time of life' (Chambers 1995:164), rather than a change that is percolating successively through the generations, as would be expected with an incoming grammatical change.

The monotone recess with increasing age is the most common pattern of LIKE use, but it is not universal. Clause-initial LIKE in IrE, JamE, and PhiE, for instance, peak not among the youngest speakers, but among speakers in their late twenties and early thirties. Interestingly, this pattern is gender-specific in IrE as it applies only to females while males show the more common linear trajectory. The existence of a late peak is suggestive, because it may indicate that LIKE is becoming unfashionable and rejected by adolescents due to its association with older cohorts. The late peak is, in this interpretation, the product of the adolescent's extrication from the family circle and as an externalization of rebellion against authority, as proposed by Chambers (1995:170-171):

> Rebellion can be expressed superficially in distinctive outer markings such as green-dyed hair, nose-rings and purposely torn jeans. It is also marked in a linguistically superficial way, by the use of a distinctive vocabulary called slang, in which terms become fashionable and serve as markers of in-group membership, and then quickly become outmoded in order to mark their users as outsiders.

In contrast to younger speakers, older speakers unanimously exhibit the lowest rates of this pragmatic marker feature. Hence, LIKE, although not limited to younger speakers, is significantly less likely to be used by older speakers. This even applies to clause-final uses, which are, despite my own expectation, not most common among non-mobile, older rural males (NORMs), but among younger speakers. It is safe to assume that the higher rates of clause-

final LIKE are not a result of acts of identity (cf. Le Page 1968), but reflect a parasitic effect linked to the increase in clause-medial LIKE. In this view, the increase in clause-final LIKE is not caused by its association with a specific reference category, but the inability of speakers to differentiate between functionally distinct uses of LIKE.

The correlation between LIKE use and age leads to the question of "whether these linguistic items really represent linguistic change or whether they are simply items that distinguish teen-talk from the language of both younger and older speakers" (Tagliamonte 2005:1897). The answer to this question is more complex than it may superficially appear. At first sight, the monotonic decline in apparent-time adds support to the assumption that LIKE use "is presumed ephemeral and temporally banded, appropriate for a certain stage of life and then shrugged off when its suitability wanes" (D'Arcy 2007:398). However, while it is true that we observe a highly stable and consistent association of LIKE use with young adults and adolescents, the data also show that clause-initial LIKE is, as discussed above, most prevalent among speakers in their late twenties and early thirties in IrE, JamE and PhiE. This late peak is inconsistent with the hypothesis that LIKE functions universally as an identity marker among adolescents.

In addition, the rate of real-time change in IrE, which is most pronounced among speakers in their late twenties and early thirties, suggests that this age group continues to use LIKE as they grow older, while unintentionally introducing this virus-like element into formerly constrained registers and styles. Based on this finding, it is not implausible to hypothesize that LIKE may be undergoing a change with respect to its primary function: a loss of social meaning and a simultaneous increase in pragmatic weight. In other words, LIKE seems to lose its status as an index of youth culture and appears to become a fully functional part of IrE grammar without a distinct sociolinguistic profile. Indeed, this incipient trend is not confined to IrE, but similar observations apply to JamE and PhiE. The remarkable increase in LIKE in IrE, JamE, and PhiE over less than fifteen years, confirms that it is not only

adolescents who have adopted this form, but that it is already a functional element among speakers up to their forties. This observation implies that LIKE has transgressed teen-talk and become a widespread and salient feature of contemporary vernacular more generally.

An additional point which contradicts LIKE' being solely a marker of teen-talk relates to the linearity of the recess: a steady decline suggests that the functionality of LIKE is not limited to its use as a social index among adolescents, but that it diffuses into older cohorts due to its pragmatic versatility. If LIKE was used only as a social index, then we should observe a marked decline in the frequency of LIKE around age twenty to twenty-five. Contrary to this expectation, there is not a steep decline, but a steady and monotone recess.

Hence, asserting that LIKE is simply a feature used solely by teenagers is too simplistic and can neither account for the existence of several late peaks nor rapid real-time changes affecting speakers beyond adolescence. The notable increase over a brief ten- to fifteen-year period also rules out both age-grading and generational change as the only types of change at work. In other words, the evidence suggests additional communal change. Communal change refers to situations "when individuals themselves shift the frequency of linguistic features over their lifetimes alongside incremental change from one age group to another" (Tagliamonte & D'Arcy 2007:202).

In fact, this is exactly what recent analyses suggest:

> However, many recent studies have shown that adults do have the capacity to change their linguistic systems to a significant degree after this critical period (Sankoff 2004). Real-time replications consistently show some adult movement in the direction of the change (Labov 1994:Ch. 4). A real-time restudy of Montreal French (Sankoff et al. 2001) found a shift from apical to uvular showed the total conversion to uvular /r/ that was characteristic of many preadolescents. (Labov 2007:349-350)

Tagliamonte and D'Arcy (2007:202) elaborate on this aspect and suggest that

> young adults continue to advance ongoing linguistic changes well into adulthood (Nordberg & Sundgren 1998; Nevalainen & Raumolin-Brunberg 2003). In other words, the most recent research on linguistic change suggests that generational change and communal change can progress simultaneously. (Tagliamonte & D'Arcy 2007:202)

If the grammar of individual speakers is thus more flexible than previously assumed (cf. also Labov 2010:311), the drastic increase observed in the case of LIKE is best attributed to (partial) communal rather than generational change. Adding weight to this proposition is the fact that apparent-time studies have systematically underestimated the degree of change when comparing apparent-time analyses to the results of real-time studies re-visiting the original location (cf. e.g. Cedergren 1984, 1987; Sankoff et al. 2001; Sankoff 2004 and particularly Trudgill 1988).

The complex relationship between apparent-time and real-time, therefore, poses serious issues concerning the validity of premature extrapolations and attestations of change from apparent-time data alone. One approach addressing this issue is offered by Labov (2001), attesting that age stratification alone is insufficient to distinguish stable age-grading from generation change. As a consequence, Labov (2001) proposes that age stratification in apparent-time requires either additional information of social class distributions (Labov 2001:77), knowledge of gender differences, or complementary real-time analyses.

In the present case, the additional real-time analyses have validated the apparent-time trajectories and provided a more detailed understanding of the processes shaping ongoing change. However, the analysis of LIKE use in real-time indicates that the global diffusion of LIKE progresses at different rates in geographically noncontinuous populations. Indeed, the relation between linguistic variation in apparent-time and in real-time has emphasized the

variety-specificity of the trajectories of change. For example, varieties with rather similar sociolinguistic profiles like CanE and JamE differ markedly with respect to the rate of change in real-time. While LIKE use did not significantly change over time in CanE, it increased notably over a brief fifteen-year period in JamE. A still different trajectory appears to apply for the IrE data. Here, LIKE use increased significantly, but this increase is limited to speakers younger than 34 years of age.

With respect to future developments, the substantial frequency of LIKE use among adults and even older speakers of CanE suggests that its use will stabilize and slowly diffuse through older cohorts by means of generational change. This trajectory is indeed congruent with findings from analyses on the institutionalization of lexical innovations (cf. Fischer 1998:174), which show that lexical innovations settle at a lower rate after having diffused through all social starta, i.e. after they are institutionalized. Further support for this interpretation can be seen in the significant degree of social stratification among CanE speakers, which indicates that LIKE has not as yet diffused into more conservative social strata and registers.

At least two alternative, or perhaps complementary, explanations seem viable here. During its implementation in CanE, LIKE could have undergone re-evaluation as well as re-interpretation, thereby gaining an additional function, i.e. indexing socio-economic status, which LIKE did not possess previously in AmE. This perspective is in line with authors (e.g. Kachru 1992; Buchstaller 2008; Buchstaller & D'Arcy 2009; Meyerhoff & Niedzielski 2003), who emphasize the fact that the adoption of globally available variants is not merely a passive act of reception or borrowing, but a proactive process during which the social meaning of forms undergoes modification and is adapted to the local norms and practices.

Another viable approach draws attention to the fact that social stratification is suggestive of incomplete diffusion, thereby indicating that the incoming form has not yet intruded into all social groups. This latter

perspective assumes a conflict model of social organization which focuses particularly on class-specific language use as, for example, Labov (1966, 1972). Both views are coherent and offer valid accounts of the subject matter. Hence, they are taken to be complementary, and not exclusive in their explanatory power.

8.2.2 Discussion of hypothesis 2: LIKE as a marker of female speech

The second hypothesis (cf. section 5.4), which proposes that *LIKE is typically a female feature and thus gendered*, also addresses LIKE's function as a marker of identity. The evidence only partially supports or even challenges hypothesis 2 in a narrow reading.

The multivariate analyses have shown that the interplay between LIKE use and gender is rather complex and sensitive to culturally shaped gender roles: the results challenge the notion of a uniformly consistent effect and instead reveal variety-specific gender differences. For instance, in AmE none of the uses of LIKE are correlated with gender, while the opposite holds true for IndE. But, the association of LIKE with gender is not random: once a functional variant of LIKE is perceived as being typical for female speech, this perception appears to become increasingly associated with other variants as well. Thereby the gender distributions of LIKE corroborate the assertion that it is not the specific linguistic element or behavior which is typical for either males or females, but that these gendered practices are culture-specific and "differ considerably from culture to culture, from place to place, from group to group, living at the intersection of all other aspects of social identity" (Eckert 1998:66). In other words, "[s]ociolinguistic variation is parasitic upon linguistic variation. It is an opportunistic process that reinforces social distinctions by associating them with particular linguistic variants" (Labov 2002).

However, it is reasonable to assume, given the assertions by various scholars (e.g. Siegel 2002, Dailey-O'Cain 2000), that LIKE may have been preferred by females during earlier stages of change. Hence, it cannot be ruled out that LIKE lost its gender sensitivity once it had diffused through most social strata and emerged as the widely used pragmatically functional element it is today. The situation in CanE appears to be reminiscent of this process. In contrast to AmE, LIKE use in CanE indicates gender sensitivity. The use of clause-initial LIKE, for example, is strongly correlated with females, while other variants are not notably associated with gender. This loss of gender marking can be interpreted as the final stage of a process in which uses of LIKE have lost their association with females as the social reference group which was initially responsible for its promotion and spread. EngE, on the other hand, offers a quite distinct setting. Here, LIKE use is associated with females, although the frequency of this form suggests that the change is nearing completion. It appears that LIKE is only in the initial stage of change, and female speakers have only recently begun to take the lead in this process. Although this reasoning is altogether plausible, it should be kept in mind that the findings of the present study are tentative in this respect and require further research.

Regarding the variety-specificity of gender differences, an intriguing explanation is offered by Tagliamonte (2005:1912-1913), who proposes that gender differences are developmental rather than inherent. In other words, differences in the linguistic behavior of males and females exist neither due to the feature itself nor to the underlying mechanisms of change, but as a result of dynamics within the social network:

> These trends show that sex differences — at least with respect to these discourse/pragmatic features — are developmental, and are learned. They do not appear to be endemic to the features themselves, but are created in the speech community, within the peer group. (Tagliamonte 2005:1912-1913)

Tagliamonte's (2005) hypothesis matches the findings of the present study which corroborates that the degree and direction of gender differences are variety-specific, but they also suggest that the impact of gender depends crucially on the phase of change (cf. Labov 2001:320-322).

In addition to their culture-specificity, gender effects are dependent on the specific type of LIKE: where gender effects emerge, they are most likely to surface in the use of clause-medial LIKE. In five of the eight regional varieties, it is clause-medial LIKE which is gendered – none of the other forms exhibits such a consistent pattern.

In his study on LIKE use in IrE, Schweinberger (2012) argues that the similarity between males and females may be linked to the overt stigmatization of this form. LIKE is almost universally described as a non-standard, vernacular feature and often viewed overtly negatively (Dailey-O'Cain 2000:69–70; Schourup 1982:29), "highly stigmatized by normative grammarians" (Buchstaller 2001:2) and drawing "overt attention and commentary" (D'Arcy 2007:411). This is not surprising, as innovations are frequently met with derision (D'Arcy 2007:387). It could be the case that the stigmatization of LIKE interacts with general trends observed in the linguistic behavior of women. For example, women conform more closely to sociolinguistic norms when these are overtly prescribed, but conform less than men do when they are not (Labov 2001:272, 292). In other words, women are more likely to adopt incoming and innovative forms as long as they are not overtly stigmatized. This implies that the differences between women and men, or the lack thereof, might well be accounted for by different degrees of overt prestige or stigmatization. If LIKE is overtly stigmatized, then we do not observe the expected female lead, while we will observe a female lead where LIKE is not (yet) marked by overt stigma.

8.2.3 Discussion of hypothesis 3: The universality of the Labovian model

The third hypothesis (cf. section 5.4), which proposes that *the Labovian model of the social motivation for linguistic change is valid with regard to pragmatic change*, addresses the issue of how appropriate the Labovian model of change from below is in cases where one is not dealing with phonological transmission, but the diffusion of pragmatic innovations such as the spread of LIKE. The evidence only partially supports or even challenges hypothesis 3 in that the analyses show that the spread of LIKE progessed more quickly and is accompanied by substantially less social stratification than Labov's model would predict.

Labov (1994, 2001, 2010) focused on language change through contact between continuous varieties, while the contact scenarios investigated here represent geographically noncontinuous contact scenarios. Thus, the contact situations investigated by Labov (2010) do not exhibit the degree of cultural diversity we are dealing with here. In contrast to the changes analyzed by Labov (2010), culturally determined social meanings of linguistic innovations represent key aspects of the underlying sociolinguistic motivation for social diffusion of lexical elements.

In addition, mechanisms which apply in cases of sound change are not necessarily involved in cases of lexical diffusion and in particular when innovations are being disseminated globally. For instance, Thomason and Kaufman (1991) show that lexical borrowing, as in the present case, is distinct from phonological and structural borrowing in terms of both the rapidity of diffusion and the necessary amount and quality of contact: "[a]s is usual in borrowing situations, words are borrowed first and structural features later, if at all" (1991: 40).

Another inadequacy relates to the systematic over-estimation of social stratification and gender differentiation. While theories of society as a compartmentalized, layered entity are appropriate when dealing with

unconscious changes such as sound changes, they overestimate class differences in cases of pragmatic diffusion. Contrary to such "conflict models" of society, more recent approaches in contemporary sociolinguistic theory have shown a bias towards treating "social stratification as a product of shared values and broad social consensus" (Milroy & Gordon 2003:95). In the latter understanding, social classes are considered "as forming a continuum rather than sharply divided, and […] the basis of social hierarchy is not different relations to the market (as both Marx and Weber proposed), but different status values assigned to different occupations" (Milroy & Gordon 2003:95). Social class is thus better grasped as a "group of persons sharing similar occupations and incomes, life-styles and beliefs" (Milroy & Gordon 2003:95). The minor role of socio-economic status in the present analysis adds to the empirical support for the inadequacy of conflict models. The absence of social stratification suggests that the class boundaries are not insurmountable, but readily crossed by socially mobile speakers. These socially mobile speakers are decisive in spearheading local diffusion, during which the predictive power of social class lessens rather swiftly in cases of lexical borrowing. Like syntactic variability, lexical variables are less likely to be socially indexical (Milroy & Gordon 2003:171). In other words, in cases of lexical diffusion, variants are less prone to exhibit significant degrees of gender differentiation and social stratification.

Moreover, the impact of mass media, which is negligible in cases of phonological transmission (cf. Labov 2001:228) since it reflects "change rather than generating it" (Labov 2010:195), appears to be a crucial factor contributing to pragmatic diffusion (Muhr 2003:123). Based on the similarities of the (age) distributions, the functional employment and the syntagmatic positioning, it is highly plausible that the media, rather than face-to-face contact, served as a means of diffusion. This finding runs counter to the interactive principle put forth by Labov (cf. Labov 2001:229) which presumes a marked dominance of the effect of face-to-face transmission as opposed to

diffusion via the mass media (cf. Labov 2001:228-229, 362-363, 385; 2010:195).

Despite the issues addressed above, both the distributional patterns of extra-linguistic variables in apparent-time and the relationship between apparent-time and real-time are very well accounted for in a Labovian framework. In addition, variationist approaches have proven to be adequate, particularly in cases of change from below. Complementary perspectives such as, for instance, Eckert's (2001) analysis of prestige and identity, contribute greatly to a more detailed understanding of mechanisms underlying ongoing change. In view of these considerations, the evidence only partially corroborates Hypothesis I, stating that *the sociolinguistic mechanisms of phonological change can be applied to the study of the discourse marker LIKE and other morpho-syntactic and lexical features*.

8.2.4 Discussion of hypothesis 4: Diffusion and social stratification

The fourth hypothesis (cf. section 5.4), which proposes that *LIKE is socially stratified particularly in locales where it has only recently been introduced* addresses the issue of how socially stratified the use of pragmatic innovations is and how social stratification relates to the degree of diffusion of speech communities. The evidence challenges hypothesis 4 because (a) LIKE is either only insignificantly stratified or, if it is, not as the Labovian model would predict; and (b) the analyses show that the most substantial social stratification is observable in locales where LIKE is well established.

The fact that occupation – as an approximate index of the socio-economic status of speakers – more often than not fails to significantly affect LIKE use is unexpected. This absence of sharp social stratification is particularly remarkable considering Labov's assertion that "most of the linguistic changes in progress studied in the 2nd half of the 20th century show a high degree of social stratification" (Labov 2002).

Considering this intimate relation between social stratification and linguistic variation, we expected LIKE to be distributed rather heterogeneously among social strata. Based on the curvilinear hypothesis described by Labov (1994, 2001), the group using LIKE most was expected to be located in the centre of the socio-economic scale, i.e. speakers in clerical and managerial positions. Consider Labov's curvilinear hypothesis: "[W]hile stable sociolinguistic variables showed a monotonic social class distribution, a monotonic distribution in age groups was associated with a curvilinear pattern in the socioeconomic hierarchy" (Labov 2001:32).

However, this is not what the analysis showed – at least in the majority of cases and for most forms of LIKE. Despite the fact that age predominantly followed the monotonic distribution, the prevalent scenario is that LIKE does not show any sign of social stratification such as a curvilinear pattern. This absence of social layering is, however, in line with research on syntactic changes which proposes that syntactic variables are less likely to be socially marked (Hudson 1996:45).

Despite the hypothesis that social stratification is most salient in locales in which LIKE has only recently been introduced, the most consistent effect of social class has not been detected in IndE or PhiE, but in NZE and CanE. Furthermore, only in NZE was clause-medial LIKE socially stratified, while it was the use of clause-initial and non-clausal LIKE in CanE.

There are several possible explanations for why the occupation of speakers appears to be a negligible factor in the present investigation. For instance, we may ask whether the current occupation of speakers accurately reflects the socio-economic status of speakers. Although this objection is plausible, there are some issues which need to be addressed. Students have not been considered in the analysis of the social stratification of LIKE use, as the occupations of their parents were not available. Neglecting the predominantly younger cohort not only substantially reduced the data set but also cut off the most relevant speakers, since LIKE use is particularly salient among younger

speakers. Unfortunately, this was unavoidable given that the information on socio-economic status is not accessible to researchers.

8.2.5 Discussion of Hypothesis V: LIKE use is modified during local implementation

The fifth hypothesis (cf. section 5.4), which proposes that *the regional locale and the cultural norms and practices shape the local implementation of LIKE* addressed processes taking place during the diffusion of globally available elements through geographically discontinuous and culturally diverse speech communities. The findings support Buchstaller and D'Arcy's (2009) hypothesis that innovations are not passively borrowed, but that, on entering new speech communities, the pragmatic functions of linguistic innovations are re-negotiated.

A case in point is the remarkably high proportion of non-clausal LIKE use in IndE. The frequency with which non-clausal LIKE is used in IndE suggests not only that LIKE has undergone adaptation to the local system, because the positional profile of IndE and thus the functional employment of LIKE in IndE differs markedly from the positional profiles of other varieties. There are several aspects which may help shed light on why IndE exhibits such an idiosyncratic profile. It is likely that the unusual Indian pattern suggests incomplete acquisition of the functional diversity of LIKE among adult speakers due to limited contact with LIKE use resulting from the absence of LIKE in EngE - the target variety of standard IndE particularly among graduate students and university staff (cf. Schneider 2003:171). Subsequently, these adults transmit their functionally limited use of LIKE to L1 learners, thereby perpetuating functionally limited use of LIKE throughout generations (cf. Labov 2010:333). Functionally limited refers to an over-representation of non-clausal LIKE while other variants are underrepresented due to avoidance strategies, or simply due to a lack of knowledge about possible functions and constraints of LIKE in other varieties.

Whatever the exact causes for the idiosyncratic use of LIKE in IndE, it is safe to assume that the variation across varieties of English is confined to differences in frequency of certain forms and their association with specific social reference groups, while the functions and syntactic constraints are mostly invariant. This finding complies with Thomason and Kaufman's hypothesis that it is not language-internal factors which condition linguistic inference but social ones (1991:35). According to this view, "it is the sociolinguistic history of speakers, and not the structure of their language, that is the primary determinant of the linguistic outcome of language contact" (1991:35).

8.3 Discussion of other findings

In addition to testing the aforementioned hypotheses, this study has brought to light several other intruiging findings that have so far not been addressed. The following section addresses additional findings that the study has unearthed.

8.3.1 Priming, accommodation and language change

The most consistent correlation in the present study is that between LIKE use and the PAI index. This moderate but highly stable correlation suggests that LIKE use is significantly affected by priming and accommodation. The consistency of the size and direction of this measure suggests that priming may affect language change more than commonly assumed.

It should be kept in mind, however, that the PAI index is only a rough and approximate measure. The correlation between the PAI index and the use of LIKE is tentative but not conclusive, because further studies using more precise measures would be required to confirm a consistent effect of priming on language change. The rather weak effect of the PAI index indicates that its effect on proccesess of language change is negative overall, despite being highly

significant. The results encourage a more elaborate investigation of this issue. Indeed, the consistency of the effect indicates that LIKE may well serve as a promising direction for future research which could probe more deeply into the potential of integrating psycholinguistic concepts into models of language change and variation. This may turn out to be particularly relevant as the "think-piece" by Jäger and Rosenbach (2008) has been met with profound skepticism (cf. e.g. Eckhardt 2008; Traugott 2008).

8.3.2 The global diffusion of LIKE and the media

We are now in a position to re-consider the startling observation that LIKE use was locally confined to the British Isles for at least a century before starting to spread and becoming the globally available and abundant feature it is today. In spite of the fact that the mass media has been neglected as a means of transmission of phonological innovations (Hickey 2003:341; Labov 2001:228), the timeline of LIKE's global diffusion is suggestive. Indeed, the timeline lends plausibility to the hypothesis that the rise of LIKE relates to an increase in transnational mass media. A study explicitly addressing media-induced language change is Muhr (2003). Although this study addresses the language contact between Austrian German and German German, its implications appear viable for the present analysis. Indeed, the contact scenario between Austrian German and German German shows that "the impact of this language contact is increasing and that it can be directly linked to the amount of TV-viewing time" (Muhr 2003:103). Muhr (2003:123) furthermore concludes

> that mass media, and particularly satellite television broadcasting, is creating an intensified situation of language contact that is resulting in accelerated language shift on the part of a smaller national variety of German, Austrian German, in the direction of a dominant national variety, German German. As such, it supports the sociolinguistic principle that contact leads to convergence (Weinreich 1953), and adds to the literature an illustration of how a mass medium can create new kinds of contact situations.

In a similar vein, Açikalin (2004) confirmed that television played an important role in the adoption of non-standard forms by adolescents. In this comparative longitudinal study between 1989 and 1999, Açıkalın (2004) investigated the impact of television on the spread of non-standard elements among Turkish adolescents. Tagliamonte (2011:41-42) also provides examples of changes where innovations spread through media contact rather than direct face-to-face contact and concludes:

> The results from these studies are puzzlingly equivocal, perhaps because the precise nature of media influence is difficult to define or measure. This is a tantalizing new horizon for development in future research. Further, it remains to be discovered what type of linguistic feature - lexical, morphology, syntax. pragmatic - can arise from contact with the media. (Tagliamonte 2011:42)

The analysis of LIKE suggests that the mass media appear to have affected the diffusion of LIKE. This interpretation is supported by the fact that the most visible extra-linguistic features of speakers (age) are the most constant across English-speaking speech communities around the globe. Despite Labov's claim that the media reflects change rather than generating it (2010:195), it seems that the media have been a crucial factor in facilitating the global diffusion of LIKE during the latter half of the twentieth century.

The effect of the media in this process is at least twofold. Firstly, the media has rendered the diffusion of LIKE possible by facilitating contact between geographically noncontinuous locales. In this sense, the media have made LIKE globally available to smaller, regionally-defined speech communities and familiarized speakers with this innovation. As a consequence of becoming increasingly acquainted with this new feature, speakers stopped perceiving it as a foreign feature and have integrated this innovation into their local system.

Secondly, the media have been driving the spread of LIKE by promoting its perception as a prestige variant due to its association with the United States as a culturally and economically dominant group. This is a crucial factor, in

particular with respect to the important role that adolescents play in the diffusion of LIKE. Consider Muhr (2003:123):

> Young people are especially susceptible to this perception, as they pass through a stage of restlessness and seek to define themselves in contrast to the local values represented by their parents' generation. To this we may add the prestige of new media themselves, which frequently symbolise modernity and worldliness. Thus the informal standard of dubbed films has become a reference point for the informal speech behaviour of youth between the ages of twelve and 20, who use GG swear words, slang expressions and colloquial terms to index values outside those of their regional culture.

However, the British and Indian patterns strongly suggest that even cultural and economic dominance, a shared language, and the availability of innovations due to mass media are insufficient to account for the global diffusion of LIKE. At least one factor has additionally affected this process: the refusal by EngE speakers to adopt the variant despite being available due to sufficient high-quality contact indicates that LIKE continued to be perceived as foreign. In this sense, LIKE was associated with a geographic - in this case US American - rather than social identity among the British speech community. Only when this association of LIKE with AmE waned did adolescent EngE speakers claim it as a marker of social identity. Nonetheless, the media quite probably influenced this process by familiarizing young speakers of EngE with LIKE, and they then endorsed its perception as a prestige variant. This argument is congruent with Britain (2002:618), who describes the rejection of innovations as one of three possible outcomes in cases of contact between global (or supra-local) innovations and local norms (cf. Buchstaller & D'Arcy 2009:292).

Although LIKE almost certainly did not originate in AmE (D'Arcy 2007:400), it seems that the increase in exposure to American vernacular triggered the global spread, starting in AmE and subsequently affecting the rest of the English-speaking world. Hence, the diffusion of this form is temporally

and spatially distinct from its origins. We can assume that LIKE initially crossed the Atlantic through emigration to the United States and prolonged language contact.

The question arising is why LIKE started to spread. There are at leat two reasons for LIKEs success: firstly, its ability for clause-internal modification (focusing and hedging of constituents) and, secondly, covert prestige, i.e. its association with esteemed social reference groups and social categories as well as with desirable personality traits such as attractiveness, friendliness and successfulness (cf. Dailey-O'Cain 2000), that triggered its dramatic spread across the English-speaking world (Tagliamonte 2005:1898).

Hence, it is likely that the epicenter of the twentieth-century spread of LIKE around the English-speaking world is the United States and it probably took the same path already described by Buchstaller and D'Arcy (2009) in their study of *be like* in AmE, EngE, and NZE. Buchstaller and D'Arcy (2009) argue that this quotative complementizer developed in AmE and subsequently spread to EngE and NZE, but was re-interpreted and adapted during its implementation into the local system. In a study of *be like* in Glasgow, Macaulay (2001) similarly observed that, when entering the local speech community, *be like* has "clearly been reanalysed and indigenised by speakers" (Meyerhoff & Niedzielski 2003:545). More importantly, however, is the fact that the social distribution of LIKE differs across locales indicating that the associations attached to LIKE are subject to change during indigenization. This flexibility and modification allow new social categories to claim this form as a marker of covert prestige. As Buchstaller (2008:19) points out:

> [R]esearch on linguistic leveling corroborates findings from cultural studies and economics that, while participating in global (supra-local) trends, localized communities develop new and idiosyncratic routines for doing so.

In this view, I take the complex setting of distributional patterns to reflect the homogenizing and simultaneously diversifying effect of globalization as described by Hjarvard (2004:95):

> In this sense the media are a homogenizing factor. On the other hand, the media have considerable potential for a new kind of linguistic diversity that transcends the various national languages in that innovation is attached to use of the media, not the individual user.

The changes in the social meaning of LIKE are not random and they do not affect all aspects to the same degree. For instance, the stability of the age distribution resembling the classic pattern of age-grading suggests that "only superficial aspects of the innovation are transferred" (Meyerhoff & Niedzielski 2003:538). The stability of the age distribution on the one hand and the inconsistent patterns of gender and social class on the other indicate a moderate degree of quality face-to-face contact between individuals: if speakers of geographically noncontinuous varieties have a high degree of face-to-face contact, i.e. exposure to the vernacular of different regional varieties, we would observe the same distributions worldwide. As this is not the case, we can conclude that only the most salient associations or functions of linguistic features are preserved, while more intricate details are reinvented. Applied to the present case, this translates into the preservation of its association with youth and of its function as a clause-internal modifier, while associations with gender and social status have been subject to re-negotiation. The idiosyncratic behavior of LIKE in IndE is possibly the most perplexing case in terms of substantiating Kachru's (1992) assertion that "when a linguistic resource is adopted by new speakers, there is always some kind of transformation in meaning and form" (cited in Meyerhoff & Niedzielski 2003:538). What is remarkable here is that the way LIKE is used matches exactly what is expected when supra-local features are not simply borrowed wholesale, but adapted to the local norms and practices of the speech community.

Meyerhoff and Niedzielski's (2003) framework represents a meaningful contribution to the phenomenon and offers a valuable handle on the seemingly contradictory effect of simultaneous homogenization and diversification in globalization. According to their analysis, the emergence of similarities in linguistic behavior, i.e. borrowing, requires substantial quality face-to-face

contact. In short: the more contact, the more similar the performance. Conversely, less quality contact leads to a higher degree of diversity, as the innovative form is affected more strongly by modification, i.e. re-analysis and re-interpretation.

However, the discrepancy between the consistency in the pattern of age distributions on the one hand and the variety-specific effects of gender and social class on the other cannot be fully accounted for by differences in quality and salience of contact. There are at least two additional aspects to be considered.

Firstly, the attitudes toward and covert prestige attached to the donor variety by speakers of the receptor variety is a main factor determining to what degree an innovative form is indigenized. In other words, the acceptance of an innovation depends crucially on the attitudes attached to it and how speech communities perceive it. The rejection of LIKE by post-adolescent speakers of EngE, for example, contradicts the assertion that there is a simple correlation between the amount and quantity of contact on the one side, and the acceptance of a new variant on the other. Given the results of Buchstaller's (2006b) attitudinal study of *be like* in the United Kingdom, it is more plausible, that national or local identity issues are major contributing factors for the success of incoming forms. Secondly, lexical bleaching of pragmatically functional elements allows for variety-specific re-interpretation and thus greater flexibility. Consider cases of 'true' lexical borrowings. In such cases, it is lexically heavy elements which are borrowed. Although such elements may indeed adapt to the local system –for example, Anglicisms in German (cf. Onysko 2007) – their semantic content is less flexible than in cases of lexically empty or light elements such as pragmatic markers. Flexibility in use and pragmatic meaning allows innovations to fill pragmatic niches that less flexible forms may be constrained from filling. Hence, pragmatic flexibility increases the probability of successful incrementation of global innovations into local systems.

8.3.3 The ICE as a sociolinguistic data source

With respect to the innovative employment of the ICE, the present study illustrates how the ICE components can be used to match the requirements of fine-grained sociolinguistic research. The increase in versatility, achieved by splitting the ICE components into individual speaker corpora, offers a highly accurate depiction of regionally distinct usage patterns. In this sense, the ICE components represent ideal resources for cross-varietal analyses and satisfy the need for comparative data sets and comparable methods expressed by Buchstaller and D'Arcy (2009:298):

> What is needed, therefore, are reliable and comparable methods applied rigorously and uniformly across datasets to uncover which constraints hold both across and within varieties of English worldwide.

The present analysis serves as a case in point for how ICE components may serve as a resource for future sociolinguistic research: they contain not only a wide variety of extra-linguistic variables, but they also represent various registers and share a common design.

Although the processed ICE version is interesting for linguists, particularly with respect to the analysis of the sociolinguistic distribution of certain forms in regionally and culturally diverse settings, there are some deficiencies. The primary deficiency relates to the moderate number of informants in each subset. However, even this modest number of informants outperforms previous data sets, which frequently suffer from a lack of comparability and represent even fewer informants in each sub-cohort.

Another shortcoming relates to the inaccessibility of the original audio files which limits the analysis to morpho-syntactic features, since phonetic or phonological mark-up is missing. The ICE components can therefore only serve as data sets for studies which focus on lexical or morphological variation but not for studies which would require fine-grained phonetic mark-up..

Despite these shortcomings, the present study has demonstrated that the edited ICE components represent a valuable resource for sociolinguistic research, particularly when the focus is placed simultaneously on regional and social variation.

Although I acknowledge that this view is too simplistic to accurately account for the complexity of the globalized, multifaceted nature of the linguistic processes involved in global diffusion and transmission, I nonetheless hope that this discussion serves as a valid starting point for further, more fine-grained analyses of how the globalized setting influences on local speech communities.

9 Conclusion and outlook

The analysis of vernacular uses of LIKE in selected varieties of English presented in the preceding chapters has shown a complex interaction of globally stable sociolinguistic tendencies and culturally motivated variety-specific usage patterns. As some recent studies suggest that linguistic innovations undergo transformation and are adapted during their implementation in culturally diverse settings, the study of vernacular variants offers a wealth of information about the mechanisms underlying global diffusion. The present study intends to add one small piece to this line of inquiry.

In theoretical terms, the variety-specificity of LIKE use substantiates previous claims, according to which the social meaning of innovative forms are not merely borrowed wholesale, but adapted and re-interpreted to local systems during the process of implementation (cf. Buchstaller 2008; Buchstaller & D'Arcy 2009; Kachru 1992). The difference between CanE and EngE, for example, implies that neither the mere existence of (high-quality) contact nor the language-internal processes, are sufficient to account for linguistic diversity. Thus, the present analysis exemplifies the necessity for considering cultural diversity and local identities when analyzing variation of linguistic behavior on a global scale. Hence, my study confirms the need for "caution in offering global interpretations of effects of locally embedded variables such as status and gender" (Milroy & Gordon 2003:178).

The findings of the present analysis add weight to the hypothesis that the discourse marker LIKE is currently undergoing change and its analysis has contributed to a more detailed account of the exact trajectory of this process on the global as well as regional level. The evidence confirms that the global diffusion of LIKE affects all varieties of English included in the present sample. In contrast to phonological changes, the spread of the discourse marker LIKE, i.e. a pragmatic innovation, is neither necessarily accompanied by gender differentiation nor social stratification. This observation is in line with

previous research suggesting that lexical change, similar to syntactic variation, is "less likely to be socially indexical" (Milroy & Gordon 2003:171; cf. also Hudson 1996). Hence, my study corroborates research asserting "that many contextual factors over and above gender can determine the language of men and women" (Palomares 2008:263). The evidence shows moreover that the strength and direction of gender effects and social class differences are not only variety-specific, but are also sensitive to the age of speakers. Methodologically, the present analysis illustrates how the ICE components can be used as valuable resources for sociolinguistic studies of ongoing change. The computationally edited version of the ICE used here may offer intriguing opportunities for studying the local implementation of globally available forms, in line with studies such as Meyerhoff and Niedzielski (2003), Buchstaller (2008) and Buchstaller and D'Arcy (2009). The increase in versatility, achieved by splitting the ICE components into individual speaker corpora, offers a highly accurate depiction of regionally distinct usage patterns. Indeed, the present investigation rests upon the analysis of 1,925 speakers from eight geographically noncontinuous speech communities and 4,661 instances of LIKE which in terms of sheer quantity vastly overshadows previous studies of similar discourse phenomena. In this sense, the ICE components represent ideal resources for cross-varietal analyses and satisfy the need for both matching data sets and comparable methods, as expressed by Buchstaller and D'Arcy (2009:298):

> What is needed, therefore, are reliable and comparable methods applied rigorously and uniformly across datasets to uncover which constraints hold both across and within varieties of English worldwide.

Hence, the analysis of LIKE serves as a case in point for how ICE components may serve as resources for future research in sociolinguistics. Indeed, they represent ideal databases, as they offer a wide variety of extra-linguistic variables and represent various registers paired with a matching design.

The fact that we observe both highly stable tendencies – the negative correlation between LIKE use and the age of speakers – and variety-specific differences – the degree and direction of gander differences - raises the question of whether the patterns emerging for LIKE are isolated phenomena, or whether similar trajectories can be observed for other innovations. Research addressing this issue could, for example, focus on quite distinct innovations and test the degree to which their implementation matches the trajectory observed for LIKE. Ideally, these analyses use comparable methods to the speaker-based sub-corpora of the ICE components.

The consistent albeit weak correlation between the use of LIKE and the conglomerate measure for priming, accommodation, and idiosyncratic overuse is particularly intriguing. Although the PAI index was included as a measure to prevent over-estimating the effect of traditional extra-linguistic variables such as the age and gender of speakers, the consistency of the effect indicates a weak but significant effect of the psychological underpinning on mechanisms of change and variation. The effect of the PAI index suggests that the sociolinguistic analysis of language change might benefit greatly from integrating psycholinguistic concepts. Nonetheless, the impact of priming has until recently been ignored in studies of ongoing change (cf. e.g. Jäger & Rosenbach 2008a, 2008b). In this view, future research may profit from including psychological factors, as this allows for a more detailed account of mechanisms underlying orderly heterogeneity.

Moreover, this study focused on the diffusion of a linguistic innovation in geographically noncontinuous varieties, which necessarily relates to media induced-contact rather than face-to-face contact; it thereby draws attention to the impact of media on language change. Although difficult to quantify and reliably test for the exact impact of media effects on language change, the rapidity of LIKE's global diffusion suggests that the media represent a significant factor in similar scenarios. This implication is meaningful given that most sociolinguistic analyses assume face-to-face contact as the prototypical scenario in cases of diffusion (cf. e.g. Labov 2001:228):

> A uniform increase in contact with other dialects may also be an effect of the mass media. But all of the evidence generated in this volume and elsewhere points to the conclusion that language is not systematically affected by the mass media, and is influenced primarily in face-to-face interaction with peers.

However, more recent accounts put stronger emphasis on the mass media as a means of diffusion (cf. Tagliamonte 2011:41-42). In her study on the diffusion of lexical elements, Fischer (1998), for instance, makes a compelling point that the media are particularly relevant with respect to the spread of lexical innovations such as neologisms or blends (other studies which supstantiate this claim are e.g. Muhr 2003 and Açikalin 2004).

In some cases, the present study did not arrive at definitive answers. This is not surprising, given that sociolinguistic work on both regional and global diffusion, even of a language as well documented as English, is always an exceedingly difficult enterprise. In the present case, the existence of matching, easily available data sources, gives the researcher a relatively privileged position. The ICE components, which formed the empirical basis for most of this work, offer a wealth of extra-linguistic information about speakers in geographically distinct locales. While the use of the ICE components for comparative analyses is by no means a recent approach, the edited version created for the present purpose allows for fine-grained sociolinguistic analyses from a global perspective. The edited ICE version used here has reached a size and systematicity which, in terms of detail and comparability, was not available to previous researchers in the field. This is true not only for studies based on the apparent-time construct, but also for work analyzing changes in real-time.

In addition, the ICE lends itself to more fine-grained sociolinguistic investigations of regional and sociolinguistic variation even more now, due to improved comparability of findings with respect to applying a standardized methodology.

Further systematic analyses of the mechanisms of global diffusion could both enrich our understanding of the local actuation and incrementation of innovations, and also help to comprehend more fully the sociolinguistic factors responsible for the spread of forms across Englishes world-wide.

10 References

Abraham, Werner

1991 Discourse particles in German: How does their illocutive force come about. In Werner Abraham (ed.), *Discourse Particlcs: Description and theoretical investigation on the logical, syntactic and pragmatic properties of discoursec particles in German*. (Pragmatics and Beyond, New Series 12.), Amsterdam-Philadelphia: John Benjamins, 203-252.

Açıkalın, Isil

2004 The perpetuity trend of nonstandard linguistic forms. *International Journal of the Sociology of Language* 165: 143–154.

Aijmer, Karin

2002 English discourse particles: Evidence from a corpus. Amsterdam: John Benjamins.

Aitchinson, Jean

[3]2001 *Language change: progress or decay?* Cambridge: Cambridge University Press.

Amador-Moreno, Carolina P.

2010a *An Introduction to Irish English*. London: Equinox.

Amador Moreno, Carolina P.

2010b How can Corpora be used to explore literary speech representation?. In O'Keefee & McCarthy (eds.), 531-544.

Andersen, Gisle

1997 They like wanna see like how we talk and all that. The use of like as a discourse marker in London teenage speech. In Magnus Ljung (ed.), *Corpus-based studies in English*. Amsterdam: Rodopi, 37-48.

Andersen, Gisle

1998 The pragmatic marker like from a relevance-theoretic perspective. In Andreas H. Jucker & Yael Ziv (eds.), *Discourse markers: descriptions and theory*. Amsterdam & New York: John Benjamins, 147-170.

Andersen, Gisle

2000 *The role of the pragmatic marker like in utterance interpretation*. Amsterdam & New York: John Benjamins.

Andersen, Gisle

2001 *Pragmatic Markers and Sociolinguistic Variation. A relevance-theoretic approach to the language of adolescents*. Amsterdam & New York: John Benjamins.

Anderwald, Lieselotte

2004 English in the Southeast of England: morphology and syntax. In Bernd Kortmann et al. (eds.), 175–195.

Audretsch, David

2000 Knowledge, globalization, and regions: An economist's perspective. In John H. Dunning (ed.), *Regions, Globalisation, and the Knowledgebased Economy*. Oxford, U.K.: Oxford University Press, 63–81.

Auer, Peter & Susanne Günthner

2005 Die Entstehung von Diskursmarkern im Deutschen - ein Fall von Grammatikalisierung. In Torsten Leuschner, Tanja Mortelmans & Sarah de Groodt (eds.), 335-362.

Baayen, Harald

2008 *Analyzing Linguistic Data. A practical introduction to statistics.* Cambridge: Cambridge University Press.

Bailey, Guy, Tom Wikle, Jan Tillery & Lori Sand

1991 The apparent time construct. *Language Variation and Change* 3: 241-264.

Bakir, Murthada

1986 Sex differences in the approximation to Standard Arabic: a case study. *Anthropological Linguistics* 28: 3-9.

Bauer, Laurie

2002 *An Introduction to International Varieties of English*. Edinburg: Edinburg University Press.

Bayley, Robert & Otto Santa Ana

2004 Chicano English: morphology and syntax. In Bernd Kortmann et al. (eds.), 374–390.

Beal, Joan

2004 English dialects in the North of England: morphology and syntax. In Bernd Kortmann et al. (eds.), 114–141.

Bell, Allan

1984 Language style as audience design. *Language in Society* 13.2: 145-204.

Bell, Allan

2001 Back in style: Re-working Audience Design. In Penelope Eckert & John R Rickford (eds.), 139-69.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan

1999 *Grammar of Spoken and Written English*. London: Longman.

Bisang, Walter, Nikolaus P. Himmelmann & Björn Wiemer (eds.)

2004 *What makes Grammaticalization? A Look from its Fringes and its Components: A Look from Its Fringes and Its Components* (Trends in Linguistics). Berlin & New York: Mouton de Gruyter.

Blakemore, Diane

1987 *Semantic constraints on relevance.* Oxford: Blackwell.

Blakemore, Diane

1988a The organization of discourse. In Frederick J. Newmeyer (ed.), 229-250.

Blakemore, Diane

1988b Constraints on interpretation. In Kira Hall et al. (eds.), 363-370.

Blakemore, Diane

1990 *Relevance Relations in discourse. A study with special reference to Sissala*. Cambridge: Cambridge University Press.

Blakemore, Diane

2004 Discourse markers. In Lawrence Horn & Gregory Ward (eds), *The Handbook of Pragmatics*. Oxford: Blackwell, 221–40.

Blyth, Carl, Jr., Sigrid Recktenwald & Jenny Wang

1990 I'm like 'Say What?!'. A new quotative in American oral narrative. *American Speech* 65: 215-227.

Breivik, Leiv Egil & Ernst Håkon Jahr (eds.)

1989 *Language change: contributions to the study of its causes.* Berlin & New York: Mouton de Gruyter.

Brinton, Laurel

1996 *Pragmatic markers in English*. Berlin: Mouton de Gruyter.

Brinton, Laurel

2000 The Importance of Discourse Types in Grammaticalization: The Case of Anon. In Susan C. Herring, Pieter Th. van Reenen & Lene Schøsler (eds.), 139-162.

Brinton, Laurel & Elizabeth Traugott

2005 *Lexicalization and language change*. Cambridge: Cambridge University Press.

Britain, David

2002 Space and spatial diffusion. In Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), 603-637.

Brizuela, Maquela, Elaine Andersen, Lynne Stallings

1999 Discourse markers as indicators of register. *Hispania* 82.1: 128-142.

Buchstaller, Isabelle

2001 An alternative view of *like*: its grammaticalisation in conversational American English and beyond. *Edinburgh Working Papers in Applied Linguistics* 11: 21-41.

Buchstaller, Isabelle

2002 *He Goes and I'm Like: The New Quotatives Revisited*. Internet Proceedings of the University of Edinburgh Postgraduate Conference.

http://www.ling.ed.ac.uk/~pgc/archive/2002/proc02/buchstaller02.pdf. 28.8.2010.

Buchstaller, Isabelle

2006a Diagnostics of Age-Graded Linguistic Behaviour: The Case of the Quotative System. *Journal of Sociolinguistics* 10: 3–30.

Buchstaller, Isabelle

2006b Social Stereotypes, Personality Traits and Regional Perception Displaced: Attitudes Towards the 'New' Quotatives in the U.K. *Journal of Sociolinguistics* 10: 362–381.

Buchstaller, Isabelle

2008 The localization of global linguistic variants. *English World-Wide* 291: 15–44.

Buchstaller, Isabelle & Alexandra D'Arcy

2009 Localized globalization: A multi-local, multivariate investigation of quotative *be like*. *Journal of Sociolinguistics* 13.3: 291–331.

Bühler, Karl

1934 *Sprachtheorie*. Jena: Fischer.

Bybee, Joan

2002 Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14: 261–290.

Caron-Prague, Josiane & Caron, Jean

1991 Psychopragmatics vs. sociopragmatics: The function of pragmatic markers in thinking-aloud protocols. In Jef Verschueren (ed.), *Pragmatics at Issue: Selected Papers of the International Pragmatics Conference, Antwerp, August 17–22, 1987*. Amsterdam: John Benjamins, 29–36.

Cedergren, Henrietta

1973 *The interplay ofsocial and linguistic factors in Panama*. Unpublished Cornell University PhD dissertation.

Cedergren, Henrietta

1984 Panama Revisited: Sound Change in Real Time. Paper presented at NWAVE, Philadelphia, 1984.

Cedergren, Henrietta

1987 The Spread of Language Change: Verifying Inferences of Linguistic Diffusion. In Peter H. Lowenberg (ed.), *Language Spread and Language Policy: Issues, Implications and Case Studies. Georgetown University Round Table on Language and Linguistics.* Washington, D.C.: Georgetown University Press, 45-60.

Chambers, Jack K.

[1]1995 *Sociolinguistic Theory: Linguistic Variation and Its Social Significance*. Oxford: Blackwell.

Chambers, Jack K.

1998 Social embedding of changes in progress. *Journal of English Linguistics* 26, 1: 5-26.

Chambers, Jack K.

[2]2003 *Sociolinguistic Theory: Linguistic Variation and Its Social Significance*. Oxford: Blackwell.

Chambers, Jack K. & Peter Trudgill

[2]1998 *Dialectology*. Cambridge: Cambridge University Press.

Chambers, Jack K., Peter Trudgill and Natalie Schilling-Estes (eds.)

2002 *The Handbook of Language Variation and Change*. Oxford: Blackwell.

Chapman, Robert L. (ed.)

1986 *New dictionary of American slang*. New York: Harper & Row.

Cheshire, Jenny (ed.)

1991 *English around the World: Sociolinguistic Perspectives*. Cambridge: Cambridge University Press.

Cheshire, Jenny

1991 Introduction. In Jenny Cheshire (ed.), 1-12.

Coates, Jennifer (ed.)

1998 *Language and Gender: A Reader*. Oxford: Blackwell.

Columbus, Georgie

2009 Irish *like* as an invariant tag: evidence from ICE-Ireland. Paper presented at the AACL 2009 (American Association for Corpus Linguistics), October 9th 2009 in Edmonton, Alberta, Canada.

Coulmas, Florian (ed.)

1998 *The Handbook of Sociolinguistics*. Oxford: Blackwell.

Coupland, Nikolas, Anne Williams & Peter Garrett

1999 'Welshness' and 'Englishness' as attitudinal dimensions of English language varieties in Wales. In Dennis Preston (ed.), *Handbook of Perceptual Dialectology*. Philadelphia, Pennsylvania & Amsterdam: John Benjamins, 333-345.

Croucher, Stephen M.

2004a Like, You Know, What I'm Saying: A Study of Discourse Marker Frequency in Extemporaneous and Impromptu Speaking. (access via http: //cas.bethel.edu/dept/comm/nfa/journal/vol22no2-3.pdf, October 10th, 2010)

Croucher, Stephen M.

2004b I uh know what like you are saying: An analysis of discourse markers in limited preparation events. *National Forensics Journal*, 41.4: 38-52.

Crystal, David

1988 Another look at well, you know... *English Today* 13: 47-49.

Crystal, David & Derek Davy

1975 *Advanced conversational English*. London: Longman.

Dailey-O'Cain, Jennifer

2000 The sociolinguistic distribution of and subjective attitudes toward focuser like and quotative like. *Journal of Sociolinguistics* 4.1: 60-80.

D'Arcy, Alexandra

2005 *LIKE. Syntax and Development*. University of Toronto PhD dissertation.

D'Arcy, Alexandra

2006 Lexical Replacement and the Like(s). *American Speech* 81: 339–357.

D'Arcy, Alexandra

2007 LIKE and language ideology: disentangling the fact form fiction. *American Speech* 82: 386–419.

D'Arcy, Alexandra

2008 Canadian English as a window to the rise of 'like' in discourse. *Anglistik* (*International Journal of English Studies*) 19.2: 125-140. [Special issue; Meyer, Matthias L. G. (ed.), *Focus on Canadian English*. Heidelberg: Winter]

De Mauro, Tullio, Federico Mancini, Massimi Vedovello, and Miriam Voghera.

1993 *Lessico di frequenza dell'italiano parlato*. Milan: Etaslibri.

Denning, Keith M., Sharon Inkelas, Faye C. McNair-Knox & John R. Rickford (eds.)

1987 *Variation in Language: NWAVE-XV at Stanford*. Stanford: Department of Linguistics, Stanford University.

Diamond, S.J. [first name unavailable]

2000 Like It or Not, 'Like' Is Probably Here to Stay. You may call it 'Valley Girl' slang. Academics link the all-purpose verbal space filler with Americans' adolescent joy in sounding inarticulate. *Los Angeles Times*, August 21st. (access via http://articles.latimes.com/2000/aug/21/news/cl-7683/2, May 11th, 2011)

Dion, Nathalie & Poplack, Shana

2007 Linguistic mythbusting: The role of the media in diffusing change. Paper presented at the *New Ways of Analyzing Variation* (NWAVE) 36. University of Pennsylvania, Philadelphia, PA.

Diewald, Gabriele

1999 *Die Modalverben im Deutschen: Grammatikalisierung und Polyfunktionalität*. Tübingen: Niemeyer.

Diewald, Gabriele

2002 A model for relevant types of contexts in grammaticalization. New Reflections on Grammaticalization. In Ilse Wischer & Gabriele Diewald (eds.), 103-120.

Diewald, Gabriele

2010 On some problem areas in grammaticalization theory. In Katerina Stathi, Elke Gehweiler & Ekkehard König (eds.), *Grammaticalization: Current Views and Issues*. Amsterdam & Philadelphia: John Benjamins, 17-50.

Diewald, Gabriele

2011 Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics* 49.2: 365–390.

Downes, William

1998 *Language and Society*. Cambridge: Cambridge University Press.

Dubois Sylvie & Barbara M.Horvath

1999 When the music changes you change too: Gender and language change in Cajun English. *Language Variation and Change* 11.3: 287-313.

Du Bois, John W., Wallace L. Chafe, Charles Meyer & Sandra A. Thompson

2000 *Santa Barbara corpus of spoken American English, Part 1*. Philadelphia: Linguistic Data Consortium.

Du Bois, John W., Wallace L. Chafe, Charles Meyer, Sandra A. Thompson & Nii Martey

2003 *Santa Barbara corpus of spoken American English, Part 2*. Philadelphia: Linguistic Data Consortium.

Du Bois, John W. & Robert Englebretson

2004 *Santa Barbara corpus of spoken American English, Part 3*. Philadelphia: Linguistic Data Consortium.

Du Bois, John W. & Robert Englebretson

2005 *Santa Barbara corpus of spoken American English, Part 4*. Philadelphia: Linguistic Data Consortium.

Eckert, Penelope

1988 Adolescent social structure and the spread of linguistic change. *Language in Society* 17.2: 183–207.

Eckert, Penelope

1989 *Jocks and Burnouts: Social Categories and Identities in the High School.* New York: Teachers College Press.

Eckert, Penelope

1998 Gender and sociolinguistic variation. In Jennifer Coates (ed.), *Language and Gender: A Reader*. Oxford: Blackwell, 64-75.

Eckert, Penelope

2000 *Linguistic Variation as Social Practice*. Oxford: Blackwell.

Eckert, Penelope

2001 Style and social meaning. In Penelope Eckert & John Rickford (eds.), *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press, 206-219.

Eckert, Penelope & John Rickford (eds.),

2001 *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press.

Egil Breivik, Leiv & Ernst Håkon Jahr (eds.)

1989 *Language change: contributions to the study of its causes*. Berlin & New York: Mouton de Gruyter.

Eid, Michael, Mario Gollwitzer & Manfred Schmitt

2010 *Statistik und Forschungsmethoden*. Wentheim & Basel: Beltz.

Erman, Britt

1987 *Pragmatic expressions in English: A study of 'You know', 'you see' and 'I mean' in face-to-face conversation*. (Acta Universitatis Stockholmiensis, Stockholm Studies in English 69.) Stockholm: Almqvist & Wiksell.

Erman, Britt

2001 Pragmatic markers revisited with a focus on you know in adult and adolescent talk. *Journal of Pragmatics* 33.9: 1337–1359.

Ferrara, Kathleen & Barbara Bell

1995 Sociolinguistic variation and discourse function of constructed dialogue introducers: the case of *be* + *like*. *American Speech* 70: 265-290.

Fischer. Roswitha

1998 *Lexical Change in Present-Day English: A Corpus-Based Study of the Motivation, Institutionalization, and Productivity of Creative Neologisms.* Tübingen: Günther Narr.

Fleischman, Suzanne

1999 Pragmatic markers in comparative perspective: A contribution to cross-language pragmatics. Paper presented at the PRAGMA99, Tel Aviv, June 1999.

Fleischman, Suzanne & Marina Yaguello

2004 Discourse markers across languages. Evidence from English and French. In Carol Lynn Moder & Aida Martinovic-Zic (eds.). *Discourse across languages and cultures*. Amstrerdam & New York: John Benjamins.

Fox, Barbara A. & Jessica Robles

2010 It's like mmm: Enactments with it's like. *Discourse studies* 12.6: 715-738.

Fraser, Bruce

1987 Pragmatic formatives. In Jef Verschueren & Marcella Bertucelli-Papi (eds.), *The pragmatic perspective*. Amsterdam: John Benjamins, 179–194.

Fraser, Bruce

1988 Types of English discourse markers. *Acta Linguistica Hungarica* 38: 19-33.

Fraser, Bruce

1990 An approach to discourse markers. *Journal of Pragmatics* 14.3: 383-395.

Fraser, Bruce

1998 Contrastive discourse markers. In: Jucker, Andreas H. & Yael Ziv (eds.), *Discourse markers: descriptions and theory*. Amsterdam & New York: John Benjamins, 301-326.

Fraser, Bruce

1999 What are discourse markers? *Journal of Pragmatics* 31: 931-952.

Fuller, Janet M.

2003 Use of the discourse marker like in interviews. *Journal of Sociolinguistics* 7.3: 365-377.

Gauchat, Louis

1905 L'Unité phonétique dans le patois d'une commune. In Ernest Bovet, Ernst Brugger, Wilhelm Degen, Arturo Farinelli, Adolf Fluri, Louis Gauchat, Jkob Jud, Jules Jeanjaquet, Emil Keller, Martha Langkavel, Marie J. Minckwitz, Kaethe Schirmacher, Ernst Tappolet & Louis P. Betz (eds.), *Aus romanischen Sprachen und Literaturen: Festschrift Heinrich Morf*. Niemeyer: Halle.

Givón, Talmy

1971 'Historical syntax and synchronic morphology: An archeologist's field trip'. *Papers from the Chicago Linguistic Society* 7: 394-415.

Görlach, Manfred

1987 Colonial Lag? The Alleged Conservative Character of American English and Other 'Colonial' Varieties. *English World-Wide* 8: 41–60.

Goldberg, Julia Anna

1980 *Discourse particles: An analysis of the role of y'kmow, I mean, well, and actuall in conversation*. Unpublished Cambridge University PhD dissertation.

Golato, Andrea

2000 An innovative German quotative for reporting on embodied actions: Und ich so/und er so 'and I'm like/and he's like'. *Journal of Pragmatics* 32: 29-54.

Grant, William & James Main Dixon

1921 *Manual of Modern Scots*. Cambridge: Cambridge University Press.

Greenbaum, Sidney

1988 A proposal for an international computerized corpus of English. *World Englishes* 7.3: 315.

Gries, Stephan

2009 *Statistics for Linguistics with R. A Practical Introduction*. Berlin & New York: Mouton de Gruyter.

Gross, Thomas

2009 Bedeutende Laute. Na, ja und so: Linguisten diskutieren über Interjektionen. *Frankfurter Allgemeine Zeitung*, March 18th.

Günthner, Susanne & Katrin Mutz

2004 Grammaticalization vs. pragmaticalization? The development of pragmatic markers in German and Italian. In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.), 77-108.

Haeri, Niloofar

1987 Male/female differences in speech: an alternative interpretation. In Keith M. Denning et al. (eds.), *Variation in Language: NWAVE-XV at Stanford*. Stanford: Department of Linguistics, Stanford University, 173-182.

Haeri, Niloofar

1996 *The Sociolinguistic Market of Cairo: Gender, Class and Education*. London: Kegan Paul International.

Haiman, John

1989 Alienation in Grammar. *Studies in Language* 13: 129-170.

Hall, Kira, Jean Pierre Koenig, Michael Meacham, Sondra Reiman & Laurel A. Sutton (eds.)

1990 *Proceedings of the sixteenth annual meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society.

Halliday, M.A.K. & Ruqaiya Hasan

1976 *Cohesion in English*. (English Language Series 9) London: Longman.

Harris, John

1993 The grammar of Irish English. In James Milroy & Leslie Milroy (eds.), 139-186.

Hasund, Ingrid Kristine

2002 Congratulations, like! - Gratulerer, liksom! Pragmatic particles in English and Norwegian. In Leiv Egil Breivik & Angela Hasselgren (eds.), *From the COLT's Mouth ... and Others*. Amsterdam & New York: Rodopi, 125-140.

Hazen, Kirk

2010 Sociolinguistics in the United States of America. In Martin J. Ball (ed.), *The Routledge Handbook of Sociolinguistics Around the World*, Abingdon (UK) & New York: Routeledge, 7-24.

Hedevind, Bertil

1967 *The dialect of Dentdale in the West Riding of Yorkshire*. [Societas Anglistica Upsaliensis 5]. Uppsala: University of Uppsala.

Heine, Bernd

2002 On the role of context in grammaticalization. In Ilse Wischer & Gabriele Diewald (eds.), 83-102.

Heine, Bernd & Tania Kuteva

2002 *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva

2005 *Language Contact and Grammatical Change.* Cambridge: University of Cambridge Press.

Herring, Susan C., Pieter van Reenen & Lene Schøsler (eds.)

2000 *Textual Parameters in Older Languages*. Amsterdam & New York: John Benjamins.

Hickey, Raymond

2001 Language Change. In Jef Verschueren, Jan-Ola Östman & Jan Blommaert (eds.) *Handbook of pragmatics*. Amsterdam: John Benjamins, 1-35.

Hickey, Raymond

2002 *A Source Book for Irish English*. Amsterdam: John Benjamins.

Hickey, Raymond

2003 *Corpus Presenter. Processing Software for Language Analysis with a Manual and A Corpus of Irish English as Sample Data*. Amsterdam: John Benjamins.

Hickey, Raymond

2007 *Irish English. History and present-day forms*. Cambridge: Cambridge University Press.

Hjarvard, Stig

2004 The Globalization of Language. How the Media Contribute to the Spread of English and the Emergence of Medialects. *Nordicom Review* 1.2: 75-97.

Hölker, Klaus

1991 Französisch: Partikelforschung. *Lexikon der Romanistischen Linguistik*, Vol. 4.1: 77-88.

Holmes, Janet

1997 Setting new standards: Sound change and gender in New Zealand English. *English World-Wide* 18.1: 107-142.

Hopper, Paul & Elizabeth Closs Traugott

1993 *Grammaticalization*. Cambridge: Cambridge University Press.

Horvath, Barbara M.

1985 *Variation in Australian English*. Cambridge: Cambridge University Press.

House, Juliane & Gabriele Kasper

1981 Zur Rolle der Kognition in Kommunikationskursen. *Die Neueren Sprachen* 80: 42-55.

Hovy, Eduard H. 1994 The multifunctionality of discourse markers. (access via http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.141.4423&rep=rep1&type=pdf, March 10th, 2011).

Hudson, Rachel

²1996 *Sociolinguistics*. Cambridge: Cambridge University Press.

Iyeiri, Yoko, Michiko Yaguchi & Hiroko Okabe

2005 Gender and Style: The Discourse Particle like in the Corpus of Spoken Professional American English. *English Corpus Studies* 12: 37-52.

Jäger, Gerhard & Anette Rosenbach

2008a Priming and unidirectional language change. *Theoretical Linguistics* 34.2: 85-113.

Jäger, Gerhard & Anette Rosenbach

2008b Priming as a testing ground for historical linguists? – A reply to Chang, Eckardt, and Traugott. *Theoretical Linguistics* 34.2: 143-156.

James, Allan R.

1983 Compromisers in English: A cross-disciplinary approach to their interpersonal significance. *Journal of Pragmatics 7.2:* 191-206.

Jespersen, Otto

1954 *A Modern English Grammar on Historical Principles: Part 6 - Morphlogy*. London: Allen & Unwin.

Johnson, Kirk

1998 Correspondence/Duh; Today's Kids Are, Like, Killing the Language. Yeah, Right. *The New York Times*, August 9th.

Johnson, Daniel Ezra

2010 Why step-wise isn't so wise. Paper presented at NWAV 39 ("New" New Ways of Analyzing Variation), November 4-6, 2010 in San Antonio, Texas.

Jucker, Andreas & Yael Ziv (eds.)

1998 *Discourse markers. Description and Theory*. Amsterdam: John Benjamins.

Jucker, Andreas, Sarah W. Smith & Tanja. Lüdge

2003 Interactive aspects of vagueness in conversation. *Journal of Pragmatics* 35: 1737–1769.

Kachru, Braj B. (ed.)

[2]1992 *The Other Tongue: English across Cultures*. Urbana, Chicago: University of Illinois Press.

Kastronic, Laura

2011 *Discourse Like in Quebec English*. *University of Pennsylvania Working Papers in Linguistics* 17.2: 105-114.

Keller, Eric

1979 Gambits: Conversational strategy signals. *Journal of Pragmatics* 3.3-4: 219-238.

Knott, Alistair & Robert Dale

1994 Using Linguistic Phenomena to Motivate a Set of Coherence Relations. *Discourse Processes* 18.1: 35-62.

Kortmann, Bernd

2004 Synopsis: morphological and syntactic variation in the British Isles. In Bernd Kortmann et al. (eds.), 1089–1103.

Kortmann, Bernd, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton (eds.)

2004 *A Handbook of Varieties of English. Volume 1: Phonology, Volume 2: Morphology and Syntax*. Berlin & New York: Mouton de Gruyter.

Kortmann, Bernd & Benedikt Szmrecsanyi

2004 Global synopsis: Morphological and syntactic variation in English. In Bernd Kortmann et al. (eds.), 1142–1202.

Kortmann, Bernd & Kerstin Lunkenheimer (eds.)

2011 *The electronic World Atlas of Varieties of English [eWAVE]*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (access via http://www.ewave-atlas.org/, April 9th, 2012)

Kotsinas, Ulla-Britt

1994 *Ungdomsspråk*. Uppsala: Hallgren & Fallgren.

Kroch, Anthony S.

1978 Toward a theory of social dialect variation. *Language in Society* 7: 17-36.

Kroon, Caroline H. M

1995 *Discourse Particles in Latin. A Study of* nam*,* enim*,* autem*,* vero *and* at. Amsterdam: Gieben.

Labov, William

1963 The social motivation of a sound change. *Word* 19: 273-309.

Labov, William

1966 *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.

Labov, William

1972 *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William

1982 Building on empirical foundations. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Perspectives on Historical Linguistics*. Amsterdam: John Benjamins, 17-92.

Labov, William

1990 The intersection of sex and social class in the course of linguistic change. *Language and Variation* 2: 205-251.

Labov, William

1994 *Principles of Linguistic Change. Vol. 1, Internal Factors*. Oxford: Blackwell.

Labov, William

2001 *Principles of Linguistic Change. Vol. 2, Social Factors*. Oxford: Blackwell.

Labov, William

2002 Driving forces in linguistic change. Paper presented at the *International Conference on Korean Linguistics*.

Labov, William

2003 Pursuing the Cascade Model. In Britain, David & Jenny Cheshire (eds.), *Social Dialectology: Honour of Peter Trudgill.* Amsterdam: John Benjamins, 9–22.

Labov, William

2007 Transmission and Diffusion. *Language* 83.2: 344-387.

Labov, William

2010 *Principles of Linguistic Change, Volume 3: Cognitive and Cultural Factors.* Oxford: Blackwell.

Labov, William, Sharon Ash & Charles Boberg

2006 *The Atlas of North American English: Phonetics, phonology and sound change.* Berlin: Mouton de Gruyter.

Lakoff, George

1987 *Women, Fire, and Dangerous Things*. Chicago & London: University of Chicago Press.

Lehmann, Christian

[2]1982 *Thoughts on Grammaticalization*. Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt: Erfurt. (access via http://www.scribd.com/doc/24853253/Thoughts-on-grammaticalization-Christian-Lehmann, March 10th, 2011)

Lehmann, Christian

1991 Grammaticalization and related changes in contemporary German, In Elizabeth Closs Traugott & Bernd Heine (eds.), 493-535.

Lehmann, Winfred P & Yakov Malkiel (eds.)

1982 *Perspectives on Historical Linguistics*. Amsterdam: John Benjamins.

Lenk, Uta

1997 *Marking discourse coherence: functions of discourse markers in English*. Tübingen: Narr.

Leuschner, Torsten, Tanja Mortelmans & Sarah de Groodt (eds.)

2005 *Grammatikalisierung im Deutschen*. Berlin & New York: Mouton de Gruyter.

Levey, Bob

1999 It's, like, an answer to like addiction. *The Washington Post*, September 13th.

Levey, Stephen

2003 He's like 'Do it now!' and I'm like 'No!': Some Innovative Quotative Usage among Young People in London. *English Today* 19: 24–32.

Levey, Stephen

2006 The sociolinguistic distribution of discourse marker *like*. *Multiulingua* 25: 413-441.

Le Page, R. B. [first name not available]

1968 Problems of the descriptiuon of multilingual communities. *Transactions of the Philological Society for 1968* 67.1: 189-212.

Le Page, R. B. [first name not available]

1998 The Evolution of a Sociolinguistic Theory of Language. In Jennifer Coulmas (ed.), 15-32.

Le Page, R. B. [first name not available] & Andrée Tabouret-Keller

1985 *Acts of Identity*. Cambridge: Cambridge University Press.

Macaulay, Ronald K.S.

1977 *Language, Social Class and Education: A Glasgow Study*. Edinburgh: University of Edinburgh Press.

Macaulay, Ronald K.S.

1978 Variation and consistency in Glaswegian English. In Peter Trudgill (ed.), *Sociolinguistic Patterns in British English*. London: Arnold, 132-143.

Macaulay, Ronald K.S.

2001 You're like 'why not?' The quotative expressions of Glasgow adolescents. *Journal of Sociolinguistics* 5: 3-21.

Macaulay, Ronald K.S.

2005 *Talk That Counts. Age, Gender and Social Class Differences in Discourse*. Oxford: Oxford University Press.

Mair, Christian

2009 Corpus linguistics meets sociolinguistics. Studying educated spoken usage in Jamaican on the basis of the International Corpus of English. In Thomas Hoffmann & Lucia Siebers (eds.), *World Englishes – Problems, Properties and Prospects*. Amsterdam: John Benjamins, 39-60.

Maschler, Yael

2001 Veke'ilu haraglayim sh'xa nitka'ot bifnim kaze ('and Like Your Feet Get Stuck Inside Like'): Hebrew kaze ('Like'), ke'ilu ('Like'), and the Decline of Israeli Dugri ('Direct') Speech. *Discourse Studies: An Interdisciplinary Journal for the Study of Text and Talk* 3: 295-326.

McArthur, Tom

1998 *The English Languages*. Oxford: Oxford University Press.

McArthur, Tom

2002 *Oxford Guide to World English*. Oxford: Oxford University Press.

Meehan, Teresa

1991 It's like, 'What's happening in the evolution of like?' A theory of grammaticalization. *Kansas Working Papers in Linguistics* 16: 37-51.

Mehren, Elizabeth

1999 Colleagues, like, focus on speech. *The Los Angeles Times*, March 22nd.

Meyerhoff, Miriam & Nancy Niedzielski

1998 The syntax and semantics of *olsem* in Bislama. In: Pearson, Matthew (ed.), *Recent Papers in Austronesian Linguistics UCLA Occasional papers in Linguistics 21*. Los Angeles: UCLA Department of Linguistics, 235-243.

Meyerhoff, Miriam & Nancy Niedzielski

2003 The globalization of vernacular variation. *Journal of Sociolinguistics* 7: 534-555.

Miller, Casey & Kate Swift

1976 *Words and Women*. Garden City, NY: Anchor Press.

Miller, Jim

2004 Scottish English: morphology and syntax. In Bernd Kortmann et al. (eds.), 47–72.

Miller, Jim

2009 *Like* and other discourse markers. In Pam Peters, Peter Collins & Adam Smith (eds.), *Comparative Studies in Australian and New Zealand English: Grammar and beyond*. 2009. Amsterdam: John Benjamins, 317-338.

Miller, Jim & Regina Weinert

1995 The function of *like* in dialogue. *Journal of Pragmatics* 23: 365-93.

Miller, Jim & Regina Weinert

1998 *Spontaneous Spoken Language: Syntax and Discourse*. Oxford: Clarendon Press.

Milroy, Lesley

2004 Language Ideologies and Linguistic Change. In Carmen Fought (ed.), *Sociolinguistic Variation: Critical Reflections*. New York: Oxford University Press, 161–77.

Milroy, Lesley & Matthew Gordon

2003 *Sociolinguistics: method and interpretation*. Oxford: Blackwell.

Milroy, James & Lesley Milroy

1978 Belfast: Change and variation in an urban vernacular. In Peter Trudgill (ed.), *Sociolinguistic Patterns in British English*. London: Edwin Arnold, 19–36.

Milroy, James & Lesley Milroy

1993a Mechanisms of Change in Urban Dialects: The Role of Class, Social Network and Gender. *Journal of Applied Linguistics* 3.1: 57-77.

Milroy, James & Lesley Milroy (eds.)

1993b *Real English: The Grammar of English Dialects in the British Isles*. London & New York: Longman.

Moder, Carol Lynn & Aida Martinovic-Zic (eds.)

2004 *Discourse across languages and cultures*. Amstrerdam & New York: John Benjamins.

Modaressi, Yahya

1978 *A sociolinguistic investigation of modern Persian*. University of Kansas PhD dissertation.

Moravcsik, Edith A.

1978 Language contact. In Joseph H. Greenberg (ed.), *Universals of human language, vol.1*. Stanford CA: University of Stanford Press, 93-122.

Morris, William (ed.)

1969 *American Heritage Dictionary of the English Language*. Boston: Houghton Mifflin.

Müller, Simone

2005 Discourse *markers in Native and Non-native English Discourse*. Ansterdam & New York: John Benjamins.

Mufwene, Salikoko S.

2001 The Ecology of Language Evolution. Cambridge: Cambridge University Press.

Muhr, Rudolf

2003 Language change via satellite: The influence of German television broadcasting on Austrian German. *Journal of Historical Pragmatics* 4.1: 103-128. [Special Issue; Herring, Susan C. (ed.), Media and Language Change]

Nevalainen, Terrtu & Helena Raumolin-Brunberg

2003 *Sociolinguistics: Language Change in Tudor and Stuart England* (Longman Linguistics Library). Harlow: Pearson Education.

Newman, Edwin

1974 *Strictly speaking: will America be the death of English?*. Bobbs-Merrill: Indianapolis.

Newmeyer, Frederick J. (ed.)

1988 *Linguistics: The Cambridge survey. Vol. IV; Language: The cultural context.* Cambridge: Cambridge University Press.

Newmeyer, Frederick J.

1998 *Language form and language function*. Cambridge, MA: MIT Press.

Nordberg, Bengt & Eva Sundgren

1998 *On observing real-time language change: A Swedish case study*. [FUMS report no. 190]. Uppsala: Enheten för sociolingvistik, Institutionen för Nordiska Språk vid Uppsala Universitet.

Östman, Jan-Ola

1982 The symbiotic relationship between pragmatic particles and impromptu speech. In Nils Erik Enkvist (ed.), *Impromptu speech: A symposium*. (Meddelanden fran Stiftelsens for Abo Akademi Forskningsinstitut 78.) Abo: Abo Akademi, 147-177.

Östman, Jan-Ola

1995 Pragmatic particles twenty years after. In Brita Warvik, Sanna-Kaisa Tansbnen, Risto Hiltunen (eds.), *Organization in Discourse. Proceedings from the Turku Conference*. (Anglicana Turkuensia 14). Turku: University of Turku, 95-108.

O'Keeffe, Anne & Michael McCarthy (eds.)

2010 *The Routledge Handbook of Corpus Linguistics*. Abingdon (UK) & New York: Routeledge.

Onysko, Alexander

2007 *Anglicisms in German. Borrowing, Lexical Productivity, and Written Codeswitching.* Berlin: Mouton de Gryter.

Osmers-Gordon, Elisabeth

2008 LIKE it's versatile. *The Press,* November 8th.

Overstreet, Maryann

1999 *Whales, Candlelight, and Stuff Like That. General extenders in English discourse*. Oxford: Oxford University Press

Oxford English Dictionary

1989 *Oxford English Dictionary*. Oxford: Clarendon.

Palomares, Nicholas A.

2008 Explaining Gender-Based Language Use: Effects of Gender Identity Salience on References to Emotion and Tentative Language in Intra- and Intergroup Contexts. *Human Communication Research* 34: 263–286.

Partridge, Eric

1984 *A dictionary of slang and unconventional English*. London: Routledge & Kegan Paul.

Pentillä, Esa & Heli Paulasto (eds.)

2009 *Language Contacts meet English Dialects. Studies in Honour of Markku Filppula*. Cambridge: Cambridge Scholars Publishing.

Peters, Mark

2008 Language: Like, Yikes! The misunderstood word 'like' is not just for teens or women. *Pyschology Today*, May 1st.

Petersen, A.

2004 It's, Like, Taking over the Entire Planet: Young Adults Abandon the Speech Crutch 'LIKE'. *Wall Street Journal*, February 9th.

Pichler, Heike & Stephen Levey

2010 Variability in the Co-occurrence of Discourse Features. *University of Reading Language Studies Working Papers* 2: 17-27.

Pickering, Martin J. & Garrod, Simon

2004 Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27: 169–226.

Platt, John, Heidi Weber & Ho Mian Lian (eds.)

1984 *The New Englishes*. London & Melbourne: Routledge & Kegan Paul.

Pride, John (ed.)

1982 *New Englishes*. Rowle. London & Tokyo: Newbury House.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik

1985 *A Comprehensive Grammar of the English language*. London: Longman.

R Development Core Team (R Foundation for Statistical Computing)

2010 R: A Language and Environment for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0. (access via http: //www.R-project.org, August 17th, 2010)

Randall, Bernice

1988 *Webster's New World Guide to Current American Usage*. New York: Simon & Schuster.

Redeker, Gisela

1990 Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics* 14.3: 367-381.

Redeker, Gisela

1991 Review of Schiffrin 1987. Linguistics 29: 1139-1172.

Romaine, Suzanne

1994 *Language in Society*. Oxford: Oxford University Press.

Romaine, Suzanne

2005 Historical sociolinguistics. Historische Soziolinguistik. In Peter Trudgill, Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier (eds.), *Sociolinguistics/ Soziolinguistik. An International Handbook of the Science of Language and Society/ Ein Internationales Handbuch Zur Wissenschaft Von Sprache Und Gesellschaft.* Berin & New York: Mouton de Gruyter, 1696-1703.

Romaine, Suzanne & Deborah Lange

1991 The use of *like* as a marker of reported speech and thought. A case of grammaticalization in process. *American Speech* 66.3: 227-79.

Ross, John Robert & William E. Cooper

1979 Like syntax. In William E. Cooper & Edward C. T. Walker (eds.), *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*. New York: Erlbaum Associates, 343-418.

Sankoff, Gillian

2004 Adolescentes, young adults and the critical period: Two case studies from 'Seven Up'. In Carmen Fought (ed.), *Sociolinguistic variation: Critical Reflections*. Oxford: Oxford University press, 638-668.

Sankoff, Gillian, Pierrette Thibault, Naomi Nagy, Helene Blondeau, Marie-Odile Fonollosa & Lucie Gagnon

1997 Variation in the use of discourse markers in a language contact situation. *Language Variation and Change* 9: 191-217.

Sankoff, Gillian, Hélène Blondeau & Anne Charity

2001 Individual roles in a real-time change: Montreal (r>R) 1947–1995. 'r-atics: Sociolinguistic, phonetic and phonological characteristics of /r/. *Etudes & Travaux* 4: 141–58 (special issue, eds. Hans van de Velde and Roeland van Hout).

Schiffrin, Deborah

1986 Functions of 'and' in discourse. *Journal of Pragmatics* 10.1: 41-66.

Schiffrin, Deborah

1987 *Discourse markers*. Cambridge: Cambridge University Press.

Schiffrin, Deborah

2001 Discourse markers. In Deborah Schiffrin, Deborah Tannen & Heidi E. Hamilton (eds.), *The Handbook of Discourse Analysis*. Oxford: Blackwell, 54-76.

Schilling-Estes, Natalie

1998 Investigating 'self-concious' speech: The performance register in Ocracoke English. *Language in Society* 27: 53-83.

Schilling-Estes, Natalie

1999 Reshaping economies, reshaping identities: Gender-based patterns of language variation in Ocracoke English. In Suzanne Wertheim, Ashlee C. Bailey & Monica Corston-Oliver (eds.), *Engendering Communication: Proceedings of the Fifth Berkeley Women and Language Conference.* Berkeley: Berkeley Women and Language Group, 509-520.

Schleef, Erik

2004 Gender, Power, Discipline, and Context: On the Sociolinguistic Variation of *okay*, *right*, *like*, and *you know* in English Academic Discourse. *Texas Linguistic Forum* 48: 177-186.

Schneider, Edgar W.

2003 The dynamics of New Englishes. Form identity construction to dialect birth. *Language* 79: 233-281.

Schneider, Edgar W.

2004 Synopsis: morphological and syntactic variation in the Americas and the Caribbean. In Bernd Kortmann et al. (eds.), 1104–1115.

Schneider, Edgar W.

2007 *Postcolonial English. Varieties around the World*. Cambridge: Cambridge University Press.

Schourup, Lawrence C.

1982 *Common Discourse Particles in English Conversation*. Ohio State University PhD dissertation.

Schourup, Lawrence C.

1985 *Common Discourse Particles in English Conversations "Like", "Well", "y'know"*. New York & London: Garland.

Schourup, Lawrence C.

1999 Discourse makers. Tutorial overview. *Lingua* 107: 227-265.

Schweinberger, Martin

2010 Epistemic Modality and Discourse Markers. A quantitative analysis of the distribution and function of discourse marker LIKE across varieties of English. Paper presented at ModE4 (Modality in English 4), Madrid, 9.–11.9.2010.

Schweinberger, Martin

2012 The discourse marker LIKE in Irish English. In Bettina Migge & Maire Ni Chiosain (eds.), *New Perspectives on Irish English*. Amsterdam: John Benjamins.

Schweinberger, Martin

2013 A sociolinguistics analysis of discourse marker LIKE in Northern Irish English. A look behind the scenes of quantitative reasoning. In Markus Bieswanger & Amei Koll-Stobbe (eds.), *New Approaches to the Study of Linguistic Variability*. Frankfurt am Main: Lang.

Sharifian, Farzad & Ian G. Malcolm

2003 The pragmatic marker like in English teen talk: Australian Aboriginal usage. *Pragmatics and Cognition* 11.2: 327-344.

Siegel, Muffy E. F.

2002 Like: The Discourse Particle and Semantics. *Journal of Pragmatics* 19: 35-71.

Siemund, Peter, Georg Maier & Martin Schweinberger

2009 Towards a more fine-grained analysis of the areal distributions of non-standard features of English. In Esa Pentillä & Heli Paulasto (eds.), 19-46.

Sperber, Dan & Deirdre Wilson

1986 *Relevance: communication and cognition*. Oxford: Blackwell.

Steger, Manfred B.

2003 *Globalization. A Very Short Introduction*. Oxford: Oxford University Press.

Streek, Jürgen

2002 Grammars, words, and embodied meanings: On the uses and evolution of *so* and *like*. *Journal of Communication* 52.3: 581-596.

Stubbs, Michael

1983 *Discourse analysis: The sociolinguistic analysis of natural language*. (Language in Society Series 4.) Oxford: Blackwell.

Sture Urleand, P.

1989 Some contact structures in Scandinavian, Dutch and Raeto-Romansh: inner-linguistic and/or contact causes of language change. In Leiv Egil Breivik & Ernst Håkon Jahr (eds.), *Language change: contributions to the study of its causes*. Berlin & New York: Mouton de Gruyter, 239-276.

Sturtevant, Edgar

1947 *An Introduction to Linguistic Science*. New Haven: Yale University Press.

Szmrecsanyi, Benedikt

2006 *Morphosyntactic Persistence in Spoken English - A Corpus Study at the Intersection of Variationist Sociolinguistics, Psycholinguistics, and Discourse Analysis*. Berlin & New York: Mouton de Gruyter.

Tagliamonte, Sali

2005 So who? Like how? Just what? Discourse markers in the conversation of young Canadians. *Journal of Pragmatics* 37: 1896-1915.

Tagliamonte, Sali

2012 *Variationist Sociolinguistics. Change, Observation, Interpretation.* Malden & Oxford: Wiley-Blackwell.

Tagliamonte, Sali & Alex D'Arcy

2004 He's like, She's like: The quotative system in Canadian youth. *Journal of Sociolinguistics* 8: 493-514.

Tagliamonte, Sali & Alexandra D'Arcy

2007 Frequency and Variation in the Community Grammar: Tracking a New Change through the Generations. *Language Variation and Change* 19: 199–217.

Tagliamonte, Sali A. & Alexandra D'Arcy

2009 Peaks Beyond Phonology: Adolescence, Incrementation, and Language Change. *Language* 85.1: 58-108.

Tagliamonte, Sali & Derek Denis

2010 The Stuff of Change: General Extenders in Toronto, Canada. *Journal of English Linguistics* 38.4: 335-368.

Tagliamonte, Sali & Rachel Hudson

1999 Be like et al. beyond America: The quotative system in British and Canadian youth. *Journal of Sociolinguistics* 3: 147-172.

Tagliamonte, Sali & Chris Roberts 2005 So cool, so weird, so innovative! The use of intensifiers in the television series *Friends*. *American Speech* 80.3: 280–300.

Tarde, Gabriel

1873 *Les Lois d'imitation*. English translation, New York: Henry Holt, 1903.

Thomason, Sarah G. & Terrence Kaufman

1991 *Language Contact, Creolization, and Genetic Linguistics*. Berkeley & Los Angeles: University of California Press.

Traugott, Elizabeth Closs & Bernd Heine (eds.)

1991 *Approaches to Grammaticalization, 2 Vols*. Amsterdam: John Benjamins.

Traugott, Elizabeth Closs & Ekkehard König

1991 The semantics-pragmatics of grammaticalization revisited. In Elizabeth Closs Traugott &Bernd Heine (eds.), 189-218.

Traugott, Elizabeth Closs

1982 From prepositional to textual to expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In Winfred Lehmann &

Yakov Malkiel (eds.), *Perspectives on Historical Linguistics*. Amsterdam: Benjamins, 245-271.

Traugott, Elizabeth Closs

1995 The role of the development of discourse markers in a theory of grammaticalization. Paper presented at the 12th International Conference on Historical Linguistics, Manchester, August 1995.

Traugott, Elizabeth Closs

1999 The rhetoric of counter-expectation in semantic change: A study in subjectification. *Historical Semantics and Cognition* 12: 1-12.

Traugott, Elizabeth Closs

2008 Testing the hypothesis that priming is a motivation for change. *Theoretical Linguistics* 34.2: 135–142.

Trudgill, Peter

1974 *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Trudgill, Peter

1988 Norwich revisited: recent linguistic changes in an English urban dialect. *English World Wide* 9: 33-49.

Trudgill, Peter

1999 A Window on the Past: 'Colonial Lag' and New Zealand Evidence for the Phonology of Nineteenth-Century English. *American Speech* 74, 3: 227-239.

Underhill, Robert

1988 Like Is, Like, Focus. *American Speech*, 63.3: 234-246.

Valentine, Tamara M.

1991 Getting the Message Across: Discourse Markers in Indian English. *World Englishes* 10.3: 325-334.

Watts, Richard J.

1988 A relevance-theoretic approach to commentary pragmatic markers: The case of actually, really and basically. *Acta linguistica Hungaria* 38: 235-260.

Weinreich, Uriel, William Labov & Marvin Herzog

1968 Empirical foundations for a theory of language change. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Directions for Historical Linguistics*. Austin: University of Texas Press, 97-188.

Weinreich, Uriel

1953 *Languages in Contact. Findings and Problems* [Publications of the Linguistic Circle of New York 1]. New York, NY: Linguistic Circle of New York.

Wentworth, Harold & Stuart Berg Flexner

1967 *Dictionary of American Slang*. London: George G. Harap & Co.

Wertheim, Suzanne, Ashlee C. Bailey & Monica Corston-Oliver (eds.)

1999 *Engendering Communication: Proceedings of the Fifth Berkeley Women and Language Conference.* Berkeley: Berkeley Women and Language Group.

White, William

1955 Wayne University slang. *Arnerican Speech* 30.4: 301-305.

Whitney, William Dwight

1904 *Language and the Study of Language: Twelve Lectures on the Principle of Language Science*. New York: Charles Scribner & Co.

Wierzbicka, Anna

1986a Introduction. *Journal of Pragmatics* 10.5: 519-534.

Wierzbicka, Anna

1986b Precision in vagueness: The semantics of English 'approximatives'. *Journal of Pragmatics* 10.5: 597-614.

Williams, Angie, Peter Garrett &Nikolas Coupland

1999 Dialect recognition. In Dennis Preston (ed.), *Handbook of Perceptual Dialectology*. Philadelphia & Amsterdam: John Benjamins, 345-359.

Wilson, Sir James

1915 *Lowland Scotch as spoken in the Lower Strathearn district of Perthshire*. London: Oxford University Press.

Wischer, Ilse & Gabriele Diewald (eds.)

2002 *New reflections on Grammaticalization*. Amsterdam & Philadelphia: John Benjamins.

Wischer, Ilse & Gabriele Diewald

2002 Introduction. In Wischer, Ilse & Gabriele Diewald (eds.), ix-xiv.

Wodak, Ruth & Gertraud Benke

1998 Gender as a Sociolinguistic variable: New Perspectives on Variation studies. In Florian Coulmas (ed.), *The Handbook of Sociolinguistics*. Oxford: Blackwell.

Wolfram, Walt

2008 When islands lose dialects. The Case of the Ocracoke Brogue*. Shima: The International Journal of Research into Island Cultures* 2.1: 1-13.

Wolfram, Walt & Natalie Schilling-Estes

1995 Moribund dialects and the endangerment canon: The case of the Ocracoke brogue. *Language* 71: 696-721

Wright, Joseph

1902 *The English dialect dictionary. Volume IV*. London: Henry Frowde.

Wright, Thomas

1857 *Dictionary of obsolete and provincial English. Volume II.* London: Henry G. Bohn.

11 Appendix

Table 137: Summary of significant and insignificant variables in the present study

Variety of English	Type	PAI	Age	Gender	Occupation	L1	Ethnicity	Region	Date of data compilation
American	ALL	-	√	X	X	-	X	X	-
	INI	√	X	X	X	-	X	X	-
	MED	√	√	X	X	-	X	X	-
	NON	√	√	X	√	-	X	X	-
Canadian	ALL	-	√	X	√	X	-	-	X
	INI	√	√	√	√	X	-	-	X
	MED	√	√	√	X	X	-	-	X
	NON	√	X	X	√	X	-	-	X
British	ALL	-	X	√	X	-	-	-	-
Indian	ALL	-	√	√	X	X	-	-	?
	INI	√	√	√	X	X	-	-	?
	MED	X	√	√	X	X	-	-	?
	FIN	√	√	√	X	X	-	-	?
	NON	√	√	√	X	X	-	-	?
Irish	ALL	-	√	X	X	-	-	X	√
	INI	√	√	X	X	-	-	X	√
	MED	√	√	√	X	-	-	√	√
	FIN	√	√	√	X	-	-	X	√
	NON	√	X	X	√	-	-	X	√
Jamaican	ALL	-	√	X	√	√	-	-	√
	INI	√	√	X	X	X	-	-	√
	MED	√	√	X	X	√	-	-	√
	NON	√	√	X	X	X	-	-	X
New Zealand	ALL	-	√	√	X	-	X	-	-
	INI	√	√	√	X	-	X	-	-
	MED	√	√	√	√	-	X	-	-
	FIN	-	√	X	-	-	-	-	-
	NON	√	√	X	X	-	X	-	-
Filipino	ALL	-	√	X	X	X	-	-	√
	INI	-	√	√	-	-	-	-	X
	MED	√	√	√	X	X	-	-	√
	NON	√	√	X	X	X	-	-	√

Table 138: Final data set of AmE

ID	VAR	FILE	SPEAKER	SEX	age.orig	AGEG	AGE	REG	REL	ETH	EDU	OCC	PAI	words	ini.n	med.n	fin.m	non.n	na.n	all.n	all	ini	med	fin	non
2051	USA	sbc0001	D	F	50	4	50	OTH	NA	WHI	SEC	SLM	0.190	717	1	0	0	0	0	1	1.395	1.395	0.000	0.000	0.000
2052	USA	sbc0001	Y	F	19	1	19	NA	NA	WHI	SEC	EDU	0.190	4543	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2053	USA	sbc0002	H	NA	NA	NA	NA	NA	NA	NA	NA	NA	3.705	1143	0	5	0	4	0	9	7.874	0.000	4.374	0.000	3.500
2054	USA	sbc0002	J	F	30	2	30	CA	NA	WHI	UNI	ACD	3.705	1399	1	3	0	1	0	5	3.574	0.715	2.144	0.000	0.715
2055	USA	sbc0002	M	M	NA	NA	NA	NA	NA	NA	NA	NA	3.705	2692	5	4	0	2	0	11	4.086	1.857	1.486	0.000	0.743
2056	USA	sbc0002	P	M	36	3	36	CA	NA	WHI	NA	EDU	3.705	547	0	0	0	1	0	1	1.828	0.000	0.000	0.000	1.828
2059	USA	sbc0003	M	F	33	2	33	CA	NA	WHI	NA	NA	5.414	2686	0	3	0	2	1	6	2.234	0.000	1.117	0.000	0.745
2057	USA	sbc0003	P	M	36	3	36	CA	NA	WHI	NA	EDU	5.414	1392	2	7	0	3	0	12	8.621	1.437	5.029	0.000	2.155
2058	USA	sbc0003	R	M	34	3	34	CA	NA	WHI	NA	CLM	5.414	1748	1	2	0	2	0	5	2.860	0.572	1.144	0.000	1.144
2060	USA	sbc0004	C	F	19	1	19	OTH	NA	WHI	SEC	EDU	9.093	1340	0	4	0	0	0	4	2.985	0.000	2.985	0.000	0.000
2061	USA	sbc0004	K	F	31	2	31	CA	NA	WHI	NA	EDU	9.093	836	0	0	0	3	0	3	3.589	0.000	0.000	0.000	3.589
2062	USA	sbc0004	S	F	24	1	24	OTH	NA	WHI	UNI	ACD	9.093	2903	11	12	0	7	1	31	10.679	3.789	4.134	0.000	2.411
2066	USA	sbc0005	D	M	33	2	33	CA	NA	WHI	UNI	CLM	0.910	1230	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2067	USA	sbc0005	P	F	38	3	38	CA	NA	WHI	UNI	CLM	0.910	2065	0	2	0	1	0	3	1.453	0.000	0.969	0.000	0.484
2069	USA	sbc0006	A	F	34	3	34	CA	NA	WHI	UNI	CLM	2.734	6076	3	12	0	0	1	16	2.633	0.494	1.975	0.000	0.000
2068	USA	sbc0006	L	F	30	2	30	CA	NA	WHI	UNI	EDU	2.734	874	1	1	0	1	0	3	3.432	1.144	1.144	0.000	1.144
2070	USA	sbc0007	A	F	28	2	28	OTH	NA	OTH	UNI	EDU	3.146	1773	3	4	0	1	2	10	5.640	1.692	2.256	0.000	0.564
2071	USA	sbc0007	M	F	27	2	27	OTH	NA	OTH	UNI	CLM	3.146	1723	0	0	0	1	0	1	0.580	0.000	0.000	0.000	0.580
2075	USA	sbc0008	A	M	NA	NA	NA	CA	NA	WHI	SEC	NA	5.520	151	0	1	0	0	0	1	6.623	0.000	6.623	0.000	0.000
2074	USA	sbc0008	R	F	31	2	31	CA	NA	WHI	UNI	ACD	5.520	2924	2	1	0	1	0	4	1.368	0.684	0.342	0.000	0.342
2072	USA	sbc0008	Y	F	NA	NA	NA	CA	NA	OTH	SEC	CLM	5.520	2511	7	8	1	8	2	26	10.354	2.788	3.186	0.398	3.186

2076	USA	sbc0009	K	F	17	1	17	OTH	NA	WHI	SEC	EDU	3.340	1479	0	3	0	2	0	5	3.381	0.000	2.028	0.000	1.352
2077	USA	sbc0009	N	M	19	1	19	OTH	NA	WHI	SEC	EDU	3.340	1515	1	2	0	2	0	5	3.300	0.660	1.320	0.000	1.320
2078	USA	sbc0010	B	M	45	4	45	NA	NA	WHI	UNI	ACD	0.693	1384	0	0	0	1	0	1	0.723	0.000	0.000	0.000	0.723
2079	USA	sbc0010	P	M	30	2	30	OTH	NA	OTH	UNI	CLM	0.693	2945	2	0	0	0	0	2	0.679	0.679	0.000	0.000	0.000
2081	USA	sbc0011	A	F	90	4	90	OTH	NA	WHI	UNI	ACD	0.455	1518	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2080	USA	sbc0011	D	F	83	4	83	OTH	NA	WHI	UNI	ACD	0.455	1598	0	0	0	1	0	1	0.626	0.000	0.000	0.000	0.626
2082	USA	sbc0011	S	F	72	4	72	OTH	NA	WHI	SEC	NA	0.455	599	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2088	USA	sbc0012	C	F	18	1	18	CA	NA	WHI	SEC	EDU	0.404	126	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2090	USA	sbc0012	F	M	24	1	24	CA	NA	WHI	UNI	CLM	0.404	1256	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2087	USA	sbc0012	G	M	22	1	22	CA	NA	OTH	SEC	EDU	0.404	217	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2085	USA	sbc0012	M	M	51	4	51	CA	NA	OTH	UNI	CLM	0.404	2355	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2092	USA	sbc0012	R	M	27	2	27	CA	NA	OTH	UNI	ACD	0.404	120	0	0	0	1	0	1	8.333	0.000	0.000	0.000	8.333
2099	USA	sbc0013	K	M	26	2	26	OTH	NA	WHI	UNI	SLM	1.249	1492	0	2	0	2	0	4	2.681	0.000	1.340	0.000	1.340
2097	USA	sbc0013	M	F	50	4	50	OTH	NA	WHI	UNI	ACD	1.249	1806	0	2	0	1	0	3	1.661	0.000	1.107	0.000	0.554
2098	USA	sbc0013	W	F	26	2	26	OTH	NA	WHI	UNI	SLM	1.249	2054	0	5	0	1	0	6	2.921	0.000	2.434	0.000	0.487
2096	USA	sbc0013	X	M	51	4	51	OTH	NA	WHI	UNI	ACD	1.249	2749	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2101	USA	sbc0014	F	M	47	4	47	OTH	NA	WHI	UNI	ACD	0.703	1598	0	1	0	0	0	1	0.626	0.000	0.626	0.000	0.000
2102	USA	sbc0014	J	M	45	4	45	OTH	NA	WHI	UNI	CLM	0.703	975	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2103	USA	sbc0014	K	M	70	4	70	OTH	NA	WHI	NA	NA	0.703	103	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2100	USA	sbc0014	X	M	41	4	41	OTH	NA	WHI	UNI	CLM	0.703	2743	1	1	0	0	0	2	0.729	0.365	0.365	0.000	0.000
2105	USA	sbc015	J	F	29	2	29	CA	NA	WHI	UNI	ACD	9.292	3717	4	18	0	7	0	29	7.802	1.076	4.843	0.000	1.883
2106	USA	sbc015	K	M	34	3	34	CA	NA	WHI	UNI	CLM	9.292	3004	3	26	0	8	0	37	12.317	0.999	8.655	0.000	2.663
2107	USA	sbc015	L	F	30	2	30	CA	NA	WHI	UNI	EDU	9.292	978	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2108	USA	sbc016	B	M	25	1	25	CA	NA	OTH	UNI	CLM	1.073	3668	3	0	0	2	0	5	1.363	0.818	0.000	0.000	0.545
2109	USA	sbc016	T	F	51	4	51	CA	NA	WHI	UNI	ACD	1.073	1924	0	1	0	0	0	1	0.520	0.000	0.520	0.000	0.000

2112	USA	sbc017	J	M	45	4	45	CA	NA	WHI	UNI	SLM	5.023	2811	1	2	0	3	0	6	2.134	0.356	0.711	0.000	1.067
2113	USA	sbc017	M	M	34	3	34	CA	NA	WHI	UNI	CLM	5.023	1768	6	5	0	6	0	17	9.615	3.394	2.828	0.000	3.394
2114	USA	sbc018	H	F	34	3	34	OTH	NA	WHI	UNI	NA	1.344	118	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2116	USA	sbc018	K	F	40	4	40	OTH	NA	WHI	UNI	CLM	1.344	490	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2117	USA	sbc018	L	F	28	2	28	OTH	NA	WHI	UNI	ACD	1.344	643	0	1	0	1	0	2	3.110	0.000	1.555	0.000	1.555
2115	USA	sbc018	M	F	28	2	28	OTH	NA	WHI	UNI	ACD	1.344	845	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2123	USA	sbc019	B	M	12	1	12	OTH	NA	WHI	UNI	EDU	2.219	783	1	0	0	1	0	2	2.554	1.277	0.000	0.000	1.277
2120	USA	sbc019	F	M	42	4	42	OTH	NA	WHI	UNI	ACD	2.219	1951	1	2	0	1	0	4	2.050	0.513	1.025	0.000	0.513
2121	USA	sbc019	J	F	40	4	40	OTH	NA	WHI	UNI	CLM	2.219	414	0	1	0	0	0	1	2.415	0.000	2.415	0.000	0.000
2122	USA	sbc019	M	F	14	1	14	OTH	NA	WHI	SEC	EDU	2.219	1509	2	2	0	0	0	4	2.651	1.325	1.325	0.000	0.000
2119	USA	sbc019	R	M	31	2	31	CA	NA	WHI	UNI	EDU	2.219	294	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2125	USA	sbc022	L	M	28	2	28	OTH	NA	OTH	UNI	CLM	1.027	1062	0	1	0	0	0	1	0.942	0.000	0.942	0.000	0.000
2124	USA	sbc022	R	M	40	4	40	OTH	NA	WHI	UNI	CLM	1.027	1860	0	1	0	1	0	2	1.075	0.000	0.538	0.000	0.538
2134	USA	sbc023	D	F	46	4	46	OTH	NA	WHI	UNI	NA	2.120	608	0	1	0	0	0	1	1.645	0.000	1.645	0.000	0.000
2132	USA	sbc023	E	F	82	4	82	OTH	NA	WHI	UNI	NA	2.120	1098	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2133	USA	sbc023	I	F	47	4	47	OTH	NA	WHI	UNI	NA	2.120	658	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2131	USA	sbc023	J	F	59	4	59	OTH	NA	WHI	UNI	CLM	2.120	655	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2137	USA	sbc023	K	F	62	4	62	OTH	NA	WHI	UNI	CLM	2.120	466	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2127	USA	sbc023	L	F	56	4	56	OTH	NA	WHI	UNI	ACD	2.120	788	0	0	0	1	0	1	1.269	0.000	0.000	0.000	1.269
2128	USA	sbc023	N	F	65	4	65	OTH	NA	NA	UNI	CLM	2.120	111	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2135	USA	sbc023	O	F	58	4	58	OTH	NA	WHI	UNI	ACD	2.120	201	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2136	USA	sbc023	P	F	53	4	53	OTH	NA	WHI	UNI	NA	2.120	794	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2129	USA	sbc023	U	F	53	4	53	OTH	NA	NA	UNI	CLM	2.120	621	0	1	0	2	0	3	4.831	0.000	1.610	0.000	3.221
2142	USA	sbc024	D	M	23	1	23	OTH	NA	WHI	SEC	SLM	1.541	1093	0	3	0	0	0	3	2.745	0.000	2.745	0.000	0.000
2143	USA	sbc024	J	F	23	1	23	OTH	NA	WHI	UNI	EDU	1.541	2152	0	2	0	0	0	2	0.929	0.000	0.929	0.000	0.000

2145	USA	sbc028	F	M	78	4	78	CA	NA	WHI	UNI	ACD	14.939	2695	18	11	0	16	0	45	16.698	6.679	4.082	0.000	5.937
2144	USA	sbc028	J	F	23	1	23	OTH	NA	WHI	UNI	EDU	14.939	2727	20	8	0	8	0	36	13.201	7.334	2.934	0.000	2.934
2146	USA	sbc029	L	M	53	4	53	CA	NA	WHI	UNI	CLM	2.087	2720	0	0	0	0	1	1	0.368	0.000	0.000	0.000	0.000
2147	USA	sbc029	S	M	33	2	33	CA	NA	WHI	UNI	CLM	2.087	3031	2	7	0	2	0	11	3.629	0.660	2.309	0.000	0.660
2149	USA	sbc031	B	F	25	1	25	OTH	NA	WHI	UNI	CLM	1.104	2375	1	0	0	0	0	1	0.421	0.421	0.000	0.000	0.000
2151	USA	sbc031	J	F	NA	NA	NA	OTH	NA	NA	NA	SLM	1.104	227	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2148	USA	sbc031	S	F	28	2	28	OTH	NA	WHI	UNI	EDU	1.104	1584	2	0	0	0	0	2	1.263	1.263	0.000	0.000	0.000
2150	USA	sbc031	R	F	57	4	57	OTH	NA	WHI	UNI	ACD	0.000	1633	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2152	USA	sbc032	A	M	60	4	60	OTH	NA	WHI	UNI	CLM	0.000	1310	1	0	0	0	0	1	0.763	0.763	0.000	0.000	0.000
2153	USA	sbc032	B	M	70	4	70	OTH	NA	WHI	UNI	ACD	0.000	3784	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2154	USA	sbc032	C	M	68	4	68	OTH	NA	OTH	UNI	ACD	0.000	2015	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2158	USA	sbc033	B	F	NA	NA	NA	OTH	NA	WHI	UNI	ACD	3.444	554	0	0	0	1	0	1	1.805	0.000	0.000	0.000	1.805
2161	USA	sbc033	D	M	60	4	60	OTH	NA	WHI	UNI	CLM	3.444	172	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2159	USA	sbc033	I	M	52	4	52	OTH	NA	WHI	UNI	ACD	3.444	417	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2157	USA	sbc033	J	F	23	1	23	OTH	NA	WHI	UNI	EDU	3.444	729	1	0	0	1	0	2	2.743	1.372	0.000	0.000	1.372
2160	USA	sbc033	L	F	54	4	54	OTH	NA	WHI	UNI	ACD	3.444	343	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2162	USA	sbc033	N	F	24	1	24	OTH	NA	WHI	UNI	EDU	3.444	528	1	2	0	0	0	3	5.682	1.894	3.788	0.000	0.000
2166	USA	sbc034	K	F	22	1	22	OTH	NA	WHI	UNI	SLM	3.816	2394	1	1	0	1	0	3	1.253	0.418	0.418	0.000	0.418
2167	USA	sbc034	S	M	24	1	24	OTH	NA	WHI	UNI	CLM	3.816	751	2	5	0	2	0	9	11.984	2.663	6.658	0.000	2.663
2170	USA	sbc035	E	F	17	1	17	OTH	NA	WHI	SEC	EDU	2.537	187	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2172	USA	sbc035	L	F	NA	NA	NA	NA	NA	NA	NA	NA	2.537	778	0	1	0	1	0	2	2.571	0.000	1.285	0.000	1.285
2171	USA	sbc035	M	F	NA	NA	NA	NA	NA	NA	NA	NA	2.537	443	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2168	USA	sbc035	P	F	53	4	53	OTH	NA	WHI	UNI	ACD	2.537	2082	1	1	0	0	0	2	0.961	0.480	0.480	0.000	0.000
2174	USA	sbc035	R	M	81	4	81	OTH	NA	OTH	UNI	ACD	2.537	101	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2169	USA	sbc035	S	F	16	1	16	OTH	NA	WHI	SEC	EDU	2.537	2264	1	5	0	0	0	6	2.650	0.442	2.208	0.000	0.000

2176	USA	sbc036	K	M	22	1	22	CA	NA	WHI	UNI	EDU	21.084	1161	1	6	0	5	0	12	10.336	0.861	5.168	0.000	4.307
2177	USA	sbc036	L	F	20	1	20	OTH	NA	WHI	UNI	EDU	21.084	2593	9	22	0	22	1	54	20.825	3.471	8.484	0.000	8.484
2178	USA	sbc036	M	F	20	1	20	OTH	NA	OTH	UNI	SLM	21.084	3478	16	34	0	23	1	74	21.277	4.600	9.776	0.000	6.613
2179	USA	sbc037	D	F	56	4	56	OTH	NA	OTH	UNI	CLM	2.381	1617	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2181	USA	sbc037	J	F	80	4	80	OTH	NA	OTH	SEC	CLM	2.381	1297	0	0	0	1	0	1	0.771	0.000	0.000	0.000	0.771
2182	USA	sbc037	K	F	20	1	20	OTH	NA	OTH	UNI	EDU	2.381	144	0	1	0	0	0	1	6.944	0.000	6.944	0.000	0.000
2180	USA	sbc037	S	M	26	2	26	OTH	NA	OTH	UNI	EDU	2.381	696	0	1	0	0	0	1	1.437	0.000	1.437	0.000	0.000
2184	USA	sbc039	A	F	NA	NA	NA	NA	NA	NA	NA	NA	1.460	337	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2185	USA	sbc039	D	M	NA	NA	NA	NA	NA	NA	NA	NA	1.460	284	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2183	USA	sbc039	K	F	29	2	29	OTH	NA	WHI	UNI	ACD	1.460	3731	1	1	0	0	0	2	0.536	0.268	0.268	0.000	0.000
2186	USA	sbc039	L	F	NA	NA	NA	NA	NA	NA	NA	NA	1.460	504	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2187	USA	sbc039	N	F	NA	NA	NA	NA	NA	NA	NA	NA	1.460	181	0	0	0	0	1	1	5.525	0.000	0.000	0.000	0.000
2190	USA	sbc041	K	F	35	3	35	CA	NA	WHI	UNI	CLM	6.674	2641	3	2	0	1	0	6	2.272	1.136	0.757	0.000	0.379
2191	USA	sbc041	P	F	37	3	37	CA	NA	WHI	UNI	EDU	6.674	955	7	8	0	3	0	18	18.848	7.330	8.377	0.000	3.141
2196	USA	sbc042	C	M	56	4	56	OTH	NA	WHI	UNI	CLM	0.000	288	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2197	USA	sbc042	G	F	17	1	17	OTH	NA	OTH	SEC	EDU	0.000	192	0	1	0	0	0	1	5.208	0.000	5.208	0.000	0.000
2192	USA	sbc042	K	F	15	1	15	OTH	NA	WHI	SEC	EDU	0.000	620	0	1	0	0	0	1	1.613	0.000	1.613	0.000	0.000
2193	USA	sbc042	S	F	18	1	18	OTH	NA	WHI	SEC	EDU	0.000	731	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2195	USA	sbc042	T	F	41	4	41	OTH	NA	WHI	UNI	CLM	0.000	1393	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2198	USA	sbc043	A	F	49	4	49	OTH	NA	WHI	UNI	ACD	1.421	3316	0	1	0	1	0	2	0.603	0.000	0.302	0.000	0.302
2199	USA	sbc043	T	F	24	1	24	OTH	NA	WHI	UNI	CLM	1.421	3722	0	3	0	5	0	8	2.149	0.000	0.806	0.000	1.343
2200	USA	sbc044	C	M	30	2	30	OTH	NA	WHI	UNI	CLM	5.566	1300	2	8	0	1	1	12	9.231	1.538	6.154	0.000	0.769
2201	USA	sbc044	L	M	35	3	35	OTH	NA	OTH	UNI	CLM	5.566	5707	11	14	0	2	0	27	4.731	1.927	2.453	0.000	0.350
2203	USA	sbc045	C	F	26	2	26	OTH	NA	OTH	UNI	EDU	17.544	2451	8	31	0	17	1	57	23.256	3.264	12.648	0.000	6.936
2202	USA	sbc045	P	M	25	1	25	OTH	NA	WHI	UNI	CLM	17.544	2451	2	23	0	4	0	29	11.832	0.816	9.384	0.000	1.632

2205	USA	sbc046	D	M	25	1	25	CA	NA	WHI	UNI	EDU	0.589	1020	1	0	0	0	0	1	0.980	0.980	0.000	0.000	0.000
2204	USA	sbc046	R	M	52	4	52	CA	NA	WHI	UNI	CLM	0.589	2376	1	0	0	0	0	1	0.421	0.421	0.000	0.000	0.000
2208	USA	sbc047	F	M	NA	NA	NA	CA	NA	OTH	NA	SLM	1.816	1550	3	1	0	4	0	8	5.161	1.935	0.645	0.000	2.581
2207	USA	sbc047	R	M	32	2	32	CA	NA	OTH	SEC	SLM	1.816	3405	1	0	0	0	0	1	0.294	0.294	0.000	0.000	0.000
2212	USA	sbc048	D	M	32	2	32	CA	NA	WHI	SEC	CLM	0.446	312	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2211	USA	sbc048	J	F	28	2	28	CA	NA	WHI	UNI	EDU	0.446	1052	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2209	USA	sbc048	L	F	55	4	55	CA	NA	WHI	UNI	CLM	0.446	1409	0	1	0	0	0	1	0.710	0.000	0.710	0.000	0.000
2210	USA	sbc048	T	M	60	4	60	CA	NA	WHI	UNI	CLM	0.446	835	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2217	USA	sbc049	A	M	44	4	44	OTH	NA	WHI	UNI	CLM	0.000	400	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2215	USA	sbc049	D	M	36	3	36	OTH	NA	WHI	UNI	EDU	0.000	1132	1	2	0	0	0	3	2.650	0.883	1.767	0.000	0.000
2218	USA	sbc049	I	F	42	4	42	OTH	NA	WHI	UNI	ACD	0.000	167	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2213	USA	sbc049	J	M	53	4	53	OTH	NA	WHI	SEC	CLM	0.000	832	0	1	0	0	0	1	1.202	0.000	1.202	0.000	0.000
2216	USA	sbc049	L	F	47	4	47	OTH	NA	WHI	NA	CLM	0.000	2504	2	1	0	0	0	3	1.198	0.799	0.399	0.000	0.000
2214	USA	sbc049	N	F	44	4	44	OTH	NA	WHI	UNI	CLM	0.000	245	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2219	USA	sbc049	V	M	15	1	15	OTH	NA	WHI	NA	EDU	0.000	145	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2226	USA	sbc050	A	F	21	1	21	OTH	NA	WHI	UNI	EDU	11.396	1136	11	8	0	8	1	28	24.648	9.683	7.042	0.000	7.042
2223	USA	sbc050	D	F	20	1	20	OTH	NA	WHI	UNI	EDU	11.396	1154	3	2	0	8	0	13	11.265	2.600	1.733	0.000	6.932
2224	USA	sbc050	K	F	20	1	20	OTH	NA	WHI	UNI	EDU	11.396	952	8	2	0	1	0	11	11.555	8.403	2.101	0.000	1.050
2228	USA	sbc051	A	F	47	4	47	CA	NA	WHI	UNI	ACD	0.921	1264	1	0	0	0	0	1	0.791	0.791	0.000	0.000	0.000
2230	USA	sbc051	B	M	51	4	51	CA	NA	WHI	UNI	SLM	0.921	890	0	3	0	0	0	3	3.371	0.000	3.371	0.000	0.000
2227	USA	sbc051	F	F	50	4	50	OTH	NA	WHI	UNI	CLM	0.921	3186	1	0	0	1	0	2	0.628	0.314	0.000	0.000	0.314
2229	USA	sbc051	S	M	49	4	49	CA	NA	WHI	UNI	SLM	0.921	1155	0	1	0	1	0	2	1.732	0.000	0.866	0.000	0.866
2232	USA	sbc052	A	M	47	4	47	OTH	NA	WHI	UNI	CLM	2.337	1656	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2233	USA	sbc052	C	F	47	4	47	OTH	NA	WHI	UNI	CLM	2.337	1325	0	1	0	0	0	1	0.755	0.000	0.755	0.000	0.000
2235	USA	sbc052	D	F	44	4	44	OTH	NA	WHI	SEC	SLM	2.337	3382	4	3	0	3	0	10	2.957	1.183	0.887	0.000	0.887

2240	USA	sbc056	A	F	NA	NA	NA	OTH	NA	NA	NA	NA	0.663	217	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2236	USA	sbc056	G	M	41	4	41	OTH	NA	WHI	UNI	ACD	0.663	1419	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2238	USA	sbc056	J	F	NA	NA	NA	OTH	NA	NA	NA	NA	0.663	5913	1	2	0	1	0	4	0.676	0.169	0.338	0.000	0.169
2239	USA	sbc056	K	F	NA	NA	NA	OTH	NA	NA	NA	NA	0.663	121	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2247	USA	sbc057	B	M	23	1	23	CA	NA	WHI	UNI	CLM	0.000	139	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2244	USA	sbc057	D	M	24	1	24	CA	NA	WHI	UNI	EDU	0.000	192	0	0	0	1	0	1	5.208	0.000	0.000	0.000	5.208
2245	USA	sbc057	J	M	37	3	37	CA	NA	NA	UNI	ACD	0.000	666	0	1	0	0	0	1	1.502	0.000	1.502	0.000	0.000
2243	USA	sbc057	L	M	22	1	22	CA	NA	WHI	UNI	SLM	0.000	126	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2242	USA	sbc057	N	M	37	3	37	CA	NA	OTH	UNI	SLM	0.000	3125	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2249	USA	sbc058	I	F	37	3	37	OTH	NA	WHI	UNI	EDU	1.160	2674	1	1	0	0	0	2	0.748	0.374	0.374	0.000	0.000
2250	USA	sbc058	S	M	11	1	11	OTH	NA	WHI	SEC	EDU	1.160	1636	2	0	0	1	0	3	1.834	1.222	0.000	0.000	0.611
2252	USA	sbc059	C	M	30	2	30	OTH	NA	WHI	UNI	CLM	1.320	1194	2	5	0	4	1	12	10.050	1.675	4.188	0.000	3.350
2251	USA	sbc059	F	M	32	2	32	OTH	NA	WHI	UNI	CLM	1.320	964	0	0	0	1	0	1	1.037	0.000	0.000	0.000	1.037
2253	USA	sbc059	J	F	65	4	65	OTH	NA	WHI	SEC	SLM	1.320	3096	3	3	0	1	0	7	2.261	0.969	0.969	0.000	0.323
2254	USA	sbc059	W	M	71	4	71	OTH	NA	WHI	SEC	SLM	1.320	2206	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2256	USA	sbc060	A	M	66	4	66	CA	NA	WHI	UNI	CLM	0.672	4035	1	1	0	0	0	2	0.496	0.248	0.248	0.000	0.000
2255	USA	sbc060	J	M	72	4	72	CA	NA	WHI	UNI	CLM	0.672	428	0	1	0	0	0	1	2.336	0.000	2.336	0.000	0.000

Table 139: Final data set of CanE.

ID	VAR	FILE	SPEAKER	SEX	age.orig	AGE	DATE	EDU	OCC	L1	PAI	words	ini.n	med.n	fin.m	non.n	na.n	all.n	all	ini	med	fin	non
1	CAN	S1A-001	A	F	41-50	4	2	UNI	ACD	ENG	0.000	1305	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2	CAN	S1A-001	B	F	41-50	4	2	NA	CLM	ENG	0.000	671	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
5	CAN	S1A-002	A	F	41-50	4	2	UNI	ACD	ENG	0.548	1179	0	1	0	0	0	1	0.848	0.000	0.848	0.000	0.000
6	CAN	S1A-002	B	F	41-50	4	2	NA	CLM	ENG	0.548	647	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
9	CAN	S1A-003	A	F	41-50	4	2	UNI	ACD	ENG	1.168	1088	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
10	CAN	S1A-003	B	F	41-50	4	2	NA	CLM	ENG	0.510	581	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
12	CAN	S1A-003	B	F	25-30	2	2	SEC	SLM	ENG	1.168	275	1	0	0	0	0	1	3.636	3.636	0.000	0.000	0.000
13	CAN	S1A-004	A	M	31-40	3	2	UNI	CLM	OTH	0.510	688	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
14	CAN	S1A-004	B	F	25-30	2	2	SEC	SLM	ENG	0.510	1273	0	1	0	0	0	1	0.786	0.000	0.786	0.000	0.000
15	CAN	S1A-005	A	M	31-40	3	2	UNI	ACD	ENG+	1.447	1350	0	2	0	0	0	2	1.481	0.000	1.481	0.000	0.000
16	CAN	S1A-005	B	F	41-50	4	2	UNI	ACD	ENG+	1.447	723	1	0	0	0	0	1	1.383	1.383	0.000	0.000	0.000
17	CAN	S1A-006	A	M	31-40	3	2	UNI	ACD	ENG+	0.996	1005	1	1	0	0	0	2	1.990	0.995	0.995	0.000	0.000
19	CAN	S1A-006	A	F	25-30	2	2	UNI	ACD	ENG	0.996	223	1	0	0	0	0	1	4.484	4.484	0.000	0.000	0.000
18	CAN	S1A-006	B	F	41-50	4	2	UNI	ACD	ENG+	0.996	781	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
21	CAN	S1A-007	A	F	25-30	2	2	UNI	ACD	ENG	9.500	667	5	1	0	0	0	6	8.996	7.496	1.499	0.000	0.000
22	CAN	S1A-007	B	M	31-40	3	2	UNI	ACD	ENG	9.500	1333	4	5	1	3	0	13	9.752	3.001	3.751	0.750	2.251
23	CAN	S1A-008	A	F	25-30	2	2	UNI	ACD	ENG	3.916	817	3	1	0	0	0	4	4.896	3.672	1.224	0.000	0.000
27	CAN	S1A-008	A	F	25-30	2	2	UNI	ACD	ENG	3.916	303	0	0	0	1	1	2	3.300	0.000	0.000	0.000	3.300
24	CAN	S1A-008	B	M	31-40	3	2	UNI	ACD	ENG	1.653	482	1	1	0	0	0	2	4.149	2.075	2.075	0.000	0.000
28	CAN	S1A-008	B	M	31-40	3	2	UNI	ACD	ENG	3.916	284	0	1	0	0	0	1	3.521	0.000	3.521	0.000	0.000
29	CAN	S1A-009	A	M	51+	4	2	UNI	ACD	ENG	1.653	671	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
31	CAN	S1A-009	A	M	20-24	1	2	SEC	NA	ENG	1.653	274	0	1	0	1	0	2	7.299	0.000	3.650	0.000	3.650
30	CAN	S1A-009	B	F	31-40	3	2	UNI	ACD	ENG	1.653	936	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

33	CAN	S1A-010	A	M	20-24	1	2	SEC	NA	ENG	7.380	1355	4	3	1	2	0	10	7.380	2.952	2.214	0.738	1.476
35	CAN	S1A-011	A	F	31-40	3	3	UNI	ACD	OTH	2.481	895	4	1	0	2	0	7	7.821	4.469	1.117	0.000	2.235
37	CAN	S1A-011	A	M	41-50	4	3	UNI	ACD	ENG	2.481	171	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
41	CAN	S1A-011	A	F	31-40	3	3	UNI	ACD	OTH	2.481	285	1	0	0	0	0	1	3.509	3.509	0.000	0.000	0.000
36	CAN	S1A-011	B	F	51+	4	3	UNI	ACD	ENG	1.460	258	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
38	CAN	S1A-011	B	F	41-50	4	3	UNI	ACD	ENG	1.460	258	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
42	CAN	S1A-011	B	F	51+	4	3	UNI	ACD	ENG	2.481	145	1	0	0	0	0	1	6.897	6.897	0.000	0.000	0.000
43	CAN	S1A-012	A	M	41-50	4	3	UNI	ACD	ENG	1.460	958	1	1	0	0	0	2	2.088	1.044	1.044	0.000	0.000
44	CAN	S1A-012	B	F	41-50	4	3	UNI	ACD	ENG	1.460	1097	1	0	0	0	0	1	0.912	0.912	0.000	0.000	0.000
45	CAN	S1A-013	A	M	41-50	4	3	UNI	ACD	ENG	2.198	246	0	1	0	0	0	1	4.065	0.000	4.065	0.000	0.000
47	CAN	S1A-013	A	M	41-50	4	3	UNI	ACD	ENG	2.198	916	3	5	1	1	0	10	10.917	3.275	5.459	1.092	1.092
46	CAN	S1A-013	B	F	41-50	4	3	UNI	ACD	ENG	0.962	288	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
48	CAN	S1A-013	B	F	41-50	4	3	UNI	ACD	ENG	2.198	622	1	1	0	0	0	2	3.215	1.608	1.608	0.000	0.000
49	CAN	S1A-014	A	F	41-50	4	2	UNI	ACD	ENG	0.962	928	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
50	CAN	S1A-014	B	F	31-40	3	2	UNI	ACD	ENG	0.962	1152	1	1	0	0	0	2	1.736	0.868	0.868	0.000	0.000
52	CAN	S1A-015	A	M	51+	4	2	UNI	ACD	ENG	0.000	995	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
53	CAN	S1A-015	B	F	31-40	3	2	UNI	ACD	ENG	0.000	969	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
54	CAN	S1A-016	A	M	51+	4	2	UNI	ACD	ENG	0.000	1002	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
55	CAN	S1A-016	B	F	31-40	3	2	UNI	ACD	ENG	0.000	1032	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
56	CAN	S1A-017	A	M	31-40	3	2	UNI	CLM	OTH	3.755	172	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
58	CAN	S1A-017	A	M	31-40	3	2	UNI	CLM	OTH	3.755	644	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
60	CAN	S1A-017	A	M	51+	4	2	UNI	ACD	ENG	3.755	246	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
57	CAN	S1A-017	B	F	25-30	2	2	SEC	SLM	ENG	0.000	157	1	2	0	0	0	3	19.108	6.369	12.739	0.000	0.000
59	CAN	S1A-017	B	F	25-30	2	2	SEC	SLM	ENG	0.000	683	1	2	0	0	0	3	4.392	1.464	2.928	0.000	0.000
61	CAN	S1A-017	B	F	31-40	3	2	UNI	ACD	ENG	3.755	116	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

64	CAN	S1A-018	A	M	31-40	3	2	UNI	CLM	OTH	0.000	866	0	1	0	0	0	1	1.155	0.000	1.155	0.000	0.000
63	CAN	S1A-018	B	F	25-30	2	2	SEC	SLM	ENG	0.539	631	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
65	CAN	S1A-018	B	F	25-30	2	2	SEC	SLM	ENG	0.000	397	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
66	CAN	S1A-019	A	F	25-30	2	2	UNI	ACD	ENG	0.539	172	0	1	0	0	0	1	5.814	0.000	5.814	0.000	0.000
68	CAN	S1A-019	A	M	41-50	4	2	UNI	CLM	ENG	0.539	1056	0	1	0	0	0	1	0.947	0.000	0.947	0.000	0.000
69	CAN	S1A-019	B	F	51+	4	2	UNI	CLM	ENG	0.539	800	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
70	CAN	S1A-020	A	M	20-24	1	2	SEC	NA	ENG	11.558	1367	4	6	0	0	0	10	7.315	2.926	4.389	0.000	0.000
71	CAN	S1A-020	B	M	20-24	1	2	UNI	EDU	ENG	11.558	623	4	7	1	1	0	13	20.867	6.421	11.236	1.605	1.605
72	CAN	S1A-021	A	M	20-24	1	2	SEC	NA	ENG	11.646	1251	2	7	1	3	0	13	10.392	1.599	5.596	0.799	2.398
73	CAN	S1A-021	B	M	20-24	1	2	UNI	EDU	ENG	11.646	724	5	4	0	1	0	10	13.812	6.906	5.525	0.000	1.381
74	CAN	S1A-022	A	M	20-24	1	2	SEC	NA	ENG	8.077	1389	2	5	0	2	0	9	6.479	1.440	3.600	0.000	1.440
75	CAN	S1A-022	B	M	20-24	1	2	UNI	EDU	ENG	8.077	592	3	3	0	1	0	7	11.824	5.068	5.068	0.000	1.689
76	CAN	S1A-023	A	M	20-24	1	2	SEC	NA	ENG	0.931	411	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
78	CAN	S1A-023	A	M	31-40	3	2	UNI	EDU	ENG	0.931	242	0	0	0	1	0	1	4.132	0.000	0.000	0.000	4.132
77	CAN	S1A-023	B	M	20-24	1	2	UNI	EDU	ENG	0.931	259	1	2	0	0	0	3	11.583	3.861	7.722	0.000	0.000
79	CAN	S1A-023	B	M	31-40	3	2	UNI	CLM	OTH	0.931	398	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
80	CAN	S1A-023	C	F	25-30	2	2	UNI	EDU	ENG	0.931	676	0	1	0	0	0	1	1.479	0.000	1.479	0.000	0.000
81	CAN	S1A-024	A	M	25-30	2	2	UNI	CLM	ENG	0.977	1136	1	0	0	0	0	1	0.880	0.880	0.000	0.000	0.000
82	CAN	S1A-024	B	M	25-30	2	2	UNI	ACD	ENG	0.977	912	0	1	0	0	0	1	1.096	0.000	1.096	0.000	0.000
83	CAN	S1A-025	A	M	25-30	2	2	UNI	CLM	ENG	3.865	1483	2	3	0	0	0	5	3.372	1.349	2.023	0.000	0.000
84	CAN	S1A-025	B	M	25-30	2	2	UNI	ACD	ENG	3.865	587	1	1	0	1	0	3	5.111	1.704	1.704	0.000	1.704
85	CAN	S1A-026	A	M	31-40	3	2	UNI	EDU	ENG	2.331	295	0	1	0	0	0	1	3.390	0.000	3.390	0.000	0.000
86	CAN	S1A-026	B	M	31-40	3	2	UNI	CLM	OTH	2.331	1032	2	0	0	0	0	2	1.938	1.938	0.000	0.000	0.000
87	CAN	S1A-026	C	F	25-30	2	2	UNI	EDU	ENG	2.331	684	1	1	0	0	0	2	2.924	1.462	1.462	0.000	0.000
88	CAN	S1A-027	A	M	31-40	3	2	UNI	EDU	ENG	5.249	471	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

89	CAN	S1A-027	B	M	31-40	3	2	UNI	CLM	OTH	5.249	873	2	0	0	0	0	2	2.291	2.291	0.000	0.000	0.000
90	CAN	S1A-027	C	F	25-30	2	2	UNI	EDU	ENG	5.249	651	3	0	0	3	0	6	9.217	4.608	0.000	0.000	4.608
91	CAN	S1A-028	A	M	31-40	3	2	UNI	EDU	ENG	12.949	340	1	0	0	0	0	1	2.941	2.941	0.000	0.000	0.000
92	CAN	S1A-028	B	M	31-40	3	2	UNI	CLM	OTH	12.949	648	3	1	0	2	0	6	9.259	4.630	1.543	0.000	3.086
93	CAN	S1A-028	C	F	25-30	2	2	UNI	EDU	ENG	12.949	1051	10	3	1	2	0	16	15.224	9.515	2.854	0.951	1.903
94	CAN	S1A-029	A	M	41-50	4	2	UNI	CLM	ENG	0.484	737	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
95	CAN	S1A-029	B	F	51+	4	2	UNI	CLM	ENG	0.484	1327	0	1	0	0	0	1	0.754	0.000	0.754	0.000	0.000
99	CAN	S1A-030	A	F	31-40	3	2	UNI	SLM	ENG	11.186	976	3	6	0	2	0	11	11.270	3.074	6.148	0.000	2.049
100	CAN	S1A-030	B	F	41-50	4	2	UNI	ACD	ENG	11.186	590	1	0	1	1	0	3	5.085	1.695	0.000	1.695	1.695
98	CAN	S1A-030	C	F	25-30	2	2	UNI	EDU	ENG	11.186	304	2	4	0	1	0	7	23.026	6.579	13.158	0.000	3.289
101	CAN	S1A-031	A	F	31-40	3	2	UNI	SLM	ENG	6.927	956	5	4	0	1	0	10	10.460	5.230	4.184	0.000	1.046
102	CAN	S1A-031	B	F	41-50	4	2	UNI	ACD	ENG	6.927	1065	1	3	0	0	0	4	3.756	0.939	2.817	0.000	0.000
103	CAN	S1A-032	A	F	31-40	3	2	UNI	SLM	ENG	8.704	1296	10	5	0	1	0	16	13.117	7.716	3.858	0.000	0.772
104	CAN	S1A-032	B	F	41-50	4	2	UNI	ACD	ENG	8.704	772	2	0	0	0	0	2	2.591	2.591	0.000	0.000	0.000
105	CAN	S1A-033	A	F	31-40	3	2	UNI	SLM	ENG	4.400	728	4	1	1	1	0	7	9.615	5.495	1.374	1.374	1.374
106	CAN	S1A-033	B	F	41-50	4	2	UNI	ACD	ENG	4.400	863	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
109	CAN	S1A-034	A	F	31-40	3	2	UNI	ACD	ENG	2.572	946	2	1	0	1	0	4	4.228	2.114	1.057	0.000	1.057
110	CAN	S1A-034	B	M	41-50	4	2	UNI	ACD	ENG	2.572	998	0	1	0	0	0	1	1.002	0.000	1.002	0.000	0.000
111	CAN	S1A-035	A	F	31-40	3	2	UNI	ACD	ENG	2.956	1023	1	2	0	1	0	4	3.910	0.978	1.955	0.000	0.978
112	CAN	S1A-035	B	M	41-50	4	2	UNI	ACD	ENG	2.956	1007	1	0	0	1	0	2	1.986	0.993	0.000	0.000	0.993
113	CAN	S1A-036	A	F	31-40	3	2	UNI	ACD	ENG	3.990	1255	3	4	0	1	0	8	6.375	2.390	3.187	0.000	0.797
114	CAN	S1A-036	B	M	41-50	4	2	UNI	ACD	ENG	3.990	750	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
115	CAN	S1A-037	A	F	31-40	3	2	UNI	ACD	ENG	0.925	1116	0	2	0	0	0	2	1.792	0.000	1.792	0.000	0.000
116	CAN	S1A-037	B	M	41-50	4	2	UNI	ACD	ENG	0.925	1047	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
117	CAN	S1A-038	A	M	41-50	4	2	UNI	CLM	ENG	0.000	1142	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

118	CAN	S1A-038	B	F	51+	4	2	UNI	CLM	ENG	0.000	940	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
119	CAN	S1A-039	A	F	31-40	3	2	UNI	ACD	ENG	0.000	283	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
121	CAN	S1A-039	A	F	31-40	3	2	UNI	ACD	ENG	0.000	365	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
120	CAN	S1A-039	B	M	41-50	4	2	UNI	ACD	ENG	3.982	582	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
122	CAN	S1A-039	B	F	41-50	4	2	UNI	CLM	ENG	0.000	769	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
123	CAN	S1A-040	A	F	31-40	3	2	UNI	ACD	ENG	3.982	354	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
124	CAN	S1A-040	B	F	41-50	4	2	UNI	CLM	ENG	3.982	1655	3	3	0	2	0	8	4.834	1.813	1.813	0.000	1.208
125	CAN	S1A-041	A	F	31-40	3	2	UNI	ACD	ENG	3.009	842	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
126	CAN	S1A-041	B	F	41-50	4	2	UNI	CLM	ENG	3.009	1152	3	2	1	0	0	6	5.208	2.604	1.736	0.868	0.000
127	CAN	S1A-042	A	F	31-40	3	2	UNI	ACD	ENG	2.510	879	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
128	CAN	S1A-042	B	F	41-50	4	2	UNI	CLM	ENG	2.510	1113	1	2	0	2	0	5	4.492	0.898	1.797	0.000	1.797
129	CAN	S1A-043	A	F	31-40	3	2	UNI	ACD	ENG	5.769	697	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
131	CAN	S1A-043	A	F	25-30	2	2	UNI	CLM	ENG	5.769	292	1	4	0	0	0	5	17.123	3.425	13.699	0.000	0.000
130	CAN	S1A-043	B	F	41-50	4	2	UNI	CLM	ENG	5.769	502	2	0	0	1	0	3	5.976	3.984	0.000	0.000	1.992
132	CAN	S1A-043	B	F	25-30	2	2	UNI	ACD	ENG	5.769	284	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
133	CAN	S1A-043	C	F	25-30	2	2	UNI	CLM	ENG	5.769	236	0	3	0	0	0	3	12.712	0.000	12.712	0.000	0.000
134	CAN	S1A-044	A	F	25-30	2	2	UNI	CLM	ENG	16.750	837	8	3	0	1	0	12	14.337	9.558	3.584	0.000	1.195
135	CAN	S1A-044	B	F	25-30	2	2	UNI	ACD	ENG	16.750	502	3	3	1	0	0	7	13.944	5.976	5.976	1.992	0.000
136	CAN	S1A-044	C	F	25-30	2	2	UNI	CLM	ENG	16.750	692	6	5	0	2	0	13	18.786	8.671	7.225	0.000	2.890
137	CAN	S1A-045	A	F	25-30	2	2	UNI	CLM	ENG	10.955	711	3	4	0	0	0	7	9.845	4.219	5.626	0.000	0.000
138	CAN	S1A-045	B	F	25-30	2	2	UNI	ACD	ENG	10.955	940	4	6	0	1	1	12	11.702	4.255	6.383	0.000	1.064
139	CAN	S1A-045	C	F	25-30	2	2	UNI	CLM	ENG	10.955	338	1	1	0	0	0	2	5.917	2.959	2.959	0.000	0.000
140	CAN	S1A-046	A	F	31-40	3	2	UNI	ACD	OTH	5.015	1356	2	4	0	2	0	8	5.900	1.475	2.950	0.000	1.475
141	CAN	S1A-046	B	M	31-40	3	2	UNI	ACD	ENG	5.015	638	0	0	0	2	0	2	3.135	0.000	0.000	0.000	3.135
142	CAN	S1A-047	A	F	31-40	3	2	UNI	ACD	OTH	0.498	1125	0	1	0	0	0	1	0.889	0.000	0.889	0.000	0.000

143	CAN	S1A-047	B	M	31-40	3	2	UNI	ACD	ENG	0.498	883	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
144	CAN	S1A-048	A	F	31-40	3	2	UNI	ACD	OTH	3.006	1294	1	4	0	1	0	6	4.637	0.773	3.091	0.000	0.773
145	CAN	S1A-048	B	M	31-40	3	2	UNI	ACD	ENG	3.006	702	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
146	CAN	S1A-049	A	F	31-40	3	2	UNI	ACD	OTH	2.988	1374	3	3	0	0	0	6	4.367	2.183	2.183	0.000	0.000
147	CAN	S1A-049	B	M	31-40	3	2	UNI	ACD	ENG	2.988	634	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
148	CAN	S1A-050	A	F	31-40	3	2	UNI	ACD	OTH	2.004	481	0	2	0	1	0	3	6.237	0.000	4.158	0.000	2.079
150	CAN	S1A-050	A	M	25-30	2	2	SEC	EDU	ENG	2.004	523	1	4	0	2	0	7	13.384	1.912	7.648	0.000	3.824
149	CAN	S1A-050	B	M	31-40	3	2	UNI	ACD	ENG	5.107	345	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
151	CAN	S1A-050	B	F	25-30	2	2	SEC	EDU	ENG	2.004	653	0	2	0	0	0	2	3.063	0.000	3.063	0.000	0.000
152	CAN	S1A-051	A	M	25-30	2	2	SEC	EDU	ENG	5.107	993	2	4	0	0	0	6	6.042	2.014	4.028	0.000	0.000
153	CAN	S1A-051	B	F	25-30	2	2	SEC	EDU	ENG	5.107	965	1	3	0	0	0	4	4.145	1.036	3.109	0.000	0.000
154	CAN	S1A-052	A	M	25-30	2	2	SEC	EDU	ENG	3.945	1052	4	0	0	1	0	5	4.753	3.802	0.000	0.000	0.951
155	CAN	S1A-052	B	F	25-30	2	2	SEC	EDU	ENG	3.945	976	2	1	0	0	0	3	3.074	2.049	1.025	0.000	0.000
156	CAN	S1A-053	A	M	25-30	2	2	SEC	EDU	ENG	2.532	1351	1	0	0	4	0	5	3.701	0.740	0.000	0.000	2.961
157	CAN	S1A-053	B	F	25-30	2	2	SEC	EDU	ENG	2.532	624	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
158	CAN	S1A-054	A	M	25-30	2	2	UNI	EDU	ENG	1.001	660	0	1	0	1	0	2	3.030	0.000	1.515	0.000	1.515
159	CAN	S1A-054	B	F	51+	4	2	SEC	SLM	ENG	1.001	1338	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
160	CAN	S1A-055	A	M	25-30	2	2	UNI	EDU	ENG	2.004	914	2	1	0	0	0	3	3.282	2.188	1.094	0.000	0.000
161	CAN	S1A-055	B	F	51+	4	2	SEC	SLM	ENG	2.004	1082	0	1	0	0	0	1	0.924	0.000	0.924	0.000	0.000
162	CAN	S1A-056	A	M	25-30	2	2	UNI	EDU	ENG	2.004	881	1	1	0	0	0	2	2.270	1.135	1.135	0.000	0.000
163	CAN	S1A-056	B	F	51+	4	2	SEC	SLM	ENG	2.004	1115	1	1	0	0	0	2	1.794	0.897	0.897	0.000	0.000
164	CAN	S1A-057	A	M	25-30	2	2	UNI	EDU	ENG	3.488	826	4	1	0	0	0	5	6.053	4.843	1.211	0.000	0.000
165	CAN	S1A-057	B	F	51+	4	2	SEC	SLM	ENG	3.488	1181	1	0	0	1	0	2	1.693	0.847	0.000	0.000	0.847
168	CAN	S1A-058	A	M	51+	4	2	UNI	ACD	ENG	0.000	261	0	1	0	0	0	1	3.831	0.000	3.831	0.000	0.000
167	CAN	S1A-058	B	F	51+	4	2	SEC	SLM	ENG	0.000	657	1	2	0	1	0	4	6.088	1.522	3.044	0.000	1.522

169	CAN	S1A-058	B	M	41-50	4	2	UNI	ACD	OTH	0.000	502	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
170	CAN	S1A-058	C	F	41-50	4	2	UNI	EDU	ENG	0.000	501	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
171	CAN	S1A-059	A	M	41-50	4	2	UNI	CLM	ENG	1.912	691	0	0	0	1	0	1	1.447	0.000	0.000	0.000	1.447
173	CAN	S1A-059	A	F	41-50	4	2	SEC	SLM	OTH	1.912	212	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
172	CAN	S1A-059	B	F	51+	4	2	UNI	CLM	ENG	3.956	337	1	0	0	1	0	2	5.935	2.967	0.000	0.000	2.967
174	CAN	S1A-059	B	F	31-40	3	2	UNI	CLM	ENG+	1.912	709	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
178	CAN	S1A-060	A	F	31-40	3	2	UNI	ACD	ENG	3.956	1153	1	7	0	0	0	8	6.938	0.867	6.071	0.000	0.000
179	CAN	S1A-060	B	F	51+	4	2	UNI	ACD	ENG	3.956	869	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
180	CAN	S1A-061	A	F	31-40	3	2	UNI	ACD	ENG	1.988	1032	0	0	1	0	0	1	0.969	0.000	0.000	0.969	0.000
181	CAN	S1A-061	B	F	51+	4	2	UNI	ACD	ENG	1.988	980	2	1	0	0	0	3	3.061	2.041	1.020	0.000	0.000
182	CAN	S1A-062	A	F	31-40	3	2	UNI	ACD	ENG	3.460	787	2	2	0	1	0	5	6.353	2.541	2.541	0.000	1.271
183	CAN	S1A-062	B	F	51+	4	2	UNI	ACD	ENG	3.460	1236	1	1	0	0	0	2	1.618	0.809	0.809	0.000	0.000
184	CAN	S1A-063	A	F	31-40	3	2	UNI	ACD	ENG	2.956	1086	2	2	0	0	1	5	3.683	1.842	1.842	0.000	0.000
185	CAN	S1A-063	B	F	51+	4	2	UNI	ACD	ENG	2.956	944	1	0	0	0	0	1	1.059	1.059	0.000	0.000	0.000
186	CAN	S1A-064	A	F	31-40	3	2	UNI	ACD	ENG	4.274	1238	3	0	0	0	0	3	2.423	2.423	0.000	0.000	0.000
188	CAN	S1A-064	A	F	51+	4	2	UNI	ACD	ENG	4.274	136	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
187	CAN	S1A-064	B	F	51+	4	2	UNI	ACD	ENG	4.274	566	2	1	0	0	0	3	5.300	3.534	1.767	0.000	0.000
190	CAN	S1A-065	A	F	51+	4	2	UNI	ACD	ENG	0.993	467	0	1	0	0	0	1	2.141	0.000	2.141	0.000	0.000
191	CAN	S1A-065	B	M	51+	4	2	UNI	CLM	ENG	0.993	1547	0	0	0	1	0	1	0.646	0.000	0.000	0.000	0.646
192	CAN	S1A-066	A	F	51+	4	2	UNI	ACD	ENG	1.473	683	0	0	0	0	1	1	0.000	0.000	0.000	0.000	0.000
193	CAN	S1A-066	B	M	51+	4	2	UNI	CLM	ENG	1.473	1353	2	0	0	0	0	2	1.478	1.478	0.000	0.000	0.000
194	CAN	S1A-067	A	F	51+	4	2	UNI	ACD	ENG	0.000	1022	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
195	CAN	S1A-067	B	M	51+	4	2	UNI	CLM	ENG	0.000	1185	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
196	CAN	S1A-068	A	F	51+	4	2	UNI	ACD	ENG	0.491	1025	1	0	0	0	0	1	0.976	0.976	0.000	0.000	0.000
197	CAN	S1A-068	B	M	51+	4	2	UNI	CLM	ENG	0.491	1012	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

198	CAN	S1A-069	A	F	51+	4	2	UNI	ACD	ENG	6.732	322	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
200	CAN	S1A-069	A	F	31-40	3	2	SEC	SLM	ENG	6.732	774	2	2	0	0	0	4	5.168	2.584	2.584	0.000	0.000
201	CAN	S1A-069	B	F	31-40	3	2	UNI	ACD	ENG	6.732	860	6	1	0	0	0	7	8.140	6.977	1.163	0.000	0.000
202	CAN	S1A-070	A	F	31-40	3	2	SEC	SLM	ENG	4.968	844	1	0	0	0	0	1	1.185	1.185	0.000	0.000	0.000
203	CAN	S1A-070	B	F	31-40	3	2	UNI	ACD	ENG	4.968	1169	5	3	0	1	0	9	7.699	4.277	2.566	0.000	0.855
204	CAN	S1A-071	A	F	31-40	3	2	SEC	SLM	ENG	2.959	651	2	0	0	0	0	2	3.072	3.072	0.000	0.000	0.000
205	CAN	S1A-071	B	F	31-40	3	2	UNI	ACD	ENG	2.959	1377	1	3	0	0	0	4	2.905	0.726	2.179	0.000	0.000
206	CAN	S1A-072	A	F	31-40	3	2	SEC	SLM	ENG	6.354	778	2	1	0	0	0	3	3.856	2.571	1.285	0.000	0.000
207	CAN	S1A-072	B	F	31-40	3	2	UNI	ACD	ENG	6.354	1268	5	4	0	1	0	10	7.886	3.943	3.155	0.000	0.789
208	CAN	S1A-073	A	F	31-40	3	2	SEC	SLM	ENG	8.258	347	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
210	CAN	S1A-073	A	M	25-30	2	2	UNI	CLM	ENG	8.258	507	0	0	0	1	0	1	1.972	0.000	0.000	0.000	1.972
209	CAN	S1A-073	B	F	31-40	3	2	UNI	ACD	ENG	6.277	673	1	4	0	0	0	5	7.429	1.486	5.944	0.000	0.000
211	CAN	S1A-073	B	F	25-30	2	2	UNI	CLM	ENG	8.258	538	3	2	0	0	0	5	9.294	5.576	3.717	0.000	0.000
212	CAN	S1A-074	A	M	25-30	2	2	UNI	CLM	ENG	6.277	790	0	2	0	0	0	2	2.532	0.000	2.532	0.000	0.000
213	CAN	S1A-074	B	F	25-30	2	2	UNI	CLM	ENG	6.277	1281	8	3	0	0	0	11	8.587	6.245	2.342	0.000	0.000
214	CAN	S1A-075	A	M	25-30	2	2	UNI	CLM	ENG	6.109	867	1	3	0	1	0	5	5.767	1.153	3.460	0.000	1.153
215	CAN	S1A-075	B	F	25-30	2	2	UNI	CLM	ENG	6.109	1261	5	1	1	1	0	8	6.344	3.965	0.793	0.793	0.793
216	CAN	S1A-076	A	M	25-30	2	2	UNI	CLM	ENG	3.914	636	0	2	0	0	0	2	3.145	0.000	3.145	0.000	0.000
218	CAN	S1A-076	A	F	25-30	2	2	SEC	SLM	ENG	3.914	399	2	0	0	0	0	2	5.013	5.013	0.000	0.000	0.000
217	CAN	S1A-076	B	F	25-30	2	2	UNI	CLM	ENG	9.009	795	2	1	0	1	0	4	5.031	2.516	1.258	0.000	1.258
219	CAN	S1A-076	B	M	31-40	3	2	SEC	SLM	ENG	3.914	227	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
220	CAN	S1A-077	A	F	31-40	3	2	UNI	CLM	ENG	9.009	1460	2	11	0	4	0	17	11.644	1.370	7.534	0.000	2.740
221	CAN	S1A-077	B	F	41-50	4	2	UNI	CLM	ENG	9.009	538	1	0	0	0	0	1	1.859	1.859	0.000	0.000	0.000
223	CAN	S1A-078	B	F	41-50	4	2	UNI	CLM	ENG	3.841	781	2	1	0	0	0	3	3.841	2.561	1.280	0.000	0.000
224	CAN	S1A-079	A	F	31-40	3	2	UNI	CLM	ENG	12.776	832	6	8	1	2	1	18	20.433	7.212	9.615	1.202	2.404

225	CAN	S1A-079	B	F	41-50	4	2	UNI	CLM	ENG	12.776	1203	2	4	0	2	0	8	6.650	1.663	3.325	0.000	1.663
226	CAN	S1A-080	A	F	31-40	3	2	UNI	CLM	ENG	12.395	1406	7	12	2	1	2	24	15.647	4.979	8.535	1.422	0.711
227	CAN	S1A-080	B	F	41-50	4	2	UNI	CLM	ENG	12.395	611	0	1	0	0	0	1	1.637	0.000	1.637	0.000	0.000
229	CAN	S1A-081	B	F	41-50	4	2	UNI	CLM	ENG	0.851	1175	1	0	0	0	0	1	0.851	0.851	0.000	0.000	0.000
232	CAN	S1A-082	A	F	25-30	2	2	SEC	SLM	ENG	3.984	881	3	3	0	0	0	6	6.810	3.405	3.405	0.000	0.000
231	CAN	S1A-082	B	F	41-50	4	2	UNI	CLM	ENG	7.314	291	1	1	0	0	0	2	6.873	3.436	3.436	0.000	0.000
233	CAN	S1A-082	B	M	31-40	3	2	SEC	SLM	ENG	3.984	462	1	0	0	0	0	1	2.165	2.165	0.000	0.000	0.000
234	CAN	S1A-083	A	F	25-30	2	2	SEC	SLM	ENG	7.314	981	2	2	0	3	1	8	7.136	2.039	2.039	0.000	3.058
235	CAN	S1A-083	B	M	31-40	3	2	SEC	SLM	ENG	7.314	1070	1	5	0	1	0	7	6.542	0.935	4.673	0.000	0.935
236	CAN	S1A-084	A	F	25-30	2	2	SEC	SLM	ENG	13.455	846	4	5	1	0	1	11	11.820	4.728	5.910	1.182	0.000
237	CAN	S1A-084	B	M	31-40	3	2	SEC	SLM	ENG	13.455	1235	7	5	0	4	1	17	12.955	5.668	4.049	0.000	3.239
238	CAN	S1A-085	A	F	25-30	2	2	SEC	SLM	ENG	5.780	932	0	2	1	0	0	3	3.219	0.000	2.146	1.073	0.000
240	CAN	S1A-085	A	F	51+	4	3	SEC	CLM	ENG	5.780	259	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
239	CAN	S1A-085	B	M	31-40	3	2	SEC	SLM	ENG	7.346	739	2	1	0	2	0	5	6.766	2.706	1.353	0.000	2.706
241	CAN	S1A-085	B	M	31-40	3	3	UNI	ACD	ENG	5.780	126	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
242	CAN	S1A-086	A	F	51+	4	3	SEC	CLM	ENG	7.346	856	2	1	0	0	0	3	3.505	2.336	1.168	0.000	0.000
243	CAN	S1A-086	B	M	31-40	3	3	UNI	ACD	ENG	7.346	1186	7	3	1	1	0	12	10.118	5.902	2.530	0.843	0.843
244	CAN	S1A-087	A	F	51+	4	3	SEC	CLM	ENG	3.316	1258	1	0	0	1	1	3	1.590	0.795	0.000	0.000	0.795
245	CAN	S1A-087	B	M	31-40	3	3	UNI	ACD	ENG	3.316	853	1	2	0	1	0	4	4.689	1.172	2.345	0.000	1.172
246	CAN	S1A-088	A	F	31-40	3	2	UNI	CLM	ENG+	3.651	894	1	0	0	1	0	2	2.237	1.119	0.000	0.000	1.119
247	CAN	S1A-088	B	F	41-50	4	2	SEC	SLM	OTH	3.651	1297	4	1	0	1	0	6	4.626	3.084	0.771	0.000	0.771
248	CAN	S1A-089	A	F	51+	4	3	SEC	CLM	ENG	3.722	175	0	0	0	1	0	1	5.714	0.000	0.000	0.000	5.714
250	CAN	S1A-089	A	F	31-40	3	2	UNI	CLM	ENG+	3.722	1250	2	2	0	2	0	6	4.800	1.600	1.600	0.000	1.600
251	CAN	S1A-089	B	F	41-50	4	2	SEC	SLM	OTH	3.722	362	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
252	CAN	S1A-090	A	F	31-40	3	2	UNI	CLM	ENG+	3.945	1523	3	1	0	1	0	5	3.283	1.970	0.657	0.000	0.657

253	CAN	S1A-090	B	F	41-50	4	2	SEC	SLM	OTH	3.945	505	2	0	1	0	0	3	5.941	3.960	0.000	1.980	0.000
255	CAN	S1A-091	B	M	Okt 24	1	1	UNI	CLM	ENG	6.061	495	0	2	0	1	0	3	6.061	0.000	4.040	0.000	2.020
256	CAN	S1A-092	A	F	20-24	1	1	UNI	EDU	ENG+	14.905	1358	7	9	1	2	1	20	13.991	5.155	6.627	0.736	1.473
257	CAN	S1A-092	B	M	Okt 24	1	1	UNI	CLM	ENG	14.905	856	0	10	1	1	1	13	14.019	0.000	11.682	1.168	1.168
258	CAN	S1A-093	A	F	20-24	1	1	SEC	EDU	ENG	3.751	850	0	0	0	2	0	2	2.353	0.000	0.000	0.000	2.353
259	CAN	S1A-093	B	M	25-30	2	1	UNI	EDU	ENG	3.751	1283	4	1	1	0	0	6	4.677	3.118	0.779	0.779	0.000
260	CAN	S1A-094	A	F	20-24	1	1	UNI	EDU	ENG+	9.162	1597	7	9	1	1	0	18	11.271	4.383	5.636	0.626	0.626
261	CAN	S1A-094	B	M	Okt 24	1	1	UNI	CLM	ENG	9.162	586	2	0	0	0	0	2	3.413	3.413	0.000	0.000	0.000
262	CAN	S1A-095	A	F	20-24	1	1	UNI	EDU	ENG+	10.155	836	7	2	0	1	0	10	11.962	8.373	2.392	0.000	1.196
263	CAN	S1A-095	B	M	Okt 24	1	1	UNI	CLM	ENG	10.155	1429	7	4	1	1	0	13	9.097	4.899	2.799	0.700	0.700
264	CAN	S1A-096	A	F	20-24	1	1	SEC	EDU	ENG	5.537	616	2	1	0	0	0	3	4.870	3.247	1.623	0.000	0.000
266	CAN	S1A-096	A	F	20-24	1	1	UNI	EDU	ENG	5.537	546	3	6	0	0	0	9	16.484	5.495	10.989	0.000	0.000
265	CAN	S1A-096	B	M	25-30	2	1	UNI	EDU	ENG	9.059	709	1	3	0	0	0	4	5.642	1.410	4.231	0.000	0.000
267	CAN	S1A-096	B	F	20-24	1	1	SEC	EDU	ENG	5.537	194	0	1	0	0	0	1	5.155	0.000	5.155	0.000	0.000
270	CAN	S1A-097	A	M	Okt 24	1	1	SEC	EDU	ENG	9.059	653	0	1	0	0	0	1	1.531	0.000	1.531	0.000	0.000
269	CAN	S1A-097	B	M	Okt 24	1	1	UNI	CLM	ENG	6.386	458	3	2	0	0	0	5	10.917	6.550	4.367	0.000	0.000
271	CAN	S1A-097	B	F	20-24	1	1	SEC	EDU	ENG	9.059	977	2	5	0	1	0	8	8.188	2.047	5.118	0.000	1.024
273	CAN	S1A-098	B	F	20-24	1	1	SEC	EDU	ENG	6.386	783	1	3	1	0	0	5	6.386	1.277	3.831	1.277	0.000
274	CAN	S1A-099	A	M	Okt 24	1	1	SEC	EDU	ENG	1.362	734	1	0	0	0	0	1	1.362	1.362	0.000	0.000	0.000
276	CAN	S1A-100	A	M	Okt 24	1	1	SEC	EDU	ENG	14.423	590	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
277	CAN	S1A-100	B	F	20-24	1	1	SEC	EDU	ENG	14.423	1490	12	16	1	1	0	30	20.134	8.054	10.738	0.671	0.671

Table 140: Final data set of EngE.

ID	VAR	FILE	SPEAKER	SEX	AGE	AGEG	education	OCC	PAI	words	ini.n	med.n	fin.m	non.n	na.n	all.n	all	ini	med	fin	non
278	GB	S1A-001	A	M	26-45	3	UNI	ACD	0.000	286	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
279	GB	S1A-001	B	M	18-25	1	SEC	EDU	0.000	1758	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
281	GB	S1A-002	A	M	26-45	3	UNI	ACD	0.000	177	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
284	GB	S1A-002	A	M	26-45	3	UNI	ACD	0.000	147	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
282	GB	S1A-002	B	F	18-25	1	SEC	EDU	0.000	489	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
285	GB	S1A-002	B	F	18-25	1	SEC	EDU	0.000	502	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
283	GB	S1A-002	C	F	18-25	1	SEC	EDU	0.000	736	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
286	GB	S1A-003	A	M	26-45	3	UNI	ACD	0.000	546	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
287	GB	S1A-003	B	F	18-25	1	SEC	EDU	0.000	1589	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
289	GB	S1A-004	A	M	26-45	3	UNI	ACD	0.000	405	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
290	GB	S1A-004	B	M	18-25	1	SEC	EDU	0.000	1668	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
292	GB	S1A-005	A	F	18-25	1	UNI	EDU	0.464	1234	1	0	0	0	0	1	0.810	0.810	0.000	0.000	0.000
293	GB	S1A-005	B	F	18-25	1	UNI	EDU	0.464	919	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
294	GB	S1A-006	A	F	18-25	1	UNI	NA	0.000	1062	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
295	GB	S1A-006	B	M	18-25	1	UNI	NA	0.000	1036	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
296	GB	S1A-007	A	M	46+	4	NA	NA	0.000	641	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
297	GB	S1A-007	B	F	46+	4	NA	NA	0.000	990	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
298	GB	S1A-007	C	F	18-25	1	UNI	EDU	0.000	276	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
301	GB	S1A-008	A	M	18-25	1	UNI	NA	1.356	1438	1	1	0	0	0	2	1.391	0.695	0.695	0.000	0.000
302	GB	S1A-008	B	M	18-25	1	UNI	EDU	1.356	775	0	0	1	0	0	1	1.290	0.000	0.000	1.290	0.000
303	GB	S1A-009	A	F	46+	4	UNI	ACD	0.000	857	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
304	GB	S1A-009	B	M	26-45	3	SEC	CLM	0.000	1136	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
305	GB	S1A-010	A	F	46+	4	SEC	CLM	0.000	651	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

306	GB	S1A-010	B	F	26-45	3	SEC	CLM	0.000	1413	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
308	GB	S1A-011	A	F	26-45	3	UNI	ACD	0.000	243	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
311	GB	S1A-011	A	F	26-45	3	UNI	ACD	0.000	563	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
309	GB	S1A-011	B	F	26-45	3	UNI	ACD	0.000	267	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
312	GB	S1A-011	B	F	26-45	3	UNI	ACD	0.000	558	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
310	GB	S1A-011	C	M	26-45	3	UNI	CLM	0.000	318	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
313	GB	S1A-012	A	M	26-45	3	UNI	ACD	0.000	710	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
314	GB	S1A-012	B	F	26-45	3	UNI	ACD	0.000	391	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
315	GB	S1A-012	C	M	26-45	3	UNI	ACD	0.000	569	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
316	GB	S1A-012	D	M	26-45	3	UNI	ACD	0.000	100	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
318	GB	S1A-013	A	M	26-45	3	UNI	ACD	0.000	233	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
319	GB	S1A-013	B	F	46+	4	SEC	SLM	0.000	630	2	1	0	0	0	3	4.762	3.175	1.587	0.000	0.000
320	GB	S1A-013	C	M	46+	4	SEC	SLM	0.000	535	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
322	GB	S1A-013	E	M	NA	NA	NA	NA	0.000	593	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
323	GB	S1A-013	F	M	NA	NA	NA	NA	0.000	157	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
325	GB	S1A-014	A	F	18-25	1	UNI	EDU	1.190	365	0	1	0	0	0	1	2.740	0.000	2.740	0.000	0.000
326	GB	S1A-014	B	F	26-45	3	UNI	CLM	1.190	953	1	1	0	0	0	2	2.099	1.049	1.049	0.000	0.000
327	GB	S1A-014	C	M	18-25	1	UNI	EDU	1.190	728	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
329	GB	S1A-015	A	M	18-25	1	UNI	EDU	3.000	1040	0	1	0	0	0	1	0.962	0.000	0.962	0.000	0.000
330	GB	S1A-015	B	F	18-25	1	UNI	EDU	3.000	960	2	2	0	1	0	5	5.208	2.083	2.083	0.000	1.042
331	GB	S1A-016	A	M	26-45	3	UNI	ACD	0.000	326	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
332	GB	S1A-016	B	NA	NA	NA	NA	NA	0.000	224	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
333	GB	S1A-016	C	NA	NA	NA	NA	NA	0.000	297	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
334	GB	S1A-016	D	NA	NA	NA	NA	NA	0.000	629	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
335	GB	S1A-016	E	NA	NA	NA	NA	NA	0.000	570	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

337	GB	S1A-017	A	M	26-45	3	UNI	CLM	0.000	823	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
338	GB	S1A-017	B	F	26-45	3	UNI	CLM	0.000	816	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
339	GB	S1A-017	C	F	18-25	1	UNI	EDU	0.000	189	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
340	GB	S1A-018	A	F	18-25	1	SEC	EDU	0.000	1287	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
341	GB	S1A-018	B	F	18-25	1	NA	NA	0.000	741	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
343	GB	S1A-019	A	F	18-25	1	UNI	EDU	0.000	1157	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
344	GB	S1A-019	B	M	18-25	1	UNI	EDU	0.000	141	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
345	GB	S1A-019	C	F	26-45	3	UNI	EDU	0.000	415	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
346	GB	S1A-019	D	F	18-25	1	SEC	EDU	0.000	186	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
347	GB	S1A-019	E	F	18-25	1	SEC	EDU	0.000	237	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
348	GB	S1A-019	F	F	18-25	1	UNI	EDU	0.000	125	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
349	GB	S1A-020	A	M	NA	NA	NA	NA	0.000	418	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
350	GB	S1A-020	B	M	NA	NA	NA	NA	0.000	873	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
351	GB	S1A-020	C	M	NA	NA	NA	NA	0.000	540	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
352	GB	S1A-020	D	F	26-45	3	UNI	ACD	0.000	160	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
353	GB	S1A-021	A	M	26-45	3	UNI	ACD	0.000	204	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
358	GB	S1A-021	A	F	26-45	3	UNI	ACD	0.000	225	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
359	GB	S1A-021	B	M	26-45	3	UNI	ACD	0.000	368	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
355	GB	S1A-021	C	M	26-45	3	UNI	ACD	0.000	340	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
360	GB	S1A-021	C	F	26-45	3	UNI	ACD	0.000	153	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
356	GB	S1A-021	D	F	26-45	3	UNI	ACD	0.000	283	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
361	GB	S1A-021	D	M	26-45	3	UNI	ACD	0.000	202	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
357	GB	S1A-021	Z	NA	NA	NA	NA	NA	0.000	126	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
362	GB	S1A-022	A	F	18-25	1	SEC	EDU	0.809	393	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
363	GB	S1A-022	B	F	46+	4	UNI	ACD	0.809	251	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

364	GB	S1A-022	C	M	46+	4	UNI	ACD	0.809	360	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
365	GB	S1A-022	D	F	18-25	1	UNI	EDU	0.809	876	0	0	0	1	0	1	1.142	0.000	0.000	0.000	1.142
366	GB	S1A-023	A	M	46+	4	SEC	CLM	0.000	1098	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
367	GB	S1A-023	B	F	46+	4	SEC	CLM	0.000	1089	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
368	GB	S1A-024	A	M	46+	4	UNI	ACD	0.000	732	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
369	GB	S1A-024	B	M	46+	4	UNI	ACD	0.000	1435	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
370	GB	S1A-025	A	M	26-45	3	SEC	NA	0.000	917	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
371	GB	S1A-025	B	F	26-45	3	SEC	SLM	0.000	1257	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
372	GB	S1A-026	A	M	26-45	3	UNI	ACD	0.000	1503	2	2	0	0	0	4	2.661	1.331	1.331	0.000	0.000
374	GB	S1A-026	C	M	26-45	3	UNI	ACD	0.000	304	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
375	GB	S1A-026	D	M	26-45	3	UNI	ACD	0.000	184	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
377	GB	S1A-027	B	M	26-45	3	UNI	CLM	0.000	651	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
378	GB	S1A-027	C	M	26-45	3	UNI	ACD	0.000	552	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
379	GB	S1A-027	D	F	26-45	3	UNI	SLM	0.000	527	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
380	GB	S1A-028	A	M	46+	4	SEC	SLM	0.000	1676	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
381	GB	S1A-028	B	M	46+	4	UNI	ACD	0.000	319	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
382	GB	S1A-028	C	F	18-25	1	UNI	EDU	0.000	117	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
385	GB	S1A-029	A	M	46+	4	UNI	ACD	0.000	1723	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
386	GB	S1A-029	B	M	18-25	1	UNI	CLM	0.000	687	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
387	GB	S1A-029	C	M	18-25	1	UNI	CLM	0.000	709	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
388	GB	S1A-029	D	M	18-25	1	UNI	CLM	0.000	364	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
390	GB	S1A-030	A	NA	NA	NA	NA	NA	3.929	903	2	3	0	2	0	7	7.752	2.215	3.322	0.000	2.215
391	GB	S1A-030	B	NA	NA	NA	NA	NA	3.929	601	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
392	GB	S1A-030	C	NA	NA	NA	NA	NA	3.929	417	0	1	0	3	0	4	9.592	0.000	2.398	0.000	7.194
395	GB	S1A-031	B	F	26-45	3	UNI	ACD	0.000	1174	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

396	GB	S1A-032	A	F	18-25	1	UNI	ACD	0.000	281	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
401	GB	S1A-032	A	F	18-25	1	UNI	EDU	0.000	180	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
397	GB	S1A-032	B	F	46+	4	UNI	ACD	0.000	349	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
402	GB	S1A-032	B	F	18-25	1	SEC	EDU	0.000	228	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
398	GB	S1A-032	C	M	46+	4	NA	CLM	0.000	134	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
403	GB	S1A-032	C	F	46+	4	UNI	ACD	0.000	159	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
399	GB	S1A-032	D	F	18-25	1	SEC	EDU	0.000	352	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
406	GB	S1A-033	A	M	26-45	3	UNI	CLM	0.997	1442	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
407	GB	S1A-033	B	M	18-25	1	SEC	EDU	0.997	565	1	0	0	1	0	2	3.540	1.770	0.000	0.000	1.770
408	GB	S1A-034	A	F	26-45	3	UNI	CLM	0.000	512	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
410	GB	S1A-035	A	F	26-45	3	UNI	CLM	1.480	934	0	1	0	0	0	1	1.071	0.000	1.071	0.000	0.000
411	GB	S1A-035	B	F	18-25	1	SEC	EDU	1.480	1093	0	1	0	1	0	2	1.830	0.000	0.915	0.000	0.915
412	GB	S1A-036	A	F	26-45	3	SEC	SLM	0.481	596	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
413	GB	S1A-036	B	F	26-45	3	SEC	SLM	0.481	1482	1	0	0	0	0	1	0.675	0.675	0.000	0.000	0.000
414	GB	S1A-037	A	F	18-25	1	UNI	EDU	1.801	444	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
415	GB	S1A-037	B	F	26-45	3	UNI	NA	1.801	1777	1	3	0	0	0	4	2.251	0.563	1.688	0.000	0.000
416	GB	S1A-038	A	F	18-25	1	UNI	ACD	0.528	1001	0	0	0	1	0	1	0.999	0.000	0.000	0.000	0.999
418	GB	S1A-038	C	M	18-25	1	UNI	EDU	0.528	892	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
419	GB	S1A-039	A	F	18-25	1	UNI	ACD	0.484	1025	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
420	GB	S1A-039	B	F	18-25	1	UNI	ACD	0.484	1040	0	1	0	0	0	1	0.962	0.000	0.962	0.000	0.000
421	GB	S1A-040	A	F	18-25	1	UNI	EDU	1.203	710	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
422	GB	S1A-040	B	F	18-25	1	UNI	EDU	1.203	516	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
423	GB	S1A-040	C	F	18-25	1	UNI	ACD	1.203	366	0	1	0	0	0	1	2.732	0.000	2.732	0.000	0.000
424	GB	S1A-040	D	F	18-25	1	UNI	EDU	1.203	577	0	1	0	0	0	1	1.733	0.000	1.733	0.000	0.000
425	GB	S1A-040	E	M	18-25	1	UNI	CLM	1.203	254	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

427	GB	S1A-041	A	M	18-25	1	UNI	EDU	0.453	1273	0	0	0	1	0	1	0.786	0.000	0.000	0.000	0.786
428	GB	S1A-041	B	M	18-25	1	UNI	EDU	0.453	936	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
429	GB	S1A-042	A	F	18-25	1	UNI	EDU	0.727	621	1	0	0	0	0	1	1.610	1.610	0.000	0.000	0.000
430	GB	S1A-042	B	F	18-25	1	UNI	EDU	0.727	802	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
431	GB	S1A-042	C	F	18-25	1	UNI	EDU	0.727	573	0	1	0	0	0	1	1.745	0.000	1.745	0.000	0.000
432	GB	S1A-043	A	M	26-45	3	UNI	ACD	0.000	1128	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
433	GB	S1A-043	B	M	26-45	3	SEC	CLM	0.000	693	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
434	GB	S1A-044	A	F	46+	4	UNI	CLM	0.000	1069	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
435	GB	S1A-044	B	M	26-45	3	UNI	ACD	0.000	822	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
436	GB	S1A-045	A	M	26-45	3	UNI	ACD	0.000	553	0	0	0	1	0	1	1.808	0.000	0.000	0.000	1.808
438	GB	S1A-045	A	M	26-45	3	UNI	NA	0.000	198	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
440	GB	S1A-045	A	M	26-45	3	UNI	ACD	0.000	215	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
437	GB	S1A-045	B	F	46+	4	UNI	CLM	0.588	638	0	0	0	2	0	2	3.135	0.000	0.000	0.000	3.135
439	GB	S1A-045	B	M	26-45	3	UNI	ACD	0.588	336	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
441	GB	S1A-045	B	M	26-45	3	UNI	NA	0.000	150	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
442	GB	S1A-046	A	F	18-25	1	NA	NA	0.588	854	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
443	GB	S1A-046	B	M	46+	4	NA	NA	0.588	207	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
444	GB	S1A-046	C	F	18-25	1	NA	NA	0.588	768	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
446	GB	S1A-046	Z	NA	NA	NA	NA	NA	0.588	933	1	0	0	0	0	1	1.072	1.072	0.000	0.000	0.000
447	GB	S1A-047	A	M	26-45	3	UNI	CLM	0.000	862	2	1	0	1	0	4	4.640	2.320	1.160	0.000	1.160
448	GB	S1A-047	B	F	26-45	3	SEC	EDU	0.000	967	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
449	GB	S1A-047	C	F	46+	4	SEC	SLM	0.000	309	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
450	GB	S1A-048	A	F	26-45	3	UNI	CLM	2.625	427	0	0	0	1	0	1	2.342	0.000	0.000	0.000	2.342
451	GB	S1A-048	B	F	26-45	3	SEC	SLM	2.625	664	1	0	0	2	0	3	4.518	1.506	0.000	0.000	3.012
452	GB	S1A-048	C	F	26-45	3	SEC	EDU	2.625	860	0	0	0	1	0	1	1.163	0.000	0.000	0.000	1.163

453	GB	S1A-049	A	F	26-45	3	SEC	CLM	1.279	435	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
454	GB	S1A-049	B	F	26-45	3	SEC	SLM	1.279	1138	1	1	0	0	0	2	1.757	0.879	0.879	0.000	0.000
455	GB	S1A-049	C	F	26-45	3	SEC	SLM	1.279	426	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
456	GB	S1A-050	A	M	46+	4	UNI	CLM	0.450	569	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
457	GB	S1A-050	B	F	18-25	1	SEC	EDU	0.450	1653	0	1	0	0	0	1	0.605	0.000	0.605	0.000	0.000
458	GB	S1A-051	A	M	46+	4	UNI	ACD	0.000	193	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
460	GB	S1A-051	A	M	46+	4	UNI	ACD	0.000	218	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
462	GB	S1A-051	A	M	46+	4	UNI	ACD	0.000	187	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
464	GB	S1A-051	A	M	46+	4	UNI	ACD	0.000	372	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
459	GB	S1A-051	B	F	46+	4	SEC	SLM	4.912	315	0	1	0	0	0	1	3.175	0.000	3.175	0.000	0.000
461	GB	S1A-051	B	M	18-25	1	SEC	EDU	4.912	453	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
465	GB	S1A-051	B	M	18-25	1	SEC	EDU	0.000	155	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
469	GB	S1A-052	A	M	46+	4	SEC	SLM	4.912	859	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
467	GB	S1A-052	B	M	46+	4	SEC	SLM	0.860	666	2	2	0	1	0	5	7.508	3.003	3.003	0.000	1.502
470	GB	S1A-052	B	M	26-45	3	UNI	ACD	4.912	352	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
472	GB	S1A-053	A	M	26-45	3	UNI	EDU	0.860	1021	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
473	GB	S1A-053	B	F	26-45	3	UNI	EDU	0.860	1024	0	1	0	0	0	1	0.977	0.000	0.977	0.000	0.000
474	GB	S1A-053	C	M	46+	4	SEC	CLM	0.860	139	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
475	GB	S1A-054	A	F	46+	4	UNI	ACD	0.911	1596	1	1	0	0	0	2	1.253	0.627	0.627	0.000	0.000
476	GB	S1A-054	B	F	26-45	3	UNI	ACD	0.911	600	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
477	GB	S1A-055	A	M	26-45	3	NA	NA	0.983	487	1	0	0	0	0	1	2.053	2.053	0.000	0.000	0.000
478	GB	S1A-055	B	F	26-45	3	NA	NA	0.983	833	0	1	0	0	0	1	1.200	0.000	1.200	0.000	0.000
480	GB	S1A-055	D	F	18-25	1	UNI	CLM	0.983	184	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
482	GB	S1A-056	A	M	26-45	3	UNI	ACD	0.000	470	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
486	GB	S1A-056	A	F	26-45	3	UNI	ACD	0.000	218	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

490	GB	S1A-056	A	F	26-45	3	UNI	ACD	0.000	114	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
494	GB	S1A-056	A	M	26-45	3	UNI	ACD	0.000	191	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
487	GB	S1A-056	B	M	26-45	3	UNI	ACD	0.000	427	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
495	GB	S1A-056	B	F	26-45	3	UNI	ACD	0.000	369	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
485	GB	S1A-056	Z	NA	NA	NA	NA	NA	0.000	291	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
489	GB	S1A-056	Z	NA	NA	NA	NA	NA	0.000	125	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
497	GB	S1A-056	Z	NA	NA	NA	NA	NA	0.000	115	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
498	GB	S1A-057	A	F	26-45	3	UNI	ACD	0.000	546	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
499	GB	S1A-057	B	M	46+	4	UNI	ACD	0.000	1429	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
500	GB	S1A-057	C	M	26-45	3	UNI	ACD	0.000	155	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
501	GB	S1A-058	A	M	26-45	3	NA	NA	0.000	146	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
505	GB	S1A-058	A	M	26-45	3	NA	NA	0.000	371	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
509	GB	S1A-058	A	M	26-45	3	UNI	ACD	0.000	156	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
506	GB	S1A-058	B	M	26-45	3	NA	NA	0.000	313	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
510	GB	S1A-058	B	F	26-45	3	UNI	CLM	0.000	818	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
503	GB	S1A-058	C	F	26-45	3	UNI	CLM	0.000	111	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
511	GB	S1A-059	A	M	46+	4	UNI	CLM	0.000	1020	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
512	GB	S1A-059	B	M	18-25	1	SEC	EDU	0.000	982	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
513	GB	S1A-060	A	NA	NA	NA	NA	NA	0.461	455	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
514	GB	S1A-060	B	NA	NA	NA	NA	NA	0.461	1714	1	0	0	0	0	1	0.583	0.583	0.000	0.000	0.000
515	GB	S1A-061	A	M	26-45	3	SEC	CLM	0.000	880	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
516	GB	S1A-061	B	M	46+	4	UNI	CLM	0.000	1230	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
517	GB	S1A-062	A	M	46+	4	UNI	CLM	0.990	909	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
518	GB	S1A-062	B	F	18-25	1	SEC	EDU	0.990	1112	2	0	0	0	0	2	1.799	1.799	0.000	0.000	0.000
519	GB	S1A-063	A	M	26-45	3	UNI	ACD	0.000	1202	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

520	GB	S1A-063	B	F	26-45	3	UNI	ACD	0.000	230	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
521	GB	S1A-063	C	M	26-45	3	UNI	ACD	0.000	211	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
522	GB	S1A-063	D	F	26-45	3	SEC	SLM	0.000	274	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
523	GB	S1A-064	A	F	18-25	1	SEC	EDU	0.988	1007	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
524	GB	S1A-064	B	M	18-25	1	SEC	EDU	0.988	283	1	0	0	0	0	1	3.534	3.534	0.000	0.000	0.000
525	GB	S1A-064	C	F	18-25	1	SEC	EDU	0.988	729	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
526	GB	S1A-065	A	F	26-45	3	UNI	EDU	0.000	416	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
527	GB	S1A-065	B	F	26-45	3	SEC	SLM	0.000	332	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
528	GB	S1A-065	C	M	46+	4	UNI	EDU	0.000	307	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
530	GB	S1A-065	E	M	46+	4	UNI	SLM	0.000	478	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
531	GB	S1A-065	F	M	18-25	1	UNI	EDU	0.000	195	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
533	GB	S1A-066	A	M	46+	4	UNI	CLM	0.483	856	0	1	0	0	0	1	1.168	0.000	1.168	0.000	0.000
534	GB	S1A-066	B	F	26-45	3	UNI	EDU	0.483	1215	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
535	GB	S1A-067	A	F	26-45	3	UNI	CLM	1.926	998	0	1	0	0	0	1	1.002	0.000	1.002	0.000	0.000
536	GB	S1A-067	B	F	26-45	3	UNI	ACD	1.926	1079	1	0	1	1	0	3	2.780	0.927	0.000	0.927	0.927
537	GB	S1A-068	A	M	18-25	1	SEC	EDU	0.000	979	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
538	GB	S1A-068	B	F	18-25	1	SEC	EDU	0.000	733	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
539	GB	S1A-068	C	M	18-25	1	SEC	EDU	0.000	236	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
540	GB	S1A-069	A	M	18-25	1	SEC	EDU	1.095	479	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
542	GB	S1A-069	A	M	18-25	1	SEC	EDU	1.095	744	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
541	GB	S1A-069	B	F	18-25	1	SEC	EDU	0.000	382	0	1	0	0	0	1	2.618	0.000	2.618	0.000	0.000
543	GB	S1A-069	B	F	18-25	1	SEC	EDU	1.095	531	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
544	GB	S1A-070	A	M	18-25	1	SEC	EDU	0.000	843	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
545	GB	S1A-070	B	M	18-25	1	SEC	EDU	0.000	1013	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
546	GB	S1A-071	A	F	18-25	1	NA	NA	1.148	427	1	0	0	0	0	1	2.342	2.342	0.000	0.000	0.000

547	GB	S1A-071	B	M	18-25	1	UNI	EDU	1.148	444	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
550	GB	S1A-072	A	F	26-45	3	UNI	NA	0.508	691	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
551	GB	S1A-072	B	M	26-45	3	NA	NA	0.508	1276	0	0	0	1	0	1	0.784	0.000	0.000	0.000	0.784
552	GB	S1A-073	A	M	46+	4	UNI	CLM	0.000	580	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
553	GB	S1A-073	B	M	46+	4	UNI	ACD	0.000	682	0	2	0	0	0	2	2.933	0.000	2.933	0.000	0.000
554	GB	S1A-073	C	F	26-45	3	UNI	ACD	0.000	521	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
555	GB	S1A-073	D	F	26-45	3	UNI	ACD	0.000	172	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
556	GB	S1A-074	A	F	18-25	1	SEC	EDU	0.000	548	0	2	0	0	0	2	3.650	0.000	3.650	0.000	0.000
559	GB	S1A-074	A	M	18-25	1	UNI	CLM	0.000	119	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
564	GB	S1A-074	A	M	18-25	1	SEC	CLM	0.000	185	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
572	GB	S1A-074	A	M	18-25	1	SEC	EDU	0.000	158	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
557	GB	S1A-074	B	M	18-25	1	SEC	CLM	0.000	203	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
558	GB	S1A-074	C	M	18-25	1	UNI	CLM	0.000	168	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
575	GB	S1A-075	A	F	26-45	3	UNI	NA	0.000	604	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
576	GB	S1A-075	B	M	26-45	3	UNI	CLM	0.000	1572	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
577	GB	S1A-076	A	F	26-45	3	UNI	NA	2.775	434	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
578	GB	S1A-076	B	M	26-45	3	UNI	NA	2.775	1728	1	3	0	2	0	6	3.472	0.579	1.736	0.000	1.157
579	GB	S1A-077	A	M	26-45	3	UNI	CLM	0.000	761	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
580	GB	S1A-077	B	F	26-45	3	UNI	CLM	0.000	1272	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
585	GB	S1A-078	A	M	26-45	3	UNI	CLM	0.000	205	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
588	GB	S1A-078	A	F	26-45	3	UNI	CLM	0.000	294	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
586	GB	S1A-078	B	F	26-45	3	UNI	CLM	0.000	465	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
589	GB	S1A-078	B	M	18-25	1	SEC	EDU	0.000	131	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
591	GB	S1A-078	B	F	26-45	3	UNI	CLM	0.000	151	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
587	GB	S1A-078	C	F	18-25	1	SEC	EDU	3.988	223	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

592	GB	S1A-078	C	F	18-25	1	SEC	EDU	0.000	249	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
593	GB	S1A-079	A	F	26-45	3	UNI	CLM	3.988	1118	0	1	0	0	0	1	0.894	0.000	0.894	0.000	0.000
594	GB	S1A-079	B	M	26-45	3	UNI	CLM	3.988	597	2	0	0	1	0	3	5.025	3.350	0.000	0.000	1.675
595	GB	S1A-079	C	F	26-45	3	UNI	EDU	3.988	406	0	1	0	0	0	1	2.463	0.000	2.463	0.000	0.000
596	GB	S1A-080	A	F	18-25	1	UNI	NA	0.000	742	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
597	GB	S1A-080	B	F	26-45	3	UNI	NA	0.000	1468	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
598	GB	S1A-081	A	M	26-45	3	UNI	ACD	0.377	1485	0	1	0	0	0	1	0.673	0.000	0.673	0.000	0.000
599	GB	S1A-081	B	F	26-45	3	SEC	SLM	0.377	1171	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
600	GB	S1A-082	A	F	26-45	3	UNI	ACD	1.202	1205	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
601	GB	S1A-082	B	M	18-25	1	SEC	EDU	1.202	723	0	0	0	1	0	1	1.383	0.000	0.000	0.000	1.383
602	GB	S1A-082	C	F	18-25	1	SEC	EDU	1.202	109	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
603	GB	S1A-083	A	F	26-45	3	SEC	SLM	0.000	985	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
604	GB	S1A-083	B	F	46+	4	SEC	SLM	0.000	1064	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
605	GB	S1A-084	A	M	18-25	1	UNI	ACD	2.515	949	1	2	0	1	0	4	4.215	1.054	2.107	0.000	1.054
606	GB	S1A-084	B	F	18-25	1	SEC	EDU	2.515	1039	0	1	0	0	0	1	0.962	0.000	0.962	0.000	0.000
608	GB	S1A-085	A	M	18-25	1	SEC	EDU	1.429	776	0	1	0	0	0	1	1.289	0.000	1.289	0.000	0.000
609	GB	S1A-085	B	F	18-25	1	SEC	EDU	1.429	1323	0	2	0	0	0	2	1.512	0.000	1.512	0.000	0.000
610	GB	S1A-086	A	F	NA	NA	NA	NA	0.000	1031	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
611	GB	S1A-086	B	F	26-45	3	UNI	ACD	0.000	691	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
612	GB	S1A-086	C	F	NA	NA	NA	NA	0.000	267	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
613	GB	S1A-087	A	M	46+	4	UNI	ACD	0.000	796	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
615	GB	S1A-087	A	M	46+	4	UNI	ACD	0.000	973	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
614	GB	S1A-087	B	M	26-45	3	UNI	ACD	0.000	206	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
616	GB	S1A-087	B	M	18-25	1	SEC	EDU	0.000	326	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
617	GB	S1A-088	A	M	46+	4	UNI	ACD	0.000	1431	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

618	GB	S1A-088	B	M	26-45	3	UNI	ACD	0.000	648	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
620	GB	S1A-089	A	M	46+	4	UNI	ACD	0.000	659	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
622	GB	S1A-089	A	M	46+	4	UNI	ACD	0.000	146	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
624	GB	S1A-089	A	M	46+	4	UNI	ACD	0.000	147	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
626	GB	S1A-089	A	M	46+	4	UNI	ACD	0.000	263	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
621	GB	S1A-089	B	M	26-45	3	UNI	ACD	0.000	319	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
625	GB	S1A-089	B	F	26-45	3	UNI	EDU	0.000	144	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
627	GB	S1A-089	B	F	18-25	1	SEC	EDU	0.000	232	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
632	GB	S1A-090	A	F	18-25	1	SEC	EDU	0.000	523	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
629	GB	S1A-090	B	F	46+	4	SEC	EDU	0.000	474	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
633	GB	S1A-090	B	F	46+	4	SEC	EDU	0.000	115	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
630	GB	S1A-090	C	F	46+	4	SEC	EDU	0.000	399	1	0	0	0	0	1	2.506	2.506	0.000	0.000	0.000
634	GB	S1A-090	C	F	18-25	1	SEC	EDU	0.000	191	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
631	GB	S1A-090	D	F	46+	4	SEC	EDU	0.000	235	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
635	GB	S1A-091	A	F	26-45	3	UNI	NA	0.000	806	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
636	GB	S1A-091	B	F	26-45	3	SEC	EDU	0.000	1328	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
637	GB	S1A-092	A	F	26-45	3	UNI	NA	0.000	958	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
638	GB	S1A-092	B	M	26-45	3	UNI	NA	0.000	984	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
639	GB	S1A-093	A	F	18-25	1	UNI	EDU	1.461	1241	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
640	GB	S1A-093	B	F	18-25	1	SEC	EDU	1.461	812	0	2	0	1	0	3	3.695	0.000	2.463	0.000	1.232
641	GB	S1A-094	A	F	46+	4	SEC	NA	0.000	1527	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
642	GB	S1A-094	B	F	18-25	1	UNI	EDU	0.000	654	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
643	GB	S1A-095	A	M	18-25	1	NA	NA	0.000	367	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
645	GB	S1A-095	A	F	46+	4	NA	NA	0.000	152	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
647	GB	S1A-095	A	M	18-25	1	NA	NA	0.000	122	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

649	GB	S1A-095	A	F	46+	4	NA	NA	0.000	318	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
644	GB	S1A-095	B	M	26-45	3	UNI	CLM	1.751	215	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
646	GB	S1A-095	B	M	26-45	3	UNI	CLM	1.751	118	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
648	GB	S1A-095	B	M	26-45	3	UNI	CLM	1.751	116	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
650	GB	S1A-095	B	M	26-45	3	UNI	CLM	0.000	345	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
651	GB	S1A-096	A	M	26-45	3	SEC	SLM	1.751	1302	0	4	0	0	0	4	3.072	0.000	3.072	0.000	0.000
652	GB	S1A-096	B	F	26-45	3	UNI	CLM	1.751	983	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
653	GB	S1A-097	A	M	18-25	1	UNI	ACD	2.791	1503	1	3	0	1	0	5	3.327	0.665	1.996	0.000	0.665
654	GB	S1A-097	B	M	18-25	1	UNI	EDU	2.791	647	1	0	0	0	0	1	1.546	1.546	0.000	0.000	0.000
655	GB	S1A-098	A	F	26-45	3	UNI	ACD	0.000	332	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
657	GB	S1A-098	A	M	26-45	3	UNI	ACD	0.000	143	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
659	GB	S1A-098	A	F	26-45	3	UNI	ACD	0.000	261	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
656	GB	S1A-098	B	F	18-25	1	UNI	CLM	0.000	387	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
658	GB	S1A-098	B	F	26-45	3	UNI	ACD	0.000	263	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
660	GB	S1A-098	B	F	26-45	3	UNI	CLM	0.000	463	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
661	GB	S1A-099	A	F	18-25	1	UNI	SLM	0.000	678	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
663	GB	S1A-099	A	F	18-25	1	UNI	CLM	0.000	468	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
662	GB	S1A-099	B	M	18-25	1	UNI	ACD	0.000	597	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
664	GB	S1A-099	B	M	18-25	1	UNI	ACD	0.000	427	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
665	GB	S1A-100	A	F	46+	4	UNI	CLM	0.000	108	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
669	GB	S1A-100	A	M	26-45	3	UNI	CLM	0.000	672	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
670	GB	S1A-100	B	M	18-25	1	UNI	ACD	0.000	857	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

Table 141: Final data set of IrE.

ID	VAR	FILE	SPEAKER	SEX	age.orig	AGE	REG	REL	DATE	EDU	OCC	L1	PAI	words	ini.n	med.n	fin.m	non.n	na.n	all.n	all	ini	med	fin	non
671	IRE	S1A-001	A	M	34-49	4	N	P	1	UNI	ACD	ENG	0.484	765	0	0	0	1	0	1	1.307	0.000	0.000	0.000	1.307
672	IRE	S1A-001	B	F	34-49	4	N	P	1	SEC	CLM	ENG	0.484	1302	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
673	IRE	S1A-002	A	F	26-33	2	N	C	3	UNI	ACD	ENG	1.052	396	0	2	0	0	0	2	5.051	0.000	5.051	0.000	0.000
674	IRE	S1A-002	C	M	50+	4	N	C	3	SEC	CLM	ENG	1.052	201	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
675	IRE	S1A-002	D	F	50+	4	N	C	3	UNI	ACD	ENG	1.052	470	0	0	1	0	0	1	2.128	0.000	0.000	2.128	0.000
676	IRE	S1A-002	E	M	34-49	4	NA	C	3	UNI	ACD	ENG	1.052	641	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
677	IRE	S1A-002	F	F	26-33	2	N	C	3	UNI	ACD	ENG	1.052	310	0	0	1	0	0	1	3.226	0.000	0.000	3.226	0.000
678	IRE	S1A-003	A	M	34-49	4	NA	C	3	UNI	ACD	ENG	3.407	539	1	0	0	0	0	1	1.855	1.855	0.000	0.000	0.000
679	IRE	S1A-003	B	F	26-33	2	N	C	3	UNI	ACD	ENG	3.407	447	0	1	0	0	0	1	2.237	0.000	2.237	0.000	0.000
680	IRE	S1A-003	C	F	0-25	1	N	C	3	UNI	EDU	ENG	3.407	484	0	0	0	1	0	1	2.066	0.000	0.000	0.000	2.066
681	IRE	S1A-003	D	F	50+	4	N	C	3	UNI	ACD	ENG	3.407	279	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
682	IRE	S1A-003	E	F	26-33	2	N	C	3	UNI	ACD	ENG	3.407	403	1	0	1	0	0	2	4.963	2.481	0.000	2.481	0.000
683	IRE	S1A-003	F	M	0-25	1	NA	C	3	UNI	ACD	ENG	3.407	184	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
684	IRE	S1A-004	A	F	50+	4	N	C	3	SEC	CLM	ENG+	2.260	1284	2	0	0	0	0	2	1.558	1.558	0.000	0.000	0.000
685	IRE	S1A-004	B	F	50+	4	N	C	3	SEC	CLM	ENG	2.260	530	1	0	1	0	0	2	3.774	1.887	0.000	1.887	0.000
686	IRE	S1A-004	C	F	50+	4	N	C	3	SEC	CLM	ENG	2.260	355	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
687	IRE	S1A-005	A	M	34-49	4	N	P	3	SEC	SLM	ENG	12.980	690	2	4	4	0	1	11	15.942	2.899	5.797	5.797	0.000
688	IRE	S1A-005	B	M	26-33	2	N	P	3	UNI	CLM	ENG	12.980	1082	3	3	4	2	0	12	11.091	2.773	2.773	3.697	1.848
690	IRE	S1A-006	A	F	26-33	2	N	C	3	UNI	ACD	ENG	9.091	1053	1	0	0	2	0	3	2.849	0.950	0.000	0.000	1.899
691	IRE	S1A-006	B	F	34-49	4	NA	C	3	UNI	ACD	ENG	9.091	771	1	1	0	0	0	2	2.594	1.297	1.297	0.000	0.000
692	IRE	S1A-006	C	F	34-49	4	N	C	3	UNI	ACD	ENG	9.091	280	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
693	IRE	S1A-006	D	F	26-33	2	N	C	3	UNI	ACD	ENG	9.091	195	0	2	1	0	0	3	15.385	0.000	10.256	5.128	0.000
694	IRE	S1A-006	E	F	26-33	2	N	C	3	UNI	ACD	ENG	9.091	135	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

695	IRE	S1A-007	A	F	34-49	4	NA	C	3	UNI	ACD	ENG	2.481	644	0	2	0	0	0	2	3.106	0.000	3.106	0.000	0.000
696	IRE	S1A-007	B	F	34-49	4	N	C	3	UNI	ACD	ENG	2.481	946	0	2	0	1	0	3	3.171	0.000	2.114	0.000	1.057
697	IRE	S1A-007	C	F	34-49	4	N	C	3	UNI	ACD	ENG	2.481	142	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
698	IRE	S1A-007	D	F	26-33	2	N	C	3	UNI	ACD	ENG	2.481	211	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
699	IRE	S1A-007	E	F	26-33	2	N	C	3	UNI	ACD	ENG	2.481	235	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
700	IRE	S1A-007	F	F	26-33	2	N	C	3	UNI	ACD	ENG	2.481	168	0	0	0	1	0	1	5.952	0.000	0.000	0.000	5.952
701	IRE	S1A-008	A	F	50+	4	N	C	3	SEC	CLM	ENG	0.858	309	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
702	IRE	S1A-008	B	F	50+	4	N	C	3	SEC	NA	ENG	0.858	459	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
703	IRE	S1A-008	C	F	50+	4	N	C	3	SEC	CLM	ENG+	0.858	756	0	1	0	0	0	1	1.323	0.000	1.323	0.000	0.000
704	IRE	S1A-008	D	F	50+	4	N	C	3	SEC	CLM	ENG	0.858	409	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
705	IRE	S1A-009	A	F	50+	4	N	C	3	SEC	CLM	ENG	1.259	966	1	0	0	0	0	1	1.035	1.035	0.000	0.000	0.000
706	IRE	S1A-009	B	F	50+	4	N	C	3	UNI	ACD	ENG	1.259	623	0	0	1	0	0	1	1.605	0.000	0.000	1.605	0.000
707	IRE	S1A-010	A	F	26-33	2	NA	P	3	SEC	CLM	ENG	1.468	384	0	0	1	0	0	1	2.604	0.000	0.000	2.604	0.000
708	IRE	S1A-010	B	F	26-33	2	N	C	3	UNI	ACD	ENG	1.468	512	2	2	0	0	1	5	9.766	3.906	3.906	0.000	0.000
709	IRE	S1A-010	C	F	26-33	2	NA	C	3	UNI	CLM	ENG	1.468	531	0	1	0	0	0	1	1.883	0.000	1.883	0.000	0.000
710	IRE	S1A-010	D	F	26-33	2	N	C	3	UNI	ACD	ENG	1.468	150	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
711	IRE	S1A-011	A	F	26-33	2	N	P	3	UNI	CLM	ENG	20.496	493	4	5	0	0	0	9	18.256	8.114	10.142	0.000	0.000
712	IRE	S1A-011	B	F	NA	NA	S	C	3	NA	NA	NA	20.496	1361	6	10	7	6	0	29	21.308	4.409	7.348	5.143	4.409
713	IRE	S1A-012	A	F	NA	NA	N	NA	1	NA	NA	NA	5.786	1750	7	5	0	0	0	12	6.857	4.000	2.857	0.000	0.000
714	IRE	S1A-012	B	F	NA	NA	N	NA	1	NA	NA	NA	5.786	324	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
715	IRE	S1A-013	A	F	NA	NA	N	NA	1	NA	NA	NA	11.598	1114	0	3	1	0	0	4	3.591	0.000	2.693	0.898	0.000
716	IRE	S1A-013	B	F	NA	NA	N	NA	1	NA	NA	NA	11.598	540	4	3	1	1	0	9	16.667	7.407	5.556	1.852	1.852
717	IRE	S1A-013	C	M	NA	NA	N	NA	1	NA	NA	NA	11.598	236	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
718	IRE	S1A-014	A	F	0-25	1	N	P	1	UNI	ACD	ENG	15.548	156	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
719	IRE	S1A-014	B	F	0-25	1	N	NA	1	UNI	CLM	ENG	15.548	826	2	8	2	0	0	12	14.528	2.421	9.685	2.421	0.000

721	IRE	S1A-014	D	M	0-25	1	N	NA	1	SEC	NA	NA	15.548	589	3	1	4	2	0	10	16.978	5.093	1.698	6.791	3.396
722	IRE	S1A-015	A	M	0-25	1	N	NA	1	UNI	NA	NA	0.000	805	0	1	6	0	0	7	8.696	0.000	1.242	7.453	0.000
723	IRE	S1A-015	B	F	0-25	1	N	NA	1	UNI	NA	NA	0.000	462	0	1	0	0	0	1	2.165	0.000	2.165	0.000	0.000
724	IRE	S1A-015	C	F	0-25	1	N	NA	1	UNI	NA	NA	0.000	196	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
725	IRE	S1A-015	D	M	0-25	1	N	NA	1	UNI	NA	NA	0.000	765	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
726	IRE	S1A-016	A	F	0-25	1	NA	NA	1	UNI	EDU	NA	0.000	1023	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
727	IRE	S1A-016	B	M	0-25	1	N	NA	1	UNI	EDU	NA	0.000	470	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
728	IRE	S1A-016	C	M	NA	NA	N	NA	1	NA	CLM	NA	0.000	572	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
729	IRE	S1A-017	A	F	0-25	1	N	NA	1	UNI	EDU	NA	4.710	1043	1	2	2	0	0	5	4.794	0.959	1.918	1.918	0.000
730	IRE	S1A-017	B	F	0-25	1	N	NA	1	UNI	EDU	NA	4.710	224	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
731	IRE	S1A-017	C	M	0-25	1	N	NA	1	UNI	EDU	NA	4.710	325	0	1	0	0	0	1	3.077	0.000	3.077	0.000	0.000
732	IRE	S1A-017	D	M	0-25	1	N	NA	1	UNI	EDU	NA	4.710	312	0	0	2	0	0	2	6.410	0.000	0.000	6.410	0.000
733	IRE	S1A-018	A	M	0-25	1	N	NA	1	UNI	EDU	NA	1.832	925	1	0	1	1	0	3	3.243	1.081	0.000	1.081	1.081
734	IRE	S1A-018	B	M	0-25	1	N	NA	1	UNI	EDU	NA	1.832	822	4	0	0	0	0	4	4.866	4.866	0.000	0.000	0.000
735	IRE	S1A-018	C	F	0-25	1	NA	NA	1	UNI	EDU	NA	1.832	278	1	0	0	0	0	1	3.597	3.597	0.000	0.000	0.000
736	IRE	S1A-018	D	F	0-25	1	N	C	1	UNI	EDU	NA	1.832	268	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
737	IRE	S1A-019	A	F	26-33	2	N	C	3	UNI	ACD	ENG	5.284	752	4	0	0	0	0	4	5.319	5.319	0.000	0.000	0.000
738	IRE	S1A-019	B	F	26-33	2	N	P	3	UNI	ACD	ENG	5.284	475	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
739	IRE	S1A-019	C	F	34-49	4	N	C	3	SEC	CLM	ENG	5.284	445	0	0	2	0	0	2	4.494	0.000	0.000	4.494	0.000
740	IRE	S1A-019	D	F	26-33	2	N	C	3	UNI	ACD	ENG	5.284	312	1	1	0	0	0	2	6.410	3.205	3.205	0.000	0.000
741	IRE	S1A-020	A	F	26-33	2	N	C	3	UNI	ACD	ENG	3.125	609	2	0	2	0	0	4	6.568	3.284	0.000	3.284	0.000
742	IRE	S1A-020	B	F	26-33	2	N	P	3	UNI	ACD	ENG	3.125	294	0	1	0	1	0	2	6.803	0.000	3.401	0.000	3.401
743	IRE	S1A-020	C	F	34-49	4	N	C	3	SEC	CLM	ENG	3.125	331	1	0	0	0	0	1	3.021	3.021	0.000	0.000	0.000
744	IRE	S1A-020	D	F	26-33	2	N	C	3	UNI	ACD	ENG	3.125	949	0	2	0	1	0	3	3.161	0.000	2.107	0.000	1.054
745	IRE	S1A-021	A	F	26-33	2	N	C	3	UNI	ACD	ENG	1.626	430	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

746	IRE	S1A-021	D	M	50+	4	N	C	3	UNI	ACD	ENG	1.626	449	0	0	1	0	0	1	2.227	0.000	0.000	2.227	0.000
747	IRE	S1A-021	F	F	26-33	2	N	C	3	UNI	ACD	ENG	1.626	403	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
748	IRE	S1A-021	H	F	26-33	2	N	P	3	UNI	ACD	ENG	1.626	212	1	0	0	0	0	1	4.717	4.717	0.000	0.000	0.000
749	IRE	S1A-022	A	F	0-25	1	N	C	1	UNI	EDU	ENG	0.000	448	0	1	0	0	0	1	2.232	0.000	2.232	0.000	0.000
750	IRE	S1A-022	B	F	0-25	1	NA	C	1	UNI	EDU	ENG	0.000	542	0	1	2	0	0	3	5.535	0.000	1.845	3.690	0.000
751	IRE	S1A-022	C	F	34-49	4	N	C	1	SEC	SLM	ENG	0.000	324	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
752	IRE	S1A-022	D	M	50+	4	N	C	1	SEC	SLM	ENG	0.000	328	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
753	IRE	S1A-022	E	M	0-25	1	NA	C	1	SEC	EDU	ENG	0.000	245	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
754	IRE	S1A-023	A	F	0-25	1	N	C	1	UNI	EDU	ENG	0.000	357	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
755	IRE	S1A-023	B	F	0-25	1	NA	C	1	UNI	EDU	ENG	0.000	454	0	0	3	0	0	3	6.608	0.000	0.000	6.608	0.000
756	IRE	S1A-023	C	F	34-49	4	N	C	1	SEC	SLM	ENG	0.000	237	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
757	IRE	S1A-023	D	M	50+	4	N	C	1	SEC	SLM	ENG	0.000	371	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
758	IRE	S1A-023	E	M	0-25	1	NA	C	1	SEC	EDU	ENG	0.000	130	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
759	IRE	S1A-023	F	F	0-25	1	NA	C	1	SEC	EDU	ENG	0.000	272	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
760	IRE	S1A-024	A	F	NA	NA	N	NA	1	NA	NA	NA	5.319	673	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
761	IRE	S1A-024	B	F	NA	NA	N	NA	1	NA	NA	NA	5.319	128	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
762	IRE	S1A-024	C	M	NA	NA	N	NA	1	NA	NA	NA	5.319	681	1	0	4	1	0	6	8.811	1.468	0.000	5.874	1.468
763	IRE	S1A-024	D	M	NA	NA	N	NA	1	NA	NA	NA	5.319	447	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
764	IRE	S1A-025	A	F	NA	NA	N	NA	1	NA	NA	NA	0.000	778	0	1	0	0	0	1	1.285	0.000	1.285	0.000	0.000
765	IRE	S1A-025	B	F	NA	NA	N	NA	1	NA	NA	NA	0.000	597	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
766	IRE	S1A-025	C	M	NA	NA	N	NA	1	NA	NA	NA	0.000	275	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
767	IRE	S1A-025	D	M	NA	NA	N	NA	1	NA	NA	NA	0.000	195	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
768	IRE	S1A-026	A	F	26-33	2	N	NA	1	NA	ACD	NA	8.753	1495	7	2	3	1	0	13	8.696	4.682	1.338	2.007	0.669
769	IRE	S1A-026	B	F	NA	NA	N	NA	1	NA	ACD	NA	8.753	306	1	1	1	0	0	3	9.804	3.268	3.268	3.268	0.000
770	IRE	S1A-026	C	M	NA	NA	N	NA	1	NA	NA	NA	8.753	151	0	0	1	0	0	1	6.623	0.000	0.000	6.623	0.000

774	IRE	S1A-028	A	F	0-25	1	N	NA	1	UNI	EDU	NA	3.370	893	1	0	0	0	0	1	1.120	1.120	0.000	0.000	0.000
775	IRE	S1A-028	B	F	0-25	1	N	NA	1	UNI	EDU	NA	3.370	294	2	0	0	1	0	3	10.204	6.803	0.000	0.000	3.401
777	IRE	S1A-029	A	M	0-25	1	N	P	1	UNI	ACD	ENG	0.000	421	0	3	0	1	0	4	9.501	0.000	7.126	0.000	2.375
778	IRE	S1A-029	B	F	50+	4	N	NA	1	NA	NA	NA	0.000	1063	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
779	IRE	S1A-029	D	F	50+	4	N	NA	1	NA	NA	NA	0.000	534	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
780	IRE	S1A-030	A	M	NA	NA	N	NA	1	NA	NA	NA	2.618	492	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
781	IRE	S1A-030	B	F	NA	NA	N	NA	1	NA	NA	NA	2.618	188	0	1	0	0	0	1	5.319	0.000	5.319	0.000	0.000
782	IRE	S1A-030	C	F	NA	NA	N	NA	1	UNI	EDU	ENG	2.618	626	0	1	0	2	0	3	4.792	0.000	1.597	0.000	3.195
783	IRE	S1A-030	D	M	NA	NA	N	NA	1	NA	NA	NA	2.618	520	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
784	IRE	S1A-031	A	F	NA	NA	N	NA	1	NA	NA	NA	5.355	788	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
785	IRE	S1A-031	B	F	NA	NA	N	NA	1	NA	NA	NA	5.355	572	0	1	0	0	0	1	1.748	0.000	1.748	0.000	0.000
786	IRE	S1A-031	C	M	NA	NA	N	NA	1	NA	NA	NA	5.355	368	1	0	0	0	0	1	2.717	2.717	0.000	0.000	0.000
787	IRE	S1A-031	D	F	NA	NA	N	NA	1	NA	NA	NA	5.355	379	2	0	1	0	0	3	7.916	5.277	0.000	2.639	0.000
789	IRE	S1A-032	B	M	0-25	1	N	NA	1	UNI	EDU	NA	2.132	132	0	0	1	0	0	1	7.576	0.000	0.000	7.576	0.000
791	IRE	S1A-032	D	F	NA	NA	N	NA	1	NA	NA	NA	2.132	592	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
792	IRE	S1A-032	E	F	NA	NA	N	NA	1	NA	NA	NA	2.132	346	0	1	0	0	0	1	2.890	0.000	2.890	0.000	0.000
793	IRE	S1A-032	G	M	NA	NA	N	NA	1	NA	NA	NA	2.132	123	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
794	IRE	S1A-033	A	F	50+	4	N	C	3	SEC	CLM	ENG	1.239	162	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
795	IRE	S1A-033	B	F	50+	4	N	C	3	SEC	CLM	ENG	1.239	1268	0	0	1	0	0	1	0.789	0.000	0.000	0.789	0.000
796	IRE	S1A-033	C	F	50+	4	N	C	3	SEC	SLM	ENG	1.239	422	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
797	IRE	S1A-033	D	F	50+	4	NA	C	3	SEC	CLM	ENG+	1.239	385	1	0	0	0	0	1	2.597	2.597	0.000	0.000	0.000
798	IRE	S1A-034	A	M	NA	NA	N	NA	1	NA	NA	NA	3.876	627	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
799	IRE	S1A-034	B	M	NA	NA	N	NA	1	NA	NA	NA	3.876	470	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
800	IRE	S1A-034	C	F	NA	NA	N	NA	1	UNI	EDU	ENG	3.876	536	0	2	0	1	0	3	5.597	0.000	3.731	0.000	1.866
801	IRE	S1A-034	D	F	NA	NA	N	NA	1	NA	NA	NA	3.876	238	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

802	IRE	S1A-035	A	M	0-25	1	N	C	3	UNI	EDU	ENG	7.576	946	0	1	1	1	0	3	3.171	0.000	1.057	1.057	1.057
803	IRE	S1A-035	B	F	0-25	1	N	C	3	UNI	CLM	ENG	7.576	1166	3	10	0	0	0	13	11.149	2.573	8.576	0.000	0.000
805	IRE	S1A-036	B	F	26-33	2	N	C	3	UNI	CLM	ENG	7.183	1253	5	2	2	0	0	9	7.183	3.990	1.596	1.596	0.000
806	IRE	S1A-037	A	F	50+	4	N	C	3	SEC	CLM	ENG	2.392	206	0	0	1	0	0	1	4.854	0.000	0.000	4.854	0.000
807	IRE	S1A-037	B	F	50+	4	N	C	3	SEC	CLM	ENG	2.392	817	0	1	0	0	0	1	1.224	0.000	1.224	0.000	0.000
808	IRE	S1A-037	C	F	50+	4	N	C	3	SEC	SLM	ENG	2.392	906	1	1	0	0	0	2	2.208	1.104	1.104	0.000	0.000
809	IRE	S1A-037	D	F	50+	4	NA	C	3	SEC	CLM	ENG+	2.392	348	0	0	0	1	0	1	2.874	0.000	0.000	0.000	2.874
811	IRE	S1A-038	B	F	26-33	2	N	C	3	UNI	CLM	ENG	11.640	945	5	6	0	0	0	11	11.640	5.291	6.349	0.000	0.000
812	IRE	S1A-039	A	M	26-33	2	N	C	3	UNI	EDU	ENG	14.356	230	0	1	0	0	0	1	4.348	0.000	4.348	0.000	0.000
813	IRE	S1A-039	B	M	0-25	1	N	C	3	UNI	EDU	ENG	14.356	1122	4	4	4	5	0	17	15.152	3.565	3.565	3.565	4.456
814	IRE	S1A-039	C	F	0-25	1	N	C	3	UNI	EDU	ENG	14.356	898	7	2	1	2	0	12	13.363	7.795	2.227	1.114	2.227
815	IRE	S1A-040	A	F	50+	4	N	NA	3	NA	NA	NA	2.985	282	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
816	IRE	S1A-040	B	M	50+	4	N	C	3	NA	NA	NA	2.985	400	0	0	2	0	0	2	5.000	0.000	0.000	5.000	0.000
817	IRE	S1A-040	D	F	0-25	1	N	C	3	UNI	EDU	ENG	2.985	270	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
818	IRE	S1A-041	A	F	NA	NA	N	NA	1	NA	NA	NA	5.460	229	0	0	1	0	0	1	4.367	0.000	0.000	4.367	0.000
819	IRE	S1A-041	B	M	0-25	1	N	NA	1	NA	NA	NA	5.460	791	2	5	6	3	0	16	20.228	2.528	6.321	7.585	3.793
820	IRE	S1A-041	C	F	NA	NA	N	NA	1	NA	NA	NA	5.460	305	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
821	IRE	S1A-041	D	M	NA	NA	N	NA	1	NA	NA	NA	5.460	794	0	0	6	0	0	6	7.557	0.000	0.000	7.557	0.000
822	IRE	S1A-042	A	F	NA	NA	N	NA	1	UNI	EDU	NA	1.529	215	0	0	0	1	0	1	4.651	0.000	0.000	0.000	4.651
823	IRE	S1A-042	B	F	NA	NA	N	NA	1	NA	NA	NA	1.529	708	0	1	1	0	0	2	2.825	0.000	1.412	1.412	0.000
824	IRE	S1A-042	D	F	NA	NA	S	NA	1	NA	NA	NA	1.529	600	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
825	IRE	S1A-043	A	F	50+	4	N	C	3	SEC	SLM	ENG	2.148	1590	1	0	1	1	0	3	1.887	0.629	0.000	0.629	0.629
826	IRE	S1A-043	B	M	26-33	2	N	C	3	UNI	CLM	ENG	2.148	272	0	0	1	0	0	1	3.676	0.000	0.000	3.676	0.000
827	IRE	S1A-044	A	F	26-33	2	NA	P	3	UNI	ACD	ENG	15.616	1729	5	11	7	4	0	27	15.616	2.892	6.362	4.049	2.313
828	IRE	S1A-045	A	F	NA	NA	N	NA	3	SEC	SLM	NA	3.625	1554	0	1	2	3	0	6	3.861	0.000	0.644	1.287	1.931

829	IRE	S1A-045	B	F	0-25	1	N	C	3	UNI	EDU	ENG	3.625	377	1	0	0	0	0	1	2.653	2.653	0.000	0.000	0.000
830	IRE	S1A-046	A	F	0-25	1	S	C	1	UNI	EDU	ENG	8.573	775	0	0	3	0	0	3	3.871	0.000	0.000	3.871	0.000
831	IRE	S1A-046	B	F	0-25	1	S	C	1	UNI	EDU	ENG	8.573	858	3	0	6	2	0	11	12.821	3.497	0.000	6.993	2.331
832	IRE	S1A-047	A	F	34-49	4	S	C	1	UNI	ACD	ENG	0.000	1162	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
833	IRE	S1A-047	B	F	50+	4	S	C	1	UNI	ACD	ENG	0.000	623	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
834	IRE	S1A-048	A	F	0-25	1	S	NA	1	SEC	EDU	ENG	12.115	827	1	1	6	1	0	9	10.883	1.209	1.209	7.255	1.209
835	IRE	S1A-048	B	M	0-25	1	S	NA	1	SEC	EDU	ENG	12.115	989	2	2	7	2	0	13	13.145	2.022	2.022	7.078	2.022
836	IRE	S1A-049	A	F	0-25	1	S	C	1	UNI	ACD	ENG	4.443	2031	2	6	1	0	0	9	4.431	0.985	2.954	0.492	0.000
837	IRE	S1A-049	B	F	0-25	1	S	C	1	UNI	EDU	ENG	4.443	445	0	0	2	0	0	2	4.494	0.000	0.000	4.494	0.000
838	IRE	S1A-050	A	F	0-25	1	S	C	1	UNI	EDU	ENG	3.922	269	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
839	IRE	S1A-050	B	F	0-25	1	S	C	1	UNI	EDU	ENG	3.922	547	2	2	1	1	0	6	10.969	3.656	3.656	1.828	1.828
840	IRE	S1A-050	C	F	0-25	1	S	C	1	UNI	EDU	ENG	3.922	1238	0	0	0	1	0	1	0.808	0.000	0.000	0.000	0.808
841	IRE	S1A-051	A	F	0-25	1	NA	C	1	UNI	EDU	ENG+	6.381	631	0	0	0	1	0	1	1.585	0.000	0.000	0.000	1.585
842	IRE	S1A-051	B	F	0-25	1	S	C	1	UNI	EDU	ENG+	6.381	710	0	1	3	0	0	4	5.634	0.000	1.408	4.225	0.000
843	IRE	S1A-051	D	F	0-25	1	S	C	1	UNI	EDU	ENG+	6.381	387	0	0	3	0	0	3	7.752	0.000	0.000	7.752	0.000
845	IRE	S1A-052	A	F	0-25	1	S	C	1	UNI	EDU	ENG	0.000	643	0	0	1	0	1	2	3.110	0.000	0.000	1.555	0.000
846	IRE	S1A-052	B	M	0-25	1	S	C	1	UNI	EDU	ENG	0.000	626	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
847	IRE	S1A-052	C	M	26-33	2	S	C	1	UNI	ACD	ENG	0.000	583	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
848	IRE	S1A-053	A	F	26-33	2	S	C	1	UNI	EDU	ENG	5.714	952	2	0	0	0	0	2	2.101	2.101	0.000	0.000	0.000
849	IRE	S1A-053	B	F	0-25	1	S	C	1	UNI	EDU	ENG	5.714	182	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
850	IRE	S1A-053	C	F	0-25	1	S	C	1	UNI	EDU	ENG	5.714	417	0	0	1	0	0	1	2.398	0.000	0.000	2.398	0.000
851	IRE	S1A-053	D	F	0-25	1	S	C	1	UNI	EDU	ENG	5.714	458	0	2	2	0	0	4	8.734	0.000	4.367	4.367	0.000
852	IRE	S1A-054	A	F	26-33	2	S	C	1	UNI	EDU	ENG	2.860	770	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
853	IRE	S1A-054	C	F	0-25	1	S	C	1	UNI	EDU	ENG	2.860	412	1	0	0	0	0	1	2.427	2.427	0.000	0.000	0.000
854	IRE	S1A-054	D	F	0-25	1	S	C	1	UNI	EDU	ENG	2.860	637	0	1	1	0	0	2	3.140	0.000	1.570	1.570	0.000

855	IRE	S1A-055	A	M	50+	4	S	C	1	NA	NA	ENG	5.729	573	1	0	5	1	0	7	12.216	1.745	0.000	8.726	1.745
856	IRE	S1A-055	B	F	50+	4	S	C	1	NA	CLM	ENG	5.729	1313	1	2	3	1	1	8	6.093	0.762	1.523	2.285	0.762
857	IRE	S1A-055	D	M	50+	4	S	C	1	SEC	SLM	ENG	5.729	258	0	0	1	0	0	1	3.876	0.000	0.000	3.876	0.000
859	IRE	S1A-056	A	F	0-25	1	S	C	1	UNI	EDU	ENG	16.334	172	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
860	IRE	S1A-056	B	F	0-25	1	S	C	1	UNI	EDU	ENG	16.334	848	3	1	1	1	1	7	8.255	3.538	1.179	1.179	1.179
862	IRE	S1A-056	D	F	0-25	1	S	C	1	SEC	EDU	ENG	16.334	805	6	5	5	4	0	20	24.845	7.453	6.211	6.211	4.969
863	IRE	S1A-057	A	F	0-25	1	S	NA	1	SEC	EDU	ENG	10.363	1021	2	0	5	2	0	9	8.815	1.959	0.000	4.897	1.959
864	IRE	S1A-057	B	M	0-25	1	S	NA	1	SEC	EDU	ENG	10.363	909	1	1	6	3	0	11	12.101	1.100	1.100	6.601	3.300
865	IRE	S1A-058	A	F	0-25	1	S	C	1	UNI	EDU	ENG	8.547	238	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
866	IRE	S1A-058	B	F	0-25	1	S	C	1	UNI	EDU	ENG	8.547	711	0	0	2	0	0	2	2.813	0.000	0.000	2.813	0.000
867	IRE	S1A-058	C	F	0-25	1	S	C	1	UNI	EDU	ENG	8.547	388	1	1	1	0	0	3	7.732	2.577	2.577	2.577	0.000
868	IRE	S1A-058	D	F	0-25	1	S	C	1	SEC	EDU	ENG	8.547	665	1	0	4	1	0	6	9.023	1.504	0.000	6.015	1.504
869	IRE	S1A-059	A	F	0-25	1	S	C	1	UNI	EDU	ENG	2.045	912	0	0	1	0	0	1	1.096	0.000	0.000	1.096	0.000
870	IRE	S1A-059	B	F	50+	4	S	C	1	UNI	NA	ENG	2.045	1044	0	0	0	2	1	3	2.874	0.000	0.000	0.000	1.916
871	IRE	S1A-060	A	F	0-25	1	S	C	1	UNI	EDU	ENG	1.222	796	1	0	1	0	0	2	2.513	1.256	0.000	1.256	0.000
872	IRE	S1A-060	B	F	50+	4	S	C	1	UNI	NA	ENG	1.222	841	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
873	IRE	S1A-061	A	M	34-49	4	S	C	1	UNI	CLM	ENG	0.000	1460	1	3	0	0	0	4	2.740	0.685	2.055	0.000	0.000
874	IRE	S1A-061	B	F	0-25	1	S	C	1	UNI	EDU	ENG	0.000	383	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
875	IRE	S1A-061	C	F	0-25	1	S	P	1	UNI	EDU	ENG	0.000	307	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
876	IRE	S1A-062	A	M	0-25	1	S	NA	1	UNI	ACD	NA	8.480	336	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
877	IRE	S1A-062	B	F	0-25	1	S	NA	1	UNI	EDU	NA	8.480	1248	2	2	4	1	1	10	8.013	1.603	1.603	3.205	0.801
878	IRE	S1A-062	C	M	0-25	1	S	NA	1	UNI	ACD	NA	8.480	285	1	0	1	1	0	3	10.526	3.509	0.000	3.509	3.509
879	IRE	S1A-063	A	F	0-25	1	S	C	1	UNI	EDU	ENG	10.821	870	1	1	3	0	0	5	5.747	1.149	1.149	3.448	0.000
880	IRE	S1A-063	B	F	0-25	1	S	C	1	UNI	EDU	ENG	10.821	780	3	1	1	1	0	6	7.692	3.846	1.282	1.282	1.282
881	IRE	S1A-063	C	F	0-25	1	S	C	1	UNI	EDU	ENG	10.821	329	1	2	3	0	0	6	18.237	3.040	6.079	9.119	0.000

882	IRE	S1A-064	B	F	50+	4	S	C	1	SEC	SLM	ENG	8.929	151	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
884	IRE	S1A-064	D	F	0-25	1	S	C	1	UNI	CLM	ENG	8.929	289	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
885	IRE	S1A-064	E	F	0-25	1	S	C	1	UNI	EDU	ENG	8.929	607	1	1	4	2	0	8	13.180	1.647	1.647	6.590	3.295
886	IRE	S1A-065	A	M	0-25	1	S	NA	1	UNI	ACD	NA	7.613	412	0	0	1	0	0	1	2.427	0.000	0.000	2.427	0.000
887	IRE	S1A-065	B	M	0-25	1	S	NA	1	UNI	ACD	NA	7.613	692	0	1	5	0	0	6	8.671	0.000	1.445	7.225	0.000
888	IRE	S1A-065	C	F	0-25	1	S	NA	1	UNI	EDU	NA	7.613	1147	2	0	4	2	0	8	6.975	1.744	0.000	3.487	1.744
889	IRE	S1A-066	A	F	0-25	1	S	C	1	UNI	EDU	ENG	1.871	400	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
890	IRE	S1A-066	B	F	0-25	1	S	C	1	UNI	EDU	ENG	1.871	584	0	0	1	0	0	1	1.712	0.000	0.000	1.712	0.000
891	IRE	S1A-066	C	F	0-25	1	S	C	1	UNI	EDU	ENG	1.871	1019	1	0	0	1	0	2	1.963	0.981	0.000	0.000	0.981
892	IRE	S1A-067	A	F	0-25	1	S	C	3	UNI	EDU	ENG	0.000	907	0	1	1	1	0	3	3.308	0.000	1.103	1.103	1.103
893	IRE	S1A-067	B	M	34-49	4	S	C	3	UNI	SLM	ENG	0.000	243	0	0	1	0	0	1	4.115	0.000	0.000	4.115	0.000
894	IRE	S1A-067	C	F	50+	4	S	C	3	SEC	SLM	ENG	0.000	499	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
895	IRE	S1A-067	D	M	0-25	1	S	C	3	UNI	EDU	ENG	0.000	367	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
898	IRE	S1A-068	C	F	26-33	2	S	C	3	UNI	EDU	ENG	16.492	667	1	0	6	3	1	11	16.492	1.499	0.000	8.996	4.498
899	IRE	S1A-069	A	M	0-25	1	S	C	1	UNI	EDU	ENG	2.151	668	2	0	0	0	0	2	2.994	2.994	0.000	0.000	0.000
900	IRE	S1A-069	B	F	0-25	1	S	C	1	UNI	EDU	ENG	2.151	638	0	0	1	0	0	1	1.567	0.000	0.000	1.567	0.000
901	IRE	S1A-069	C	F	0-25	1	S	C	1	UNI	EDU	ENG	2.151	757	1	0	0	0	1	2	2.642	1.321	0.000	0.000	0.000
902	IRE	S1A-070	A	F	0-25	1	S	C	1	UNI	CLM	ENG	2.473	1096	1	0	2	0	0	3	2.737	0.912	0.000	1.825	0.000
903	IRE	S1A-070	B	F	0-25	1	S	C	1	UNI	EDU	ENG	2.473	926	0	1	0	1	0	2	2.160	0.000	1.080	0.000	1.080
904	IRE	S1A-071	A	F	0-25	1	S	C	1	UNI	EDU	ENG	2.247	707	2	3	1	0	0	6	8.487	2.829	4.243	1.414	0.000
905	IRE	S1A-071	B	M	0-25	1	S	C	1	UNI	EDU	ENG	2.247	595	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
906	IRE	S1A-071	C	M	26-33	2	S	C	1	UNI	ACD	ENG	2.247	740	1	0	0	2	0	3	4.054	1.351	0.000	0.000	2.703
907	IRE	S1A-072	A	F	0-25	1	S	C	1	SEC	CLM	ENG	4.444	563	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
908	IRE	S1A-072	B	M	0-25	1	S	C	1	SEC	EDU	ENG	4.444	1209	1	1	3	0	0	5	4.136	0.827	0.827	2.481	0.000
909	IRE	S1A-072	C	F	0-25	1	S	C	1	UNI	EDU	ENG	4.444	366	0	0	2	0	0	2	5.464	0.000	0.000	5.464	0.000

910	IRE	S1A-073	A	F	50+	4	S	C	1	UNI	CLM	ENG+	1.256	871	0	0	1	0	0	1	1.148	0.000	0.000	1.148	0.000
911	IRE	S1A-073	B	M	50+	4	S	C	1	UNI	ACD	ENG+	1.256	1288	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
912	IRE	S1A-073	C	F	0-25	1	S	C	1	UNI	EDU	ENG+	1.256	304	1	0	1	0	0	2	6.579	3.289	0.000	3.289	0.000
913	IRE	S1A-074	A	F	50+	4	S	C	1	UNI	CLM	ENG+	0.000	1029	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
914	IRE	S1A-074	B	M	50+	4	S	C	1	UNI	ACD	ENG+	0.000	767	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
915	IRE	S1A-074	C	F	0-25	1	S	C	1	UNI	EDU	ENG+	0.000	655	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
916	IRE	S1A-075	A	F	0-25	1	S	NA	1	SEC	EDU	ENG	4.587	1724	11	5	5	3	0	24	13.921	6.381	2.900	2.900	1.740
917	IRE	S1A-075	B	F	0-25	1	S	C	1	UNI	EDU	ENG	4.587	290	0	0	2	0	0	2	6.897	0.000	0.000	6.897	0.000
918	IRE	S1A-075	C	F	0-25	1	S	C	1	UNI	EDU	ENG	4.587	364	0	0	1	0	0	1	2.747	0.000	0.000	2.747	0.000
919	IRE	S1A-076	A	F	0-25	1	S	C	1	UNI	ACD	ENG	0.000	1710	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
920	IRE	S1A-076	B	F	0-25	1	S	C	1	UNI	EDU	ENG	0.000	557	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
921	IRE	S1A-077	A	F	50+	4	S	C	1	SEC	CLM	ENG	2.611	596	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
922	IRE	S1A-077	B	F	50+	4	S	NA	1	UNI	CLM	ENG	2.611	506	0	1	0	0	0	1	1.976	0.000	1.976	0.000	0.000
923	IRE	S1A-077	C	F	34-49	4	S	C	1	UNI	ACD	ENG	2.611	617	1	0	0	0	0	1	1.621	1.621	0.000	0.000	0.000
924	IRE	S1A-077	D	F	0-25	1	S	C	1	SEC	EDU	ENG	2.611	532	1	0	1	0	0	2	3.759	1.880	0.000	1.880	0.000
925	IRE	S1A-078	A	F	34-49	4	S	C	1	SEC	SLM	ENG	0.000	593	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
926	IRE	S1A-078	B	F	0-25	1	S	C	1	UNI	SLM	ENG	0.000	880	2	0	3	0	0	5	5.682	2.273	0.000	3.409	0.000
927	IRE	S1A-078	C	F	0-25	1	S	C	1	UNI	EDU	ENG	0.000	306	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
928	IRE	S1A-078	D	M	50+	4	S	C	1	SEC	SLM	ENG	0.000	105	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
929	IRE	S1A-079	A	F	0-25	1	S	C	1	UNI	EDU	ENG	7.356	750	2	5	1	1	0	9	12.000	2.667	6.667	1.333	1.333
930	IRE	S1A-079	B	F	26-33	2	S	C	1	UNI	EDU	ENG	7.356	1425	5	0	2	0	0	7	4.912	3.509	0.000	1.404	0.000
932	IRE	S1A-080	A	F	0-25	1	NA	C	1	UNI	EDU	ENG+	2.481	1100	2	0	3	3	1	9	8.182	1.818	0.000	2.727	2.727
933	IRE	S1A-080	B	F	0-25	1	S	C	1	UNI	EDU	ENG	2.481	402	1	1	0	0	0	2	4.975	2.488	2.488	0.000	0.000
934	IRE	S1A-080	C	F	0-25	1	S	NA	1	UNI	EDU	ENG	2.481	404	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
935	IRE	S1A-081	A	M	50+	4	S	C	1	UNI	SLM	ENG	0.000	346	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

936	IRE	S1A-081	B	F	50+	4	S	C	1	SEC	SLM	ENG	0.000	342	0	1	1	0	0	2	5.848	0.000	2.924	2.924	0.000
937	IRE	S1A-081	C	F	0-25	1	S	C	1	UNI	CLM	ENG	0.000	258	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
938	IRE	S1A-081	D	F	0-25	1	S	C	1	UNI	CLM	ENG	0.000	137	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
939	IRE	S1A-081	E	F	0-25	1	S	C	1	UNI	EDU	ENG	0.000	275	1	0	1	1	0	3	10.909	3.636	0.000	3.636	3.636
940	IRE	S1A-081	F	M	0-25	1	NA	C	1	SEC	EDU	ENG	0.000	225	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
941	IRE	S1A-081	G	F	0-25	1	NA	C	1	SEC	NA	ENG	0.000	219	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
942	IRE	S1A-081	H	M	0-25	1	S	C	1	UNI	CLM	ENG	0.000	125	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
943	IRE	S1A-082	A	F	26-33	2	S	C	1	UNI	SLM	ENG	1.921	578	1	0	0	1	0	2	3.460	1.730	0.000	0.000	1.730
944	IRE	S1A-082	B	M	26-33	2	S	C	1	UNI	CLM	ENG	1.921	123	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
945	IRE	S1A-082	C	F	26-33	2	S	C	1	UNI	EDU	ENG	1.921	292	0	1	0	0	0	1	3.425	0.000	3.425	0.000	0.000
946	IRE	S1A-082	E	F	26-33	2	S	C	1	UNI	EDU	ENG	1.921	546	0	1	1	0	0	2	3.663	0.000	1.832	1.832	0.000
947	IRE	S1A-082	G	F	26-33	2	S	C	1	UNI	SLM	ENG	1.921	495	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
948	IRE	S1A-083	A	M	50+	4	S	C	1	NA	NA	ENG	6.861	416	0	0	2	0	0	2	4.808	0.000	0.000	4.808	0.000
949	IRE	S1A-083	B	F	50+	4	S	C	1	NA	CLM	ENG	6.861	939	3	0	0	1	1	5	5.325	3.195	0.000	0.000	1.065
950	IRE	S1A-083	D	M	50+	4	S	C	1	SEC	SLM	ENG	6.861	268	0	0	1	0	0	1	3.731	0.000	0.000	3.731	0.000
951	IRE	S1A-083	E	F	34-49	4	S	C	1	UNI	ACD	ENG	6.861	315	0	0	2	1	0	3	9.524	0.000	0.000	6.349	3.175
952	IRE	S1A-084	A	F	0-25	1	S	C	1	UNI	EDU	ENG	1.575	244	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
953	IRE	S1A-084	B	F	0-25	1	S	C	1	UNI	EDU	ENG	1.575	850	1	1	0	0	0	2	2.353	1.176	1.176	0.000	0.000
955	IRE	S1A-084	D	F	0-25	1	S	C	1	UNI	EDU	ENG	1.575	420	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
956	IRE	S1A-085	A	F	0-25	1	S	NA	1	UNI	EDU	ENG	0.000	1012	8	9	3	2	0	22	21.739	7.905	8.893	2.964	1.976
957	IRE	S1A-085	B	F	0-25	1	S	NA	1	UNI	EDU	ENG	0.000	515	2	0	0	0	0	2	3.883	3.883	0.000	0.000	0.000
958	IRE	S1A-085	C	M	0-25	1	S	NA	1	SEC	CLM	ENG	0.000	275	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
959	IRE	S1A-085	E	F	0-25	1	S	C	1	UNI	EDU	ENG	0.000	240	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
960	IRE	S1A-086	A	F	0-25	1	S	C	1	UNI	CLM	ENG	0.000	1549	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
961	IRE	S1A-086	B	F	0-25	1	S	C	1	UNI	EDU	ENG	0.000	878	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

962	IRE	S1A-087	A	F	0-25	1	S	C	1	SEC	CLM	ENG	3.584	590	1	0	0	0	0	1	1.695	1.695	0.000	0.000	0.000
963	IRE	S1A-087	B	M	0-25	1	S	C	1	SEC	EDU	ENG	3.584	959	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
964	IRE	S1A-087	C	F	0-25	1	S	C	1	UNI	EDU	ENG	3.584	436	2	0	3	0	0	5	11.468	4.587	0.000	6.881	0.000
965	IRE	S1A-088	A	F	0-25	1	S	C	1	UNI	EDU	ENG	10.954	146	0	1	0	0	0	1	6.849	0.000	6.849	0.000	0.000
966	IRE	S1A-088	B	F	0-25	1	S	C	1	UNI	SLM	ENG	10.954	281	1	1	0	0	0	2	7.117	3.559	3.559	0.000	0.000
967	IRE	S1A-088	C	F	0-25	1	S	P	1	UNI	ACD	ENG	10.954	140	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
968	IRE	S1A-088	D	F	0-25	1	S	P	1	UNI	SLM	ENG	10.954	1126	6	1	5	3	0	15	13.321	5.329	0.888	4.440	2.664
969	IRE	S1A-088	E	F	0-25	1	S	P	1	UNI	EDU	ENG	10.954	426	0	0	2	0	0	2	4.695	0.000	0.000	4.695	0.000
970	IRE	S1A-089	A	M	NA	NA	S	NA	1	NA	NA	NA	3.487	210	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
971	IRE	S1A-089	C	F	NA	NA	S	NA	1	UNI	NA	NA	3.487	654	1	2	6	0	0	9	13.761	1.529	3.058	9.174	0.000
972	IRE	S1A-089	D	F	NA	NA	S	NA	1	UNI	NA	NA	3.487	910	3	1	0	0	0	4	4.396	3.297	1.099	0.000	0.000
973	IRE	S1A-089	E	F	NA	NA	S	NA	1	NA	NA	NA	3.487	237	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
975	IRE	S1A-090	C	F	NA	NA	S	NA	1	UNI	NA	NA	4.914	125	1	0	0	0	0	1	8.000	8.000	0.000	0.000	0.000
976	IRE	S1A-090	D	F	NA	NA	S	NA	1	UNI	NA	NA	4.914	942	2	2	1	1	0	6	6.369	2.123	2.123	1.062	1.062
977	IRE	S1A-090	E	F	NA	NA	S	NA	1	NA	NA	NA	4.914	279	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
978	IRE	S1A-091	A	F	26-33	2	N	C	3	UNI	ACD	ENG	11.463	391	0	1	1	1	0	3	7.673	0.000	2.558	2.558	2.558
979	IRE	S1A-091	B	F	26-33	2	N	C	3	UNI	ACD	ENG	11.463	301	0	0	0	0	1	1	3.322	0.000	0.000	0.000	0.000
980	IRE	S1A-091	C	F	26-33	2	N	P	3	SEC	SLM	ENG	11.463	873	2	5	4	0	0	11	12.600	2.291	5.727	4.582	0.000
981	IRE	S1A-091	D	F	26-33	2	N	C	3	UNI	ACD	ENG	11.463	610	2	1	2	1	0	6	9.836	3.279	1.639	3.279	1.639
982	IRE	S1A-092	A	F	26-33	2	N	P	3	SEC	SLM	ENG	6.828	1130	1	4	2	0	0	7	6.195	0.885	3.540	1.770	0.000
983	IRE	S1A-092	B	F	26-33	2	N	C	3	UNI	ACD	ENG	6.828	774	0	2	3	0	1	6	7.752	0.000	2.584	3.876	0.000
985	IRE	S1A-093	B	M	34-49	4	N	P	3	SEC	SLM	ENG	2.555	1174	0	1	2	0	0	3	2.555	0.000	0.852	1.704	0.000
986	IRE	S1A-094	A	F	26-33	2	NA	C	3	UNI	ACD	ENG	12.775	1433	5	4	12	1	0	22	15.352	3.489	2.791	8.374	0.698
987	IRE	S1A-094	B	F	26-33	2	N	C	3	UNI	ACD	ENG	12.775	524	0	0	3	0	0	3	5.725	0.000	0.000	5.725	0.000
988	IRE	S1A-095	A	F	26-33	2	N	C	3	UNI	CLM	ENG	8.134	1190	1	2	10	0	0	13	10.924	0.840	1.681	8.403	0.000

989	IRE	S1A-095	B	F	0-25	1	N	C	3	UNI	ACD	ENG	8.134	777	1	0	1	1	0	3	3.861	1.287	0.000	1.287	1.287
990	IRE	S1A-096	A	F	50+	4	S	C	3	SEC	SLM	ENG+	0.522	1158	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
991	IRE	S1A-096	B	F	50+	4	S	NA	3	UNI	CLM	ENG+	0.522	756	0	0	1	0	0	1	1.323	0.000	0.000	1.323	0.000
992	IRE	S1A-097	A	F	50+	4	S	NA	3	UNI	CLM	ENG+	1.401	1470	2	0	1	0	0	3	2.041	1.361	0.000	0.680	0.000
993	IRE	S1A-097	B	F	50+	4	S	C	3	SEC	SLM	ENG+	1.401	671	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
994	IRE	S1A-098	A	F	26-33	2	NA	C	3	UNI	ACD	ENG	5.058	984	1	1	3	0	0	5	5.081	1.016	1.016	3.049	0.000
995	IRE	S1A-098	B	F	0-25	1	S	C	3	SEC	CLM	ENG	5.058	993	1	0	2	1	1	5	5.035	1.007	0.000	2.014	1.007
996	IRE	S1A-099	A	F	0-25	1	S	C	3	UNI	EDU	ENG	11.500	1187	5	6	7	3	0	21	17.692	4.212	5.055	5.897	2.527
997	IRE	S1A-099	B	F	0-25	1	S	C	3	UNI	EDU	ENG	11.500	726	0	1	0	0	0	1	1.377	0.000	1.377	0.000	0.000
998	IRE	S1A-100	A	F	0-25	1	S	NA	3	SEC	EDU	ENG+	12.827	1053	6	5	0	2	0	13	12.346	5.698	4.748	0.000	1.899
999	IRE	S1A-100	B	M	0-25	1	S	NA	3	SEC	EDU	ENG	12.827	974	4	3	1	5	0	13	13.347	4.107	3.080	1.027	5.133

Table 142: Final data set of IndE.

ID	VAR	FILE	SPEAKER	SEX	age.orig	AGE	ETH	DATE	EDU	OCC	L1	PAI	words	ini.n	med.n	fin.m	non.n	na.n	all.n	all	ini	med	fin	non
1000	IND	S1A-001	A	F	42+	4	IND	1	UNI	ACD	MAR	1.948	777	0	1	0	2	0	3	3.861	0.000	1.287	0.000	2.574
1001	IND	S1A-001	B	F	34-41	3	IND	1	UNI	ACD	OTH	1.948	763	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1002	IND	S1A-002	A	M	26-33	2	IND	2	UNI	ACD	TAM	0.576	886	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1003	IND	S1A-002	B	M	18-25	1	IND	2	UNI	ACD	TEL	0.576	849	0	0	0	1	0	1	1.178	0.000	0.000	0.000	1.178
1004	IND	S1A-003	A	M	34-41	3	IND	1	UNI	CLM	MAR	0.000	1257	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1005	IND	S1A-003	B	F	34-41	3	IND	1	UNI	ACD	PUN	0.000	847	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1006	IND	S1A-004	A	F	26-33	2	IND	1	UNI	ACD	OTH	0.000	1058	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1008	IND	S1A-005	A	M	42+	4	IND	1	UNI	ACD	OTH	0.000	1113	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1007	IND	S1A-005	B	M	26-33	2	IND	1	UNI	ACD	MAR	0.000	1248	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1009	IND	S1A-005	B	M	26-33	2	IND	1	UNI	ACD	OTH	0.000	1015	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1010	IND	S1A-006	A	M	34-41	3	IND	1	UNI	ACD	HIN	0.000	1285	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1011	IND	S1A-006	B	M	26-33	2	IND	1	UNI	ACD	HIN	0.000	836	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1012	IND	S1A-007	A	F	26-33	2	IND	1	UNI	ACD	OTH	3.212	708	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1014	IND	S1A-007	C	F	26-33	2	IND	1	UNI	ACD	PUN	3.212	499	2	0	0	0	0	2	4.008	4.008	0.000	0.000	0.000
1015	IND	S1A-007	D	F	34-41	3	IND	1	UNI	ACD	OTH	3.212	435	0	0	0	1	0	1	2.299	0.000	0.000	0.000	2.299
1016	IND	S1A-008	A	F	18-25	1	IND	1	UNI	ACD	PUN	5.487	1069	3	6	0	3	0	12	11.225	2.806	5.613	0.000	2.806
1017	IND	S1A-008	B	F	26-33	2	IND	1	UNI	ACD	MAR	5.487	406	0	0	0	1	0	1	2.463	0.000	0.000	0.000	2.463
1018	IND	S1A-008	C	F	26-33	2	IND	1	UNI	ACD	PUN	5.487	472	0	0	0	3	0	3	6.356	0.000	0.000	0.000	6.356
1019	IND	S1A-008	D	NA	NA	NA	NA	1	NA	NA	NA	5.487	257	0	1	0	0	0	1	3.891	0.000	3.891	0.000	0.000
1020	IND	S1A-009	A	M	26-33	2	IND	1	UNI	ACD	TAM	0.000	1015	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1021	IND	S1A-009	B	M	26-33	2	IND	1	UNI	ACD	OTH	0.000	1216	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1022	IND	S1A-010	A	F	26-33	2	IND	1	UNI	ACD	OTH	1.078	1108	1	0	0	0	0	1	0.903	0.903	0.000	0.000	0.000
1023	IND	S1A-010	B	F	34-41	3	IND	1	UNI	ACD	OTH	1.078	110	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1024	IND	S1A-010	C	NA	NA	NA	NA	1	NA	NA	NA	1.078	818	1	0	0	0	0	1	1.222	1.222	0.000	0.000	0.000
1025	IND	S1A-011	A	F	42+	4	IND	1	UNI	ACD	TAM	0.773	265	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1027	IND	S1A-011	C	F	26-33	2	IND	1	UNI	ACD	TAM	0.773	293	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1028	IND	S1A-011	D	F	26-33	2	IND	1	UNI	ACD	TEL	0.773	324	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1029	IND	S1A-011	E	F	18-25	1	IND	1	UNI	EDU	TEL	0.773	194	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1030	IND	S1A-011	F	M	18-25	1	IND	1	UNI	ACD	TAM	0.773	801	0	1	0	0	0	1	1.248	0.000	1.248	0.000	0.000
1031	IND	S1A-011	G	M	18-25	1	IND	1	UNI	EDU	TAM	0.773	492	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1032	IND	S1A-012	A	F	26-33	2	IND	1	UNI	EDU	MAL	5.960	735	1	2	0	3	0	6	8.163	1.361	2.721	0.000	4.082
1033	IND	S1A-012	B	F	18-25	1	IND	1	UNI	EDU	PUN	5.960	1614	1	2	1	4	0	8	4.957	0.620	1.239	0.620	2.478
1034	IND	S1A-013	A	M	34-41	3	IND	1	UNI	ACD	PUN	0.645	553	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1035	IND	S1A-013	B	M	34-41	3	IND	1	UNI	ACD	MAR	0.645	1240	0	1	0	0	0	1	0.806	0.000	0.806	0.000	0.000
1036	IND	S1A-013	C	NA	NA	NA	NA	1	NA	NA	NA	0.645	311	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1037	IND	S1A-014	A	F	26-33	2	IND	2	UNI	ACD	TEL	4.690	1070	1	2	0	4	1	8	7.477	0.935	1.869	0.000	3.738
1038	IND	S1A-014	B	F	26-33	2	IND	2	UNI	ACD	OTH	4.690	1062	2	0	0	0	0	2	1.883	1.883	0.000	0.000	0.000
1039	IND	S1A-015	A	M	42+	4	IND	1	UNI	ACD	TAM	0.000	944	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1040	IND	S1A-015	B	M	42+	4	IND	1	UNI	ACD	TEL	0.000	914	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1041	IND	S1A-016	A	F	26-33	2	IND	2	UNI	ACD	TAM	2.844	1306	0	1	0	2	0	3	2.297	0.000	0.766	0.000	1.531
1042	IND	S1A-016	B	F	26-33	2	IND	2	UNI	ACD	TAM	2.844	804	1	0	1	1	0	3	3.731	1.244	0.000	1.244	1.244
1043	IND	S1A-017	A	M	18-25	1	IND	1	UNI	EDU	KANN	1.784	980	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1044	IND	S1A-017	B	M	18-25	1	IND	1	UNI	EDU	KANN	1.784	395	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1045	IND	S1A-017	C	M	18-25	1	IND	1	UNI	EDU	KANN	1.784	726	0	2	0	0	0	2	2.755	0.000	2.755	0.000	0.000
1046	IND	S1A-018	A	M	42+	4	IND	1	UNI	NA	MAR	0.478	818	0	0	1	0	0	1	1.222	0.000	0.000	1.222	0.000
1047	IND	S1A-018	B	M	42+	4	IND	1	SEC	SLM	KANN	0.478	1276	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1048	IND	S1A-019	A	M	34-41	3	IND	1	UNI	ACD	TAM	0.000	647	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1049	IND	S1A-019	B	M	34-41	3	IND	1	UNI	ACD	KANN	0.000	615	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1050	IND	S1A-019	C	M	42+	4	IND	1	UNI	ACD	OTH	0.000	287	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1051	IND	S1A-019	P	NA	NA	NA	NA	1	NA	NA	NA	0.000	400	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1056	IND	S1A-020	A	M	26-33	2	IND	1	UNI	ACD	KANN	0.000	825	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1057	IND	S1A-020	B	M	42+	4	IND	1	UNI	ACD	KANN	0.000	752	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1058	IND	S1A-020	C	M	34-41	3	IND	1	UNI	ACD	TAM	0.000	431	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1059	IND	S1A-021	A	F	18-25	1	IND	1	UNI	ACD	TAM	2.288	1181	2	1	0	2	0	5	4.234	1.693	0.847	0.000	1.693
1060	IND	S1A-021	B	F	18-25	1	IND	1	UNI	ACD	OTH	2.288	368	1	1	0	0	0	2	5.435	2.717	2.717	0.000	0.000
1061	IND	S1A-021	C	F	18-25	1	IND	1	UNI	ACD	TEL	2.288	506	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1062	IND	S1A-022	A	F	26-33	2	IND	1	UNI	ACD	OTH	1.431	846	0	2	0	0	0	2	2.364	0.000	2.364	0.000	0.000
1063	IND	S1A-022	B	F	34-41	3	IND	1	UNI	ACD	TEL	1.431	1251	0	0	0	1	0	1	0.799	0.000	0.000	0.000	0.799
1064	IND	S1A-023	A	M	26-33	2	IND	1	UNI	ACD	TAM	0.000	1593	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1065	IND	S1A-023	B	M	42+	4	IND	1	UNI	ACD	KANN	0.000	740	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1066	IND	S1A-024	A	M	26-33	2	IND	1	UNI	ACD	TEL	0.000	846	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1067	IND	S1A-024	B	M	26-33	2	IND	1	UNI	ACD	TAM	0.000	532	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1068	IND	S1A-024	C	M	42+	4	IND	1	UNI	ACD	TAM	0.000	656	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1069	IND	S1A-025	A	M	34-41	3	IND	1	UNI	ACD	MAL	1.391	690	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1070	IND	S1A-025	B	M	26-33	2	IND	1	UNI	ACD	MAL	1.391	702	0	0	0	2	0	2	2.849	0.000	0.000	0.000	2.849
1071	IND	S1A-025	C	M	34-41	3	IND	1	UNI	ACD	MAL	1.391	736	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1072	IND	S1A-026	A	M	34-41	3	IND	1	UNI	CLM	MAL	0.000	1223	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1073	IND	S1A-026	B	M	42+	4	IND	1	UNI	ACD	KANN	0.000	919	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1074	IND	S1A-027	A	M	26-33	2	IND	1	UNI	ACD	TAM	0.000	743	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1075	IND	S1A-027	B	M	34-41	3	IND	1	UNI	EDU	TEL	0.000	689	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1076	IND	S1A-027	C	M	42+	4	IND	1	UNI	ACD	TAM	0.000	554	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1077	IND	S1A-028	A	F	34-41	3	IND	2	UNI	ACD	OTH	3.439	1520	0	1	1	3	0	5	3.289	0.000	0.658	0.658	1.974
1078	IND	S1A-028	B	M	34-41	3	IND	2	UNI	ACD	OTH	3.439	806	1	1	1	0	0	3	3.722	1.241	1.241	1.241	0.000

1079	IND	S1A-029	A	F	26-33	2	IND	1	UNI	ACD	TAM	0.803	725	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1080	IND	S1A-029	B	F	42+	4	IND	1	UNI	ACD	MAL	0.803	738	0	0	0	1	0	1	1.355	0.000	0.000	0.000	1.355
1081	IND	S1A-029	C	F	34-41	3	IND	1	UNI	NA	MAL	0.803	507	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1082	IND	S1A-030	A	F	34-41	3	IND	1	UNI	ACD	TAM	0.740	685	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1083	IND	S1A-030	B	F	34-41	3	IND	1	UNI	ACD	TAM	0.740	698	0	0	0	1	0	1	1.433	0.000	0.000	0.000	1.433
1084	IND	S1A-030	C	F	26-33	2	IND	1	UNI	ACD	TAM	0.740	654	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1085	IND	S1A-031	A	F	26-33	2	IND	1	UNI	ACD	KANN	0.000	897	0	0	0	1	0	1	1.115	0.000	0.000	0.000	1.115
1086	IND	S1A-031	B	F	18-25	1	IND	1	UNI	ACD	OTH	0.000	630	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1087	IND	S1A-031	C	F	34-41	3	IND	1	UNI	ACD	TEL	0.000	441	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1088	IND	S1A-032	A	M	42+	4	IND	1	UNI	ACD	OTH	0.000	943	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1089	IND	S1A-032	B	M	42+	4	IND	1	UNI	ACD	MAR	0.000	960	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1090	IND	S1A-032	C	M	42+	4	IND	1	UNI	ACD	OTH	0.000	943	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1091	IND	S1A-032	D	M	42+	4	IND	1	UNI	ACD	MAR	0.000	960	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1092	IND	S1A-033	A	M	34-41	3	IND	1	UNI	ACD	TAM	0.000	658	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1093	IND	S1A-033	B	M	34-41	3	IND	1	UNI	ACD	MAL	0.000	1128	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1094	IND	S1A-033	C	M	42+	4	IND	1	UNI	ACD	MAR	0.000	417	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1095	IND	S1A-034	A	F	42+	4	IND	1	UNI	ACD	KANN	3.231	731	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1096	IND	S1A-034	B	M	42+	4	IND	1	UNI	ACD	OTH	3.231	296	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1097	IND	S1A-034	C	M	42+	4	IND	1	UNI	ACD	HIN	3.231	323	0	0	0	2	0	2	6.192	0.000	0.000	0.000	6.192
1098	IND	S1A-035	A	M	42+	4	IND	2	UNI	ACD	MAL	0.000	1382	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1099	IND	S1A-035	B	M	42+	4	IND	2	UNI	ACD	MAL	0.000	900	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1100	IND	S1A-036	A	M	26-33	2	IND	2	UNI	ACD	MAR	0.507	932	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1101	IND	S1A-036	B	F	34-41	3	IND	2	UNI	ACD	OTH	0.507	1042	0	0	0	1	0	1	0.960	0.000	0.000	0.000	0.960
1102	IND	S1A-037	A	F	18-25	1	IND	1	UNI	EDU	MAR	1.912	906	4	0	1	1	0	6	6.623	4.415	0.000	1.104	1.104
1103	IND	S1A-037	B	F	18-25	1	IND	1	UNI	EDU	MAR	1.912	419	0	0	0	1	0	1	2.387	0.000	0.000	0.000	2.387

1104	IND	S1A-037	C	F	NA	NA	IND	1	UNI	EDU	MAR	1.912	411	1	0	0	0	0	1	2.433	2.433	0.000	0.000	0.000
1105	IND	S1A-037	D	F	18-25	1	IND	1	UNI	EDU	MAR	1.912	112	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1106	IND	S1A-038	A	M	26-33	2	IND	1	UNI	CLM	KANN	0.000	1027	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1107	IND	S1A-038	B	M	26-33	2	IND	1	UNI	CLM	KANN	0.000	1287	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1108	IND	S1A-039	A	F	18-25	1	IND	1	UNI	EDU	OTH	11.152	1017	3	0	0	1	0	4	3.933	2.950	0.000	0.000	0.983
1109	IND	S1A-039	B	F	18-25	1	IND	1	UNI	EDU	OTH	11.152	636	4	0	1	5	0	10	15.723	6.289	0.000	1.572	7.862
1110	IND	S1A-039	C	F	18-25	1	IND	1	UNI	EDU	NA	11.152	440	1	1	0	0	0	2	4.545	2.273	2.273	0.000	0.000
1111	IND	S1A-040	A	F	18-25	1	IND	1	UNI	EDU	KANN	2.505	724	0	0	1	0	0	1	1.381	0.000	0.000	1.381	0.000
1112	IND	S1A-040	B	F	18-25	1	IND	1	UNI	EDU	KANN	2.505	1671	1	0	1	3	0	5	2.992	0.598	0.000	0.598	1.795
1113	IND	S1A-041	A	M	34-41	3	IND	1	UNI	CLM	MAL	2.073	1215	0	0	3	0	1	4	3.292	0.000	0.000	2.469	0.000
1114	IND	S1A-041	B	M	26-33	2	IND	1	UNI	CLM	HIN	2.073	715	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1116	IND	S1A-042	A	F	18-25	1	IND	1	UNI	ACD	ENG	5.268	821	3	1	0	1	0	5	6.090	3.654	1.218	0.000	1.218
1117	IND	S1A-042	B	F	18-25	1	IND	1	UNI	ACD	OTH	5.268	360	0	0	0	1	0	1	2.778	0.000	0.000	0.000	2.778
1118	IND	S1A-042	C	M	26-33	2	IND	1	UNI	CLM	HIN	5.268	779	5	0	0	0	0	5	6.418	6.418	0.000	0.000	0.000
1119	IND	S1A-043	A	F	34-41	3	IND	1	UNI	ACD	MAR	2.265	1264	1	0	0	0	0	1	0.791	0.791	0.000	0.000	0.000
1120	IND	S1A-043	B	F	18-25	1	IND	1	SEC	SLM	KANN	2.265	502	1	2	0	0	0	3	5.976	1.992	3.984	0.000	0.000
1121	IND	S1A-044	A	F	34-41	3	IND	1	UNI	NA	MAR	0.921	1086	0	0	0	1	0	1	0.921	0.000	0.000	0.000	0.921
1123	IND	S1A-045	A	M	42+	4	IND	1	UNI	ACD	MAR	2.842	1065	0	1	0	3	0	4	3.756	0.000	0.939	0.000	2.817
1124	IND	S1A-045	B	M	26-33	2	IND	1	UNI	ACD	MAR	2.842	1046	0	1	0	1	0	2	1.912	0.000	0.956	0.000	0.956
1125	IND	S1A-046	A	F	34-41	3	IND	1	UNI	ACD	HIN	0.665	589	1	1	0	0	0	2	3.396	1.698	1.698	0.000	0.000
1126	IND	S1A-046	B	F	34-41	3	IND	1	UNI	ACD	MAL	0.665	530	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1127	IND	S1A-046	C	F	34-41	3	IND	1	UNI	ACD	OTH	0.665	973	1	0	0	0	0	1	1.028	1.028	0.000	0.000	0.000
1128	IND	S1A-047	A	F	42+	4	IND	1	UNI	ACD	OTH	0.858	809	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1129	IND	S1A-047	B	M	26-33	2	IND	1	UNI	ACD	OTH	0.858	1523	1	1	0	0	0	2	1.313	0.657	0.657	0.000	0.000
1130	IND	S1A-048	A	F	18-25	1	IND	1	UNI	EDU	TAM	3.311	760	0	1	1	2	0	4	5.263	0.000	1.316	1.316	2.632

1131	IND	S1A-048	B	F	18-25	1	IND	1	UNI	EDU	ENG	3.311	509	0	1	0	0	0	1	1.965	0.000	1.965	0.000	0.000
1132	IND	S1A-048	C	F	18-25	1	IND	1	UNI	EDU	MAL	3.311	699	0	0	0	3	0	3	4.292	0.000	0.000	0.000	4.292
1133	IND	S1A-049	A	F	18-25	1	IND	1	UNI	SLM	TAM	0.000	955	0	1	0	3	0	4	4.188	0.000	1.047	0.000	3.141
1134	IND	S1A-049	B	F	18-25	1	IND	1	UNI	EDU	KANN	0.000	975	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1135	IND	S1A-049	C	F	18-25	1	IND	1	UNI	EDU	MAL	0.000	128	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1136	IND	S1A-050	A	M	18-25	1	IND	1	UNI	EDU	KANN	2.533	919	1	0	0	0	0	1	1.088	1.088	0.000	0.000	0.000
1137	IND	S1A-050	B	M	18-25	1	IND	1	UNI	EDU	KANN	2.533	1055	2	0	0	2	0	4	3.791	1.896	0.000	0.000	1.896
1138	IND	S1A-051	A	M	18-25	1	IND	1	UNI	EDU	MAR	0.941	1019	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1139	IND	S1A-051	B	M	18-25	1	IND	1	UNI	EDU	KANN	0.941	783	0	1	0	0	0	1	1.277	0.000	1.277	0.000	0.000
1140	IND	S1A-051	C	M	18-25	1	IND	1	UNI	EDU	MAR	0.941	280	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1141	IND	S1A-052	A	F	18-25	1	IND	1	UNI	EDU	KANN	1.426	1350	0	0	0	0	1	1	0.741	0.000	0.000	0.000	0.000
1142	IND	S1A-052	B	F	18-25	1	IND	1	UNI	EDU	KANN	1.426	754	1	0	1	0	0	2	2.653	1.326	0.000	1.326	0.000
1143	IND	S1A-053	A	F	18-25	1	IND	1	UNI	EDU	MAR	1.848	1177	2	1	0	1	0	4	3.398	1.699	0.850	0.000	0.850
1144	IND	S1A-053	B	F	18-25	1	IND	1	UNI	EDU	KANN	1.848	988	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1145	IND	S1A-054	A	F	18-25	1	IND	1	UNI	ACD	OTH	5.386	1789	5	5	0	1	1	12	6.708	2.795	2.795	0.000	0.559
1146	IND	S1A-054	B	F	18-25	1	IND	1	UNI	ACD	OTH	5.386	439	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1147	IND	S1A-055	A	M	18-25	1	IND	1	UNI	EDU	MAR	6.482	1170	0	0	0	1	0	1	0.855	0.000	0.000	0.000	0.855
1148	IND	S1A-055	B	F	18-25	1	IND	1	UNI	ACD	MAR	6.482	1144	7	1	0	5	1	14	12.238	6.119	0.874	0.000	4.371
1149	IND	S1A-056	A	F	18-25	1	IND	1	UNI	EDU	OTH	5.076	222	0	0	0	1	0	1	4.505	0.000	0.000	0.000	4.505
1150	IND	S1A-056	B	F	18-25	1	IND	1	UNI	EDU	OTH	5.076	668	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1151	IND	S1A-056	C	M	18-25	1	IND	1	UNI	EDU	KANN	5.076	1105	4	0	0	3	2	9	8.145	3.620	0.000	0.000	2.715
1152	IND	S1A-057	A	M	42+	4	IND	1	UNI	ACD	TAM	0.508	1446	0	1	0	0	0	1	0.692	0.000	0.692	0.000	0.000
1153	IND	S1A-057	B	M	42+	4	IND	1	UNI	ACD	TAM	0.508	524	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1154	IND	S1A-058	A	F	26-33	2	IND	1	UNI	CLM	TAM	0.000	1004	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1155	IND	S1A-058	B	M	26-33	2	IND	1	UNI	CLM	TAM	0.000	917	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1156	IND	S1A-059	A	F	18-25	1	IND	1	UNI	EDU	TEL	3.864	814	0	0	1	2	0	3	3.686	0.000	0.000	1.229	2.457
1157	IND	S1A-059	B	F	18-25	1	IND	1	UNI	EDU	TEL	3.864	480	0	0	0	2	0	2	4.167	0.000	0.000	0.000	4.167
1158	IND	S1A-060	A	F	26-33	2	IND	1	UNI	ACD	OTH	0.460	1205	0	1	0	0	0	1	0.830	0.000	0.830	0.000	0.000
1159	IND	S1A-060	B	M	42+	4	IND	1	UNI	ACD	TEL	0.460	970	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1160	IND	S1A-061	A	F	18-25	1	IND	2	UNI	ACD	KANN	1.869	1183	0	0	0	1	0	1	0.845	0.000	0.000	0.000	0.845
1161	IND	S1A-061	B	M	18-25	1	IND	2	UNI	EDU	KANN	1.869	957	0	3	0	0	0	3	3.135	0.000	3.135	0.000	0.000
1162	IND	S1A-062	A	F	18-25	1	IND	2	UNI	EDU	MAR	2.100	1686	2	0	0	2	0	4	2.372	1.186	0.000	0.000	1.186
1163	IND	S1A-062	B	F	18-25	1	IND	2	UNI	EDU	HIN	2.100	695	0	0	0	1	0	1	1.439	0.000	0.000	0.000	1.439
1164	IND	S1A-063	A	M	34-41	3	IND	1	UNI	EDU	KANN	0.000	802	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1165	IND	S1A-063	B	M	26-33	2	IND	1	UNI	ACD	KANN	0.000	1322	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1166	IND	S1A-064	A	F	26-33	2	IND	1	UNI	ACD	MAL	2.516	1166	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1167	IND	S1A-064	B	F	34-41	3	IND	1	UNI	ACD	TAM	2.516	1219	4	2	0	0	0	6	4.922	3.281	1.641	0.000	0.000
1168	IND	S1A-065	A	F	18-25	1	IND	1	UNI	ACD	HIN	3.873	1316	0	0	2	2	0	4	3.040	0.000	0.000	1.520	1.520
1169	IND	S1A-065	B	F	26-33	2	IND	1	UNI	ACD	OTH	3.873	1008	3	0	1	1	0	5	4.960	2.976	0.000	0.992	0.992
1170	IND	S1A-066	A	F	26-33	2	IND	1	UNI	ACD	MAR	0.000	1465	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1171	IND	S1A-066	B	M	42+	4	IND	1	UNI	ACD	OTH	0.000	804	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1173	IND	S1A-067	A	M	42+	4	IND	1	UNI	ACD	MAR	0.000	1247	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1174	IND	S1A-067	B	M	42+	4	IND	1	UNI	ACD	MAR	0.000	940	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1175	IND	S1A-068	A	F	42+	4	IND	1	UNI	ACD	HIN	1.804	1166	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1176	IND	S1A-068	B	M	42+	4	IND	1	UNI	ACD	HIN	1.804	1051	1	0	0	3	0	4	3.806	0.951	0.000	0.000	2.854
1178	IND	S1A-069	A	M	42+	4	IND	1	UNI	ACD	TEL	2.062	937	1	0	0	0	0	1	1.067	1.067	0.000	0.000	0.000
1179	IND	S1A-069	B	F	26-33	2	IND	1	UNI	CLM	MAR	2.062	1488	2	1	0	1	0	4	2.688	1.344	0.672	0.000	0.672
1181	IND	S1A-070	A	F	18-25	1	IND	1	UNI	EDU	KANN	0.000	1249	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1182	IND	S1A-070	B	F	18-25	1	IND	1	UNI	EDU	KANN	0.000	547	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1183	IND	S1A-070	C	F	18-25	1	IND	1	UNI	EDU	KANN	0.000	259	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1184	IND	S1A-071	A	F	18-25	1	IND	1	UNI	EDU	KANN	1.350	1528	0	0	0	1	0	1	0.654	0.000	0.000	0.000	0.654
1185	IND	S1A-071	B	F	18-25	1	IND	1	UNI	EDU	KANN	1.350	695	1	0	0	1	0	2	2.878	1.439	0.000	0.000	1.439
1186	IND	S1A-072	A	F	42+	4	IND	2	UNI	ACD	OTH	0.848	804	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1187	IND	S1A-072	B	F	34-41	3	IND	2	UNI	ACD	KANN	0.848	1554	1	1	0	0	0	2	1.287	0.644	0.644	0.000	0.000
1189	IND	S1A-073	B	F	18-25	1	IND	2	UNI	EDU	KANN	13.245	453	2	0	1	3	0	6	13.245	4.415	0.000	2.208	6.623
1190	IND	S1A-074	A	M	42+	4	IND	2	UNI	ACD	MAR	0.000	881	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1191	IND	S1A-074	B	M	34-41	3	IND	2	UNI	ACD	MAR	0.000	1416	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1192	IND	S1A-075	A	F	26-33	2	IND	2	UNI	ACD	MAR	0.476	1273	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1193	IND	S1A-075	B	F	26-33	2	IND	2	UNI	ACD	MAR	0.476	830	0	0	0	1	0	1	1.205	0.000	0.000	0.000	1.205
1195	IND	S1A-076	A	M	42+	4	IND	2	UNI	ACD	KANN	0.000	1791	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1196	IND	S1A-076	B	M	42+	4	IND	2	UNI	ACD	MAR	0.000	400	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1197	IND	S1A-077	A	M	18-25	1	IND	2	UNI	EDU	PUN	3.111	890	5	0	0	1	0	6	6.742	5.618	0.000	0.000	1.124
1198	IND	S1A-077	B	M	34-41	3	IND	2	UNI	CLM	PUN	3.111	1360	0	0	0	1	0	1	0.735	0.000	0.000	0.000	0.735
1200	IND	S1A-078	A	M	34-41	3	IND	2	UNI	ACD	PUN	0.950	1799	0	1	0	0	0	1	0.556	0.000	0.556	0.000	0.000
1201	IND	S1A-078	B	F	26-33	2	IND	2	UNI	ACD	PUN	0.950	307	0	0	0	1	0	1	3.257	0.000	0.000	0.000	3.257
1203	IND	S1A-079	A	M	42+	4	IND	2	UNI	ACD	HIN	0.000	883	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1204	IND	S1A-079	B	F	26-33	2	IND	2	UNI	ACD	HIN	0.000	1283	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1205	IND	S1A-080	A	F	42+	4	IND	2	UNI	ACD	PUN	0.486	766	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1206	IND	S1A-080	B	M	42+	4	IND	2	UNI	ACD	PUN	0.486	1292	1	0	0	0	0	1	0.774	0.774	0.000	0.000	0.000
1207	IND	S1A-081	A	M	42+	4	IND	2	UNI	ACD	PUN	0.904	1196	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1208	IND	S1A-081	B	M	42+	4	IND	2	UNI	ACD	PUN	0.904	1017	0	0	0	2	0	2	1.967	0.000	0.000	0.000	1.967
1209	IND	S1A-082	A	M	42+	4	IND	2	UNI	ACD	HIN	1.156	659	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1210	IND	S1A-082	B	M	42+	4	IND	2	UNI	ACD	PUN	1.156	1936	0	1	0	2	0	3	1.550	0.000	0.517	0.000	1.033
1211	IND	S1A-083	A	M	26-33	2	IND	2	UNI	EDU	HIN	0.000	1447	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1212	IND	S1A-083	B	M	18-25	1	IND	2	UNI	EDU	OTH	0.000	713	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1214	IND	S1A-084	A	F	26-33	2	IND	2	UNI	ACD	MAR	2.500	1049	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1215	IND	S1A-084	B	F	42+	4	IND	2	UNI	ACD	MAL	2.500	951	0	4	0	1	0	5	5.258	0.000	4.206	0.000	1.052
1216	IND	S1A-085	A	F	26-33	2	IND	2	UNI	ACD	OTH	0.456	849	0	0	0	1	0	1	1.178	0.000	0.000	0.000	1.178
1217	IND	S1A-085	B	F	18-25	1	IND	2	UNI	ACD	MAR	0.456	1345	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1218	IND	S1A-086	A	F	34-41	3	IND	2	UNI	ACD	MAR	0.000	1113	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1219	IND	S1A-086	B	F	26-33	2	IND	2	UNI	ACD	KANN	0.000	995	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1220	IND	S1A-087	A	M	26-33	2	IND	2	UNI	ACD	KANN	0.000	786	1	0	0	0	0	1	1.272	1.272	0.000	0.000	0.000
1221	IND	S1A-087	B	F	18-25	1	IND	2	UNI	ACD	MAR	0.000	1346	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1222	IND	S1A-087	C	NA	NA	NA	NA	2	NA	NA	NA	0.000	109	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1223	IND	S1A-088	A	F	42+	4	IND	2	UNI	ACD	MAR	0.412	1347	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1224	IND	S1A-088	B	F	42+	4	IND	2	UNI	ACD	MAR	0.412	1079	1	0	0	0	0	1	0.927	0.927	0.000	0.000	0.000
1225	IND	S1A-089	A	M	42+	4	IND	2	UNI	ACD	MAR	0.558	880	0	1	0	0	0	1	1.136	0.000	1.136	0.000	0.000
1226	IND	S1A-089	B	M	42+	4	IND	2	UNI	ACD	MAR	0.558	911	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1229	IND	S1A-090	B	M	18-25	1	IND	1	UNI	EDU	OTH	17.167	2097	12	3	2	19	0	36	17.167	5.722	1.431	0.954	9.061
1231	IND	S1A-091	B	F	42+	4	IND	1	UNI	ACD	TAM	0.451	1264	1	0	0	0	0	1	0.791	0.791	0.000	0.000	0.000
1232	IND	S1A-091	C	NA	NA	NA	NA	1	NA	NA	NA	0.451	953	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1233	IND	S1A-092	A	F	34-41	3	IND	1	UNI	CLM	MAR	0.000	1078	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1234	IND	S1A-092	B	F	26-33	2	IND	1	UNI	CLM	OTH	0.000	1197	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1235	IND	S1A-093	A	M	42+	4	IND	1	UNI	ACD	TEL	0.490	1291	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1236	IND	S1A-093	B	M	42+	4	IND	1	UNI	ACD	KANN	0.490	750	1	0	0	0	0	1	1.333	1.333	0.000	0.000	0.000
1237	IND	S1A-094	A	M	34-41	3	IND	1	UNI	CLM	MAR	0.709	1593	0	0	0	1	0	1	0.628	0.000	0.000	0.000	0.628
1238	IND	S1A-094	B	M	42+	4	IND	1	UNI	CLM	MAR	0.709	1228	0	0	0	1	0	1	0.814	0.000	0.000	0.000	0.814
1241	IND	S1A-095	A	M	42+	4	IND	1	UNI	ACD	OTH	0.675	654	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1242	IND	S1A-095	B	M	42+	4	IND	1	UNI	ACD	TEL	0.675	1070	0	0	0	1	0	1	0.935	0.000	0.000	0.000	0.935
1243	IND	S1A-095	C	M	42+	4	IND	1	UNI	ACD	KANN	0.675	412	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1245	IND	S1A-096(1)	A	M	34-41	3	IND	1	UNI	ACD	OTH	0.721	1314	0	1	0	1	0	2	1.522	0.000	0.761	0.000	0.761
1246	IND	S1A-096(1)	B	M	42+	4	IND	1	UNI	ACD	OTH	0.721	1131	1	0	0	0	0	1	0.884	0.884	0.000	0.000	0.000
1247	IND	S1A-096(1)	C	F	42+	4	IND	1	UNI	ACD	OTH	0.721	256	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1250	IND	S1A-097(1)	A	M	42+	4	IND	1	UNI	ACD	MAR	0.000	1280	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1251	IND	S1A-097(1)	B	F	42+	4	IND	1	UNI	ACD	KANN	0.000	960	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1252	IND	S1A-097(1)	C	F	42+	4	IND	1	UNI	ACD	KANN	0.000	129	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1256	IND	S1A-098	A	M	18-25	1	IND	1	UNI	ACD	MAL	0.000	1274	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1257	IND	S1A-098	B	F	18-25	1	IND	1	UNI	CLM	KANN	0.000	996	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1258	IND	S1A-099	A	M	42+	4	IND	1	UNI	ACD	MAR	0.778	1055	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1259	IND	S1A-099	B	M	26-33	2	IND	1	UNI	ACD	MAL	0.778	188	1	0	0	0	0	1	5.319	5.319	0.000	0.000	0.000
1260	IND	S1A-099	C	NA	NA	NA	NA	1	NA	NA	NA	0.778	1098	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1261	IND	S1A-100	A	M	26-33	2	IND	1	UNI	ACD	TEL	0.000	628	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1262	IND	S1A-100	B	M	42+	4	IND	1	UNI	ACD	TEL	0.000	1323	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

Table 143: Final data set of JamE.

ID	VAR	FILE	SPEAKER	SEX	age.orig	AGE	ETH	DATE	EDU	OCC	L1	PAI	words	ini.n	med.n	fin.m	non.n	na.n	all.n	all	ini	med	fin	non
1263	JAM	S1A-001	A	F	NA	NA	JAM	1	UNI	ACD	ENG	0.481	1418	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1264	JAM	S1A-001	B	F	NA	NA	JAM	1	SEC	CLM	NA	0.481	659	1	0	0	0	0	1	1.517	1.517	0.000	0.000	0.000
1265	JAM	S1A-002	A	M	46+	4	JAM	2	UNI	ACD	ENG	0.000	1211	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1266	JAM	S1A-002	B	F	46+	4	JAM	2	UNI	ACD	NA	0.000	289	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1267	JAM	S1A-002	C	M	26-45	3	JAM	2	UNI	ACD	ENG+	0.000	555	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1270	JAM	S1A-003	A	M	46+	4	JAM	2	UNI	ACD	ENG	0.000	620	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1271	JAM	S1A-003	B	M	46+	4	JAM	2	UNI	ACD	ENG	0.000	1553	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1274	JAM	S1A-004	A	M	46+	4	JAM	2	UNI	ACD	ENG	0.000	675	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1275	JAM	S1A-004	B	M	46+	4	JAM	2	UNI	ACD	ENG	0.000	1113	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1276	JAM	S1A-004	C	M	26-45	3	JAM	2	SEC	SLM	NA	0.000	165	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1281	JAM	S1A-005	A	M	46+	4	JAM	2	SEC	CLM	ENG	0.000	1000	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1282	JAM	S1A-005	B	M	46+	4	JAM	2	UNI	ACD	ENG	0.000	1040	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1284	JAM	S1A-006	A	M	46+	4	JAM	2	UNI	ACD	NA	2.128	1101	1	0	0	0	0	1	0.908	0.908	0.000	0.000	0.000
1285	JAM	S1A-006	B	M	26-45	3	JAM	2	UNI	ACD	ENG	2.128	826	0	3	0	0	0	3	3.632	0.000	3.632	0.000	0.000
1286	JAM	S1A-006	C	M	26-45	3	JAM	2	UNI	ACD	ENG	2.128	369	0	1	0	0	0	1	2.710	0.000	2.710	0.000	0.000
1287	JAM	S1A-006	NA	NA	NA	NA	NA	2	NA	NA	NA	2.128	101	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1288	JAM	S1A-007	A	M	17-25	1	JAM	2	UNI	CLM	ENG	4.663	2088	3	7	0	1	0	11	5.268	1.437	3.352	0.000	0.479
1291	JAM	S1A-007	Z	F	26-45	3	OTH	2	UNI	ACD	OTH	4.663	271	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1292	JAM	S1A-008	A	M	26-45	3	JAM	2	SEC	CLM	ENG	12.013	2196	10	17	0	6	0	33	15.027	4.554	7.741	0.000	2.732
1293	JAM	S1A-008	Z	F	26-45	3	OTH	2	UNI	ACD	OTH	12.013	551	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1294	JAM	S1A-009	A	F	17-25	1	JAM	3	UNI	ACD	ENG	3.872	1009	1	1	0	1	0	3	2.973	0.991	0.991	0.000	0.991
1295	JAM	S1A-009	B	M	17-25	1	JAM	3	UNI	ACD	ENG	3.872	1057	3	2	0	0	0	5	4.730	2.838	1.892	0.000	0.000
1297	JAM	S1A-010	A	F	NA	NA	JAM	3	NA	NA	NA	1.148	645	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1298	JAM	S1A-010	B	F	NA	NA	JAM	3	NA	NA	NA	1.148	453	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1299	JAM	S1A-010	C	M	26-45	3	JAM	3	NA	CLM	ENG	1.148	571	0	0	0	1	0	1	1.751	0.000	0.000	0.000	1.751
1301	JAM	S1A-010	E	M	26-45	3	JAM	3	NA	CLM	ENG	1.148	300	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1303	JAM	S1A-011	A	F	26-45	3	JAM	3	UNI	EDU	ENG+	4.970	1008	1	0	0	0	0	1	0.992	0.992	0.000	0.000	0.000
1304	JAM	S1A-011	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	4.970	162	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1306	JAM	S1A-011	D	F	26-45	3	JAM	3	UNI	CLM	ENG+	4.970	844	2	1	0	2	0	5	5.924	2.370	1.185	0.000	2.370
1308	JAM	S1A-012	A	F	26-45	3	JAM	3	UNI	CLM	ENG+	0.481	1403	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1309	JAM	S1A-012	B	F	26-45	3	JAM	3	UNI	EDU	ENG+	0.481	674	1	0	0	0	0	1	1.484	1.484	0.000	0.000	0.000
1312	JAM	S1A-013	A	F	26-45	3	JAM	3	UNI	CLM	ENG+	2.392	754	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1313	JAM	S1A-013	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	2.392	206	1	0	0	0	0	1	4.854	4.854	0.000	0.000	0.000
1314	JAM	S1A-013	C	F	26-45	3	JAM	3	UNI	EDU	ENG+	2.392	1048	0	1	0	1	0	2	1.908	0.000	0.954	0.000	0.954
1317	JAM	S1A-014	A	F	26-45	3	JAM	3	UNI	ACD	ENG	1.277	1657	7	5	0	2	0	14	8.449	4.225	3.018	0.000	1.207
1318	JAM	S1A-014	B	M	26-45	3	JAM	3	UNI	CLM	ENG	1.277	405	0	0	0	1	0	1	2.469	0.000	0.000	0.000	2.469
1319	JAM	S1A-014	Z	M	26-45	3	OTH	3	UNI	ACD	OTH	1.277	378	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1320	JAM	S1A-015	A	F	NA	NA	JAM	3	NA	NA	NA	3.398	436	1	0	0	0	0	1	2.294	2.294	0.000	0.000	0.000
1321	JAM	S1A-015	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	3.398	1624	2	4	0	0	0	6	3.695	1.232	2.463	0.000	0.000
1322	JAM	S1A-016	A	F	NA	NA	JAM	3	NA	NA	NA	1.941	342	0	1	0	0	0	1	2.924	0.000	2.924	0.000	0.000
1323	JAM	S1A-016	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	1.941	1719	0	3	0	0	0	3	1.745	0.000	1.745	0.000	0.000
1324	JAM	S1A-017	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	1.861	322	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1325	JAM	S1A-017	B	M	17-25	1	JAM	3	SEC	EDU	ENG+	1.861	1827	0	4	0	0	0	4	2.189	0.000	2.189	0.000	0.000
1326	JAM	S1A-018	A	F	NA	NA	JAM	3	NA	NA	ENG+	2.495	234	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1327	JAM	S1A-018	B	M	26-45	3	JAM	3	UNI	ACD	ENG+	2.495	1770	1	3	0	1	0	5	2.825	0.565	1.695	0.000	0.565
1328	JAM	S1A-019	A	M	26-45	3	JAM	3	UNI	ACD	ENG+	0.480	1865	1	0	0	0	0	1	0.536	0.536	0.000	0.000	0.000
1329	JAM	S1A-019	B	F	NA	NA	JAM	3	NA	NA	ENG+	0.480	219	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1330	JAM	S1A-020	A	M	26-45	3	JAM	3	UNI	ACD	ENG+	0.474	1868	0	1	0	0	0	1	0.535	0.000	0.535	0.000	0.000

1331	JAM	S1A-020	B	F	NA	NA	JAM	3	NA	NA	ENG+	0.474	241	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1332	JAM	S1A-021	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	4.367	1684	3	4	0	1	0	8	4.751	1.781	2.375	0.000	0.594
1333	JAM	S1A-021	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	4.367	377	0	1	0	0	0	1	2.653	0.000	2.653	0.000	0.000
1334	JAM	S1A-022	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	2.937	1589	2	1	0	3	0	6	3.776	1.259	0.629	0.000	1.888
1335	JAM	S1A-022	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	2.937	454	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1336	JAM	S1A-023	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	3.009	1287	1	4	0	1	0	6	4.662	0.777	3.108	0.000	0.777
1337	JAM	S1A-023	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	3.009	707	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1338	JAM	S1A-024	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	5.621	740	0	2	0	1	0	3	4.054	0.000	2.703	0.000	1.351
1339	JAM	S1A-024	B	F	17-25	1	JAM	3	UNI	EDU	PAT	5.621	1217	5	1	0	0	2	8	6.574	4.108	0.822	0.000	0.000
1340	JAM	S1A-025	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	3.028	1037	1	2	0	0	0	3	2.893	0.964	1.929	0.000	0.000
1341	JAM	S1A-025	B	F	17-25	1	JAM	3	SEC	EDU	ENG	3.028	1275	0	4	0	0	0	4	3.137	0.000	3.137	0.000	0.000
1342	JAM	S1A-026	A	F	17-25	1	JAM	3	SEC	EDU	ENG	1.936	1367	0	3	0	0	1	4	2.926	0.000	2.195	0.000	0.000
1343	JAM	S1A-026	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	1.936	699	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1344	JAM	S1A-027	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	9.246	439	0	5	0	0	1	6	13.667	0.000	11.390	0.000	0.000
1345	JAM	S1A-027	B	F	26-45	3	JAM	3	UNI	ACD	ENG+	9.246	1616	1	10	0	2	0	13	8.045	0.619	6.188	0.000	1.238
1347	JAM	S1A-028	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	2.633	989	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1348	JAM	S1A-028	B	F	17-25	1	JAM	3	UNI	EDU	PAT	2.633	1290	2	2	0	2	0	6	4.651	1.550	1.550	0.000	1.550
1349	JAM	S1A-029	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	3.351	754	1	1	0	0	0	2	2.653	1.326	1.326	0.000	0.000
1350	JAM	S1A-029	B	F	17-25	1	JAM	3	UNI	EDU	PAT	3.351	1335	0	4	0	1	0	5	3.745	0.000	2.996	0.000	0.749
1352	JAM	S1A-030	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	6.366	542	3	0	0	0	0	3	5.535	5.535	0.000	0.000	0.000
1353	JAM	S1A-030	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	6.366	1500	3	6	0	1	0	10	6.667	2.000	4.000	0.000	0.667
1354	JAM	S1A-031	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	9.955	1560	4	11	0	1	2	18	11.538	2.564	7.051	0.000	0.641
1355	JAM	S1A-031	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	9.955	449	1	0	1	0	0	2	4.454	2.227	0.000	2.227	0.000
1356	JAM	S1A-032	A	F	26-45	3	JAM	3	SEC	CLM	ENG+	1.433	1413	0	1	0	0	0	1	0.708	0.000	0.708	0.000	0.000
1357	JAM	S1A-032	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	1.433	680	2	0	0	0	0	2	2.941	2.941	0.000	0.000	0.000

1358	JAM	S1A-033	A	F	26-45	3	JAM	3	SEC	CLM	ENG+	1.826	1802	0	1	0	0	0	1	0.555	0.000	0.555	0.000	0.000
1359	JAM	S1A-033	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	1.826	388	2	1	0	0	0	3	7.732	5.155	2.577	0.000	0.000
1360	JAM	S1A-034	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.977	1155	1	0	0	0	0	1	0.866	0.866	0.000	0.000	0.000
1361	JAM	S1A-034	B	F	17-25	1	JAM	3	SEC	EDU	PAT	0.977	893	0	1	0	0	0	1	1.120	0.000	1.120	0.000	0.000
1362	JAM	S1A-035	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	6.363	1023	1	1	0	0	0	2	1.955	0.978	0.978	0.000	0.000
1363	JAM	S1A-035	B	F	17-25	1	JAM	3	SEC	EDU	PAT	6.363	1020	1	8	0	1	1	11	10.784	0.980	7.843	0.000	0.980
1364	JAM	S1A-036	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	3.411	1148	2	2	0	2	0	6	5.226	1.742	1.742	0.000	1.742
1365	JAM	S1A-036	B	M	17-25	1	JAM	3	SEC	SLM	ENG+	3.411	904	0	1	0	0	0	1	1.106	0.000	1.106	0.000	0.000
1366	JAM	S1A-037	A	M	17-25	1	JAM	3	UNI	EDU	ENG+	9.082	1677	2	13	0	2	2	19	11.330	1.193	7.752	0.000	1.193
1367	JAM	S1A-037	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	9.082	415	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1368	JAM	S1A-038	A	F	NA	NA	JAM	3	NA	NA	NA	0.000	485	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1369	JAM	S1A-038	B	F	46+	4	JAM	3	UNI	NA	ENG+	0.000	1507	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1370	JAM	S1A-039	A	F	46+	4	JAM	3	UNI	NA	ENG+	0.467	553	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1371	JAM	S1A-039	B	F	NA	NA	JAM	3	NA	NA	NA	0.467	1590	1	0	0	0	0	1	0.629	0.629	0.000	0.000	0.000
1372	JAM	S1A-040	A	F	NA	NA	JAM	3	NA	NA	NA	0.000	463	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1373	JAM	S1A-040	B	F	46+	4	JAM	3	UNI	NA	ENG+	0.000	1610	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1374	JAM	S1A-041	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.000	1552	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1375	JAM	S1A-041	B	F	NA	NA	JAM	3	NA	NA	NA	0.000	494	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1376	JAM	S1A-042	A	F	NA	NA	JAM	3	NA	NA	NA	1.927	277	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1377	JAM	S1A-042	B	M	17-25	1	JAM	3	UNI	NA	PAT	1.927	1799	2	2	0	0	0	4	2.223	1.112	1.112	0.000	0.000
1378	JAM	S1A-043	A	F	17-25	1	JAM	3	SEC	EDU	ENG+	0.000	134	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1379	JAM	S1A-043	B	F	46+	4	JAM	3	SEC	CLM	ENG	0.000	1949	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1380	JAM	S1A-044	A	F	17-25	1	JAM	3	SEC	EDU	ENG+	0.976	177	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1381	JAM	S1A-044	B	M	17-25	1	JAM	3	UNI	EDU	ENG+	0.976	1872	2	0	0	0	0	2	1.068	1.068	0.000	0.000	0.000
1382	JAM	S1A-045	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	9.606	692	0	3	0	1	0	4	5.780	0.000	4.335	0.000	1.445

1383	JAM	S1A-045	B	F	26-45	3	JAM	3	SEC	SLM	ENG+	9.606	1390	4	9	0	3	0	16	11.511	2.878	6.475	0.000	2.158
1384	JAM	S1A-046	A	F	17-25	1	JAM	3	SEC	EDU	ENG+	4.407	495	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1385	JAM	S1A-046	B	F	26-45	3	JAM	3	UNI	ACD	ENG+	4.407	1547	2	2	0	5	0	9	5.818	1.293	1.293	0.000	3.232
1386	JAM	S1A-047	A	F	26-45	3	JAM	3	UNI	ACD	ENG+	2.420	1835	1	1	0	3	0	5	2.725	0.545	0.545	0.000	1.635
1387	JAM	S1A-047	B	F	17-25	1	JAM	3	SEC	EDU	ENG+	2.420	231	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1388	JAM	S1A-048	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.000	482	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1389	JAM	S1A-048	B	M	17-25	1	JAM	3	UNI	CLM	ENG+	0.000	1617	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1390	JAM	S1A-049	A	M	17-25	1	JAM	3	UNI	CLM	ENG+	0.000	2006	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1392	JAM	S1A-050	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.468	1004	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1393	JAM	S1A-050	B	M	26-45	3	JAM	3	SEC	CLM	ENG+	0.468	1134	1	0	0	0	0	1	0.882	0.882	0.000	0.000	0.000
1394	JAM	S1A-051	A	F	17-25	1	JAM	3	UNI	EDU	PAT	7.089	1521	2	9	0	1	0	12	7.890	1.315	5.917	0.000	0.657
1395	JAM	S1A-051	B	F	NA	NA	JAM	3	NA	NA	NA	7.089	595	0	3	0	0	0	3	5.042	0.000	5.042	0.000	0.000
1396	JAM	S1A-052	A	F	17-25	1	JAM	3	UNI	EDU	PAT	4.335	1483	1	4	0	2	0	7	4.720	0.674	2.697	0.000	1.349
1397	JAM	S1A-052	B	F	NA	NA	JAM	3	NA	NA	NA	4.335	824	0	2	0	1	0	3	3.641	0.000	2.427	0.000	1.214
1399	JAM	S1A-053	B	F	26-45	3	JAM	3	SEC	SLM	ENG+	0.587	1703	0	1	0	0	0	1	0.587	0.000	0.587	0.000	0.000
1400	JAM	S1A-054	A	F	17-25	1	JAM	3	UNI	EDU	PAT	4.200	542	0	2	0	2	0	4	7.380	0.000	3.690	0.000	3.690
1401	JAM	S1A-054	B	F	26-45	3	JAM	3	SEC	SLM	ENG+	4.200	1601	3	1	0	0	1	5	3.123	1.874	0.625	0.000	0.000
1402	JAM	S1A-055	A	M	NA	NA	JAM	3	NA	NA	ENG+	6.064	650	0	1	0	0	0	1	1.538	0.000	1.538	0.000	0.000
1403	JAM	S1A-055	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	6.064	1329	4	5	0	2	0	11	8.277	3.010	3.762	0.000	1.505
1405	JAM	S1A-056	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	8.454	1152	2	9	0	1	0	12	10.417	1.736	7.813	0.000	0.868
1406	JAM	S1A-056	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	8.454	859	1	3	0	1	0	5	5.821	1.164	3.492	0.000	1.164
1408	JAM	S1A-057	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	1.867	914	1	2	0	1	0	4	4.376	1.094	2.188	0.000	1.094
1409	JAM	S1A-057	B	F	17-25	1	JAM	3	UNI	ACD	PAT	1.867	1228	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1410	JAM	S1A-058	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.996	298	0	1	0	0	0	1	3.356	0.000	3.356	0.000	0.000
1411	JAM	S1A-058	B	F	17-25	1	JAM	3	UNI	CLM	ENG+	0.996	1710	0	1	0	0	0	1	0.585	0.000	0.585	0.000	0.000

1413	JAM	S1A-059	A	F	17-25	1	JAM	3	UNI	CLM	ENG+	0.000	2039	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1414	JAM	S1A-059	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.000	131	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1415	JAM	S1A-060	A	F	17-25	1	JAM	3	UNI	CLM	ENG+	0.482	1865	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1416	JAM	S1A-060	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.482	210	0	1	0	0	0	1	4.762	0.000	4.762	0.000	0.000
1417	JAM	S1A-061	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	2.765	1058	1	0	0	0	0	1	0.945	0.945	0.000	0.000	0.000
1418	JAM	S1A-061	B	M	26-45	3	JAM	3	SEC	SLM	ENG+	2.765	1112	3	1	1	0	0	5	4.496	2.698	0.899	0.899	0.000
1419	JAM	S1A-062	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	9.434	177	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1421	JAM	S1A-062	C	F	NA	NA	JAM	3	NA	NA	NA	9.434	671	2	2	0	3	1	8	11.923	2.981	2.981	0.000	4.471
1423	JAM	S1A-063	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	13.131	534	1	1	0	1	0	3	5.618	1.873	1.873	0.000	1.873
1424	JAM	S1A-063	B	F	17-25	1	JAM	3	SEC	EDU	ENG+	13.131	1251	2	12	0	5	0	19	15.188	1.599	9.592	0.000	3.997
1425	JAM	S1A-063	C	F	NA	NA	JAM	3	NA	NA	NA	13.131	196	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1428	JAM	S1A-064	A	F	17-25	1	JAM	3	UNI	ACD	ENG+	1.455	1679	2	1	0	0	0	3	1.787	1.191	0.596	0.000	0.000
1429	JAM	S1A-064	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	1.455	383	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1430	JAM	S1A-065	A	F	17-25	1	JAM	3	UNI	ACD	ENG+	1.962	1792	1	2	0	0	0	3	1.674	0.558	1.116	0.000	0.000
1431	JAM	S1A-065	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	1.962	247	1	0	0	0	0	1	4.049	4.049	0.000	0.000	0.000
1432	JAM	S1A-066	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	2.041	1470	1	2	0	0	0	3	2.041	0.680	1.361	0.000	0.000
1437	JAM	S1A-067	A	F	17-25	1	JAM	3	UNI	NA	ENG+	7.540	658	3	3	0	1	0	7	10.638	4.559	4.559	0.000	1.520
1438	JAM	S1A-067	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	7.540	1464	3	4	0	2	0	9	6.148	2.049	2.732	0.000	1.366
1439	JAM	S1A-068	A	F	26-45	3	JAM	3	UNI	ACD	ENG+	0.981	1758	0	2	0	0	0	2	1.138	0.000	1.138	0.000	0.000
1440	JAM	S1A-068	B	F	17-25	1	JAM	3	SEC	EDU	ENG+	0.981	281	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1441	JAM	S1A-069	A	F	17-25	1	JAM	3	SEC	EDU	ENG+	6.286	319	0	1	0	0	0	1	3.135	0.000	3.135	0.000	0.000
1442	JAM	S1A-069	B	M	26-45	3	JAM	3	UNI	ACD	ENG+	6.286	1749	2	5	0	5	0	12	6.861	1.144	2.859	0.000	2.859
1443	JAM	S1A-070	A	F	26-45	3	JAM	3	UNI	ACD	ENG+	2.370	1555	3	1	0	1	0	5	3.215	1.929	0.643	0.000	0.643
1444	JAM	S1A-070	B	F	17-25	1	JAM	3	SEC	EDU	ENG+	2.370	555	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1445	JAM	S1A-071	A	M	26-45	3	JAM	3	UNI	ACD	ENG+	1.933	1307	2	2	0	0	0	4	3.060	1.530	1.530	0.000	0.000

1446	JAM	S1A-071	B	F	17-25	1	JAM	3	SEC	EDU	ENG+	1.933	762	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1447	JAM	S1A-072	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.000	620	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1448	JAM	S1A-072	B	M	17-25	1	JAM	3	SEC	EDU	ENG+	0.000	1447	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1449	JAM	S1A-073	A	M	17-25	1	JAM	3	SEC	EDU	ENG+	1.432	1252	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1450	JAM	S1A-073	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	1.432	843	0	2	0	1	0	3	3.559	0.000	2.372	0.000	1.186
1451	JAM	S1A-074	A	F	26-45	3	JAM	3	SEC	SLM	ENG+	1.892	1219	3	1	0	0	0	4	3.281	2.461	0.820	0.000	0.000
1452	JAM	S1A-074	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	1.892	895	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1453	JAM	S1A-075	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	6.513	943	0	5	0	0	0	5	5.302	0.000	5.302	0.000	0.000
1454	JAM	S1A-075	B	F	26-45	3	JAM	3	SEC	SLM	ENG+	6.513	1360	2	4	0	4	0	10	7.353	1.471	2.941	0.000	2.941
1455	JAM	S1A-076-1	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.000	311	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1456	JAM	S1A-076-1	B	M	17-25	1	JAM	3	SEC	EDU	ENG+	0.000	747	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1457	JAM	S1A-076-2	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	5.056	330	0	1	0	1	0	2	6.061	0.000	3.030	0.000	3.030
1458	JAM	S1A-076-2	B	F	26-45	3	JAM	3	SEC	SLM	ENG+	5.056	659	0	3	0	0	0	3	4.552	0.000	4.552	0.000	0.000
1459	JAM	S1A-077	A	F	17-25	1	JAM	3	SEC	EDU	ENG+	1.981	586	0	3	0	1	0	4	6.826	0.000	5.119	0.000	1.706
1460	JAM	S1A-077	B	M	17-25	1	JAM	3	UNI	ACD	ENG+	1.981	1433	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1461	JAM	S1A-078	A	F	NA	NA	JAM	3	NA	NA	NA	1.390	473	0	0	0	1	0	1	2.114	0.000	0.000	0.000	2.114
1462	JAM	S1A-078	B	F	17-25	1	JAM	3	SEC	EDU	ENG+	1.390	1686	0	2	0	0	0	2	1.186	0.000	1.186	0.000	0.000
1463	JAM	S1A-079-1	A	F	17-25	1	JAM	3	SEC	EDU	ENG+	0.000	469	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1464	JAM	S1A-079-1	B	M	17-25	1	JAM	3	UNI	ACD	ENG+	0.000	1090	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1465	JAM	S1A-079-2	A	F	NA	NA	JAM	3	NA	NA	NA	0.000	231	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1466	JAM	S1A-079-2	B	F	17-25	1	JAM	3	SEC	EDU	ENG+	0.000	632	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1467	JAM	S1A-080	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.974	1240	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1468	JAM	S1A-080	B	F	17-25	1	JAM	3	SEC	EDU	ENG+	0.974	814	0	2	0	0	0	2	2.457	0.000	2.457	0.000	0.000
1469	JAM	S1A-081	A	F	26-45	3	JAM	3	UNI	ACD	ENG+	0.000	458	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1470	JAM	S1A-081	B	F	17-25	1	JAM	3	SEC	EDU	ENG+	0.000	1623	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1471	JAM	S1A-082	A	F	26-45	3	JAM	3	UNI	ACD	ENG+	0.479	481	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1472	JAM	S1A-082	B	F	17-25	1	JAM	3	SEC	EDU	ENG+	0.479	1608	0	1	0	0	0	1	0.622	0.000	0.622	0.000	0.000
1473	JAM	S1A-083-1	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	1.524	940	0	1	0	0	0	1	1.064	0.000	1.064	0.000	0.000
1474	JAM	S1A-083-1	B	F	17-25	1	JAM	3	SEC	EDU	ENG+	1.524	372	0	1	0	0	0	1	2.688	0.000	2.688	0.000	0.000
1475	JAM	S1A-083-2	A	F	17-25	1	JAM	3	SEC	EDU	ENG+	2.283	425	0	1	0	0	0	1	2.353	0.000	2.353	0.000	0.000
1476	JAM	S1A-083-2	B	F	26-45	3	JAM	3	UNI	ACD	ENG+	2.283	451	0	1	0	0	0	1	2.217	0.000	2.217	0.000	0.000
1477	JAM	S1A-084	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	2.964	451	0	1	0	1	0	2	4.435	0.000	2.217	0.000	2.217
1478	JAM	S1A-084	B	F	26-45	3	JAM	3	UNI	ACD	PAT	2.964	1573	2	2	0	0	0	4	2.543	1.271	1.271	0.000	0.000
1479	JAM	S1A-085	A	F	NA	NA	JAM	3	NA	NA	ENG+	2.437	897	1	3	0	1	0	5	5.574	1.115	3.344	0.000	1.115
1480	JAM	S1A-085	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	2.437	1155	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1481	JAM	S1A-086-1	A	F	26-45	3	JAM	3	UNI	ACD	PAT	1.862	1291	0	1	0	1	0	2	1.549	0.000	0.775	0.000	0.775
1482	JAM	S1A-086-1	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	1.862	320	0	1	0	0	0	1	3.125	0.000	3.125	0.000	0.000
1483	JAM	S1A-086-2	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.000	286	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1485	JAM	S1A-087	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.956	281	0	1	0	0	0	1	3.559	0.000	3.559	0.000	0.000
1486	JAM	S1A-087	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.956	1812	1	0	0	0	0	1	0.552	0.552	0.000	0.000	0.000
1487	JAM	S1A-088	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.485	1726	0	1	0	0	0	1	0.579	0.000	0.579	0.000	0.000
1488	JAM	S1A-088	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.485	335	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1489	JAM	S1A-089	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.473	733	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1490	JAM	S1A-089	B	M	26-45	3	JAM	3	UNI	ACD	ENG+	0.473	1381	0	1	0	0	0	1	0.724	0.000	0.724	0.000	0.000
1491	JAM	S1A-090-1	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.000	776	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1492	JAM	S1A-090-1	B	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.000	234	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1493	JAM	S1A-090-2	A	F	17-25	1	JAM	3	UNI	EDU	ENG+	0.000	374	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1494	JAM	S1A-090-2	B	M	26-45	3	JAM	3	UNI	ACD	ENG+	0.000	866	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1495	JAM	S1A-091	A	M	26-45	3	JAM	1	UNI	CLM	NA	0.491	1408	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1496	JAM	S1A-091	B	F	26-45	3	JAM	1	SEC	NA	NA	0.491	627	0	1	0	0	0	1	1.595	0.000	1.595	0.000	0.000

1497	JAM	S1A-092	A	M	NA	NA	JAM	1	SEC	CLM	NA	0.000	956	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1498	JAM	S1A-092	B	M	NA	NA	JAM	1	SEC	NA	NA	0.000	1152	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1499	JAM	S1A-093-1	A	M	NA	NA	JAM	1	SEC	CLM	NA	1.065	1110	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1500	JAM	S1A-093-1	B	F	NA	NA	JAM	2	SEC	NA	NA	1.065	768	1	1	0	0	0	2	2.604	1.302	1.302	0.000	0.000
1501	JAM	S1A-093-2	A	M	NA	NA	JAM	1	SEC	CLM	NA	0.000	147	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1502	JAM	S1A-093-2	C	M	NA	NA	JAM	1	SEC	NA	NA	0.000	114	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1503	JAM	S1A-094	A	M	NA	NA	JAM	1	SEC	CLM	NA	0.000	738	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1504	JAM	S1A-094	B	M	NA	NA	JAM	1	SEC	NA	NA	0.000	1392	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1505	JAM	S1A-095-1	A	F	46+	4	JAM	1	UNI	CLM	NA	0.000	404	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1506	JAM	S1A-095-1	B	M	NA	NA	JAM	1	SEC	NA	NA	0.000	177	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1507	JAM	S1A-095-2	A	F	46+	4	JAM	1	UNI	CLM	NA	0.000	819	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1508	JAM	S1A-095-2	C	M	NA	NA	JAM	1	SEC	NA	NA	0.000	353	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1509	JAM	S1A-095-3	A	F	46+	4	JAM	1	UNI	CLM	NA	0.000	195	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1510	JAM	S1A-095-3	D	M	NA	NA	JAM	1	SEC	NA	NA	0.000	112	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1511	JAM	S1A-096	A	F	46+	4	JAM	1	UNI	CLM	NA	0.484	1586	0	1	0	0	0	1	0.631	0.000	0.631	0.000	0.000
1512	JAM	S1A-096	B	M	NA	NA	JAM	1	SEC	NA	NA	0.484	478	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1513	JAM	S1A-097-1	A	F	NA	NA	JAM	1	UNI	ACD	NA	0.799	474	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1514	JAM	S1A-097-1	B	F	NA	NA	JAM	1	SEC	NA	NA	0.799	777	0	1	0	0	0	1	1.287	0.000	1.287	0.000	0.000
1515	JAM	S1A-097-2	A	F	NA	NA	JAM	1	UNI	ACD	NA	0.000	438	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1516	JAM	S1A-097-2	C	F	17-25	1	JAM	1	SEC	NA	NA	0.000	438	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1517	JAM	S1A-098	A	F	NA	NA	JAM	1	UNI	ACD	NA	0.000	870	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1518	JAM	S1A-098	B	F	NA	NA	JAM	1	SEC	ACD	NA	0.000	1200	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1519	JAM	S1A-099-1	A	F	NA	NA	JAM	1	UNI	ACD	NA	0.000	914	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1520	JAM	S1A-099-1	B	F	NA	NA	JAM	1	SEC	NA	NA	0.000	210	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1521	JAM	S1A-099-2	A	F	NA	NA	JAM	1	UNI	ACD	NA	0.000	672	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1522	JAM	S1A-099-2	C	F	NA	NA	JAM	1	SEC	NA	NA	0.000	267	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1523	JAM	S1A-100	A	M	NA	NA	JAM	2	SEC	SLM	NA	3.049	1606	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1524	JAM	S1A-100	B	M	NA	NA	JAM	2	SEC	NA	NA	3.049	154	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1525	JAM	S1A-100	C	F	NA	NA	JAM	2	SEC	NA	NA	3.049	171	0	0	1	0	0	1	5.848	0.000	0.000	5.848	0.000
1526	JAM	S1A-100	D	M	NA	NA	JAM	2	SEC	NA	NA	3.049	157	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

Table 144: Final data set of NZE.

ID	VAR	FILE	SPEAKER	SEX	age.orig	AGE	ETH	DATE	EDU	OCC	PAI	words	ini.n	med.n	fin.m	non.n	na.n	all.n	all	ini	med	fin	non
1563	NZ	S1A001	F	F	20-29	1	OTH	NA	SEC	SLM	0.98	1052	0	0	0	2	0	2	1.901	0.000	0.000	0.000	1.901
1670	NZ	S1A001	M	M	20-29	1	OTH	NA	NA	EDU	0.98	989	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1530	NZ	S1A002	B	F	16-19	1	OTH	NA	UNI	SLM	4.436	1037	3	2	1	2	0	8	7.715	2.893	1.929	0.964	1.929
1529	NZ	S1A002	Q	F	16-19	1	OTH	NA	NA	EDU	4.436	992	0	0	0	1	0	1	1.008	0.000	0.000	0.000	1.008
1759	NZ	S1A003	B	M	30-39	3	OTH	NA	UNI	CLM	0.446	520	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1711	NZ	S1A003	M	M	20-29	1	OTH	NA	UNI	ACD	0.446	1722	0	1	0	0	0	1	0.581	0.000	0.581	0.000	0.000
1632	NZ	S1A004	G	F	20-29	1	MAO	NA	NA	EDU	0.485	1403	0	0	1	0	0	1	0.713	0.000	0.000	0.713	0.000
1610	NZ	S1A004	M	F	40+	4	MAO	NA	NA	EDU	0.485	659	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1536	NZ	S1A005	A	M	40+	4	OTH	NA	NA	EDU	4.772	687	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1535	NZ	S1A005	B	F	16-19	1	OTH	NA	UNI	CLM	4.772	843	3	2	1	1	0	7	8.304	3.559	2.372	1.186	1.186
1537	NZ	S1A005	C	F	16-19	1	OTH	NA	NA	EDU	4.772	624	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1588	NZ	S1A006	F	F	20-29	1	OTH	NA	NA	EDU	1.902	494	0	0	0	1	0	1	2.024	0.000	0.000	0.000	2.024
1607	NZ	S1A006	G	M	20-29	1	OTH	NA	UNI	CLM	1.902	447	0	1	0	0	0	1	2.237	0.000	2.237	0.000	0.000
1531	NZ	S1A006	M	M	20-29	1	OTH	NA	UNI	CLM	1.902	1130	1	1	0	0	0	2	1.770	0.885	0.885	0.000	0.000
1541	NZ	S1A007	G	M	40+	4	MAO	NA	SEC	NA	1.392	621	1	0	0	0	0	1	1.610	1.610	0.000	0.000	0.000
1653	NZ	S1A007	J	F	40+	4	OTH	NA	NA	NA	1.392	384	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1543	NZ	S1A007	T	F	16-19	1	OTH	NA	UNI	SLM	1.392	1053	1	0	0	1	0	2	1.899	0.950	0.000	0.000	0.950
1674	NZ	S1A008	G	M	20-29	1	OTH	NA	UNI	CLM	0.953	237	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1720	NZ	S1A008	J	M	20-29	1	OTH	NA	NA	EDU	0.953	822	1	0	0	0	0	1	1.217	1.217	0.000	0.000	0.000
1660	NZ	S1A008	L	F	20-29	1	OTH	NA	NA	NA	0.953	240	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1645	NZ	S1A008	R	M	20-29	1	OTH	NA	UNI	ACD	0.953	809	0	0	1	0	0	1	1.236	0.000	0.000	1.236	0.000
1728	NZ	S1A009	H	M	16-19	1	OTH	NA	UNI	EDU	4.271	1208	1	3	0	3	0	7	5.795	0.828	2.483	0.000	2.483
1601	NZ	S1A009	S	M	16-19	1	OTH	NA	UNI	EDU	4.271	899	0	1	1	0	0	2	2.225	0.000	1.112	1.112	0.000

1689	NZ	S1A010	D	M	20-29	1	OTH	NA	UNI	EDU	3.316	1259	1	2	0	0	0	3	2.383	0.794	1.589	0.000	0.000
1552	NZ	S1A010	H	F	16-19	1	OTH	NA	NA	CLM	3.316	852	1	0	1	2	0	4	4.695	1.174	0.000	1.174	2.347
1709	NZ	S1A011	B	F	30-39	3	MAO	NA	SEC	CLM	2.913	1328	1	2	0	0	0	3	2.259	0.753	1.506	0.000	0.000
1619	NZ	S1A011	U	F	40+	4	MAO	NA	NA	NA	2.913	732	0	2	0	1	0	3	4.098	0.000	2.732	0.000	1.366
1555	NZ	S1A012	A	M	40+	4	OTH	NA	UNI	ACD	3.145	696	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1714	NZ	S1A012	B	F	20-29	1	OTH	NA	UNI	EDU	3.145	620	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1574	NZ	S1A012	C	F	20-29	1	OTH	NA	SEC	CLM	3.145	970	2	2	0	1	0	5	5.155	2.062	2.062	0.000	1.031
1595	NZ	S1A013	E	F	20-29	1	OTH	NA	UNI	EDU	0.422	1485	0	0	0	1	0	1	0.673	0.000	0.000	0.000	0.673
1749	NZ	S1A013	H	F	20-29	1	OTH	NA	UNI	EDU	0.422	882	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1651	NZ	S1A014	A	F	20-29	1	OTH	NA	NA	EDU	5.426	1323	2	0	1	1	0	4	3.023	1.512	0.000	0.756	0.756
1560	NZ	S1A014	F	F	16-19	1	OTH	NA	UNI	SLM	5.426	1073	4	3	0	1	1	9	8.388	3.728	2.796	0.000	0.932
1533	NZ	S1A015	A	M	20-29	1	OTH	NA	NA	EDU	7.612	1424	6	4	1	4	0	15	10.534	4.213	2.809	0.702	2.809
1631	NZ	S1A015	S	F	20-29	1	OTH	NA	UNI	EDU	7.612	678	0	1	0	0	0	1	1.475	0.000	1.475	0.000	0.000
1762	NZ	S1A016	I	M	20-29	1	OTH	NA	NA	NA	4.358	1168	2	2	0	0	0	4	3.425	1.712	1.712	0.000	0.000
1764	NZ	S1A016	R	M	20-29	1	OTH	NA	UNI	SLM	4.358	897	3	1	0	1	0	5	5.574	3.344	1.115	0.000	1.115
1566	NZ	S1A017	A	F	16-19	1	OTH	NA	SEC	CLM	1.841	1738	2	1	0	0	0	3	1.726	1.151	0.575	0.000	0.000
1567	NZ	S1A017	B	F	16-19	1	OTH	NA	SEC	SLM	1.841	978	0	1	0	1	0	2	2.045	0.000	1.022	0.000	1.022
1562	NZ	S1A018	A	F	40+	4	OTH	NA	UNI	ACD	0	1137	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1532	NZ	S1A018	L	F	40+	4	OTH	NA	NA	ACD	0	1047	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1582	NZ	S1A019	A	F	30-39	3	OTH	NA	UNI	EDU	0	434	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1738	NZ	S1A019	B	F	40+	4	OTH	NA	NA	EDU	0	1661	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1726	NZ	S1A020	A	M	20-29	1	OTH	NA	UNI	NA	0.922	1126	1	0	0	0	0	1	0.888	0.888	0.000	0.000	0.000
1682	NZ	S1A020	B	M	20-29	1	OTH	NA	SEC	CLM	0.922	1044	0	1	0	0	0	1	0.958	0.000	0.958	0.000	0.000
1586	NZ	S1A021	A	F	20-29	1	OTH	NA	UNI	EDU	1.438	1474	1	0	0	1	0	2	1.357	0.678	0.000	0.000	0.678
1575	NZ	S1A021	B	F	16-19	1	OTH	NA	UNI	EDU	1.438	612	0	1	0	0	0	1	1.634	0.000	1.634	0.000	0.000

1577	NZ	S1A022	B	F	16-19	1	OTH	NA	NA	CLM	3.976	1379	3	0	1	0	0	4	2.901	2.175	0.000	0.725	0.000
1576	NZ	S1A022	H	F	16-19	1	OTH	NA	NA	EDU	3.976	633	1	2	0	1	0	4	6.319	1.580	3.160	0.000	1.580
1564	NZ	S1A023	D	F	40+	4	OTH	NA	NA	NA	0	784	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1573	NZ	S1A023	M	F	40+	4	OTH	NA	UNI	CLM	0	1262	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1740	NZ	S1A024	F	F	20-29	1	OTH	NA	UNI	CLM	4.348	625	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1542	NZ	S1A024	M	M	30-39	3	OTH	NA	UNI	CLM	4.348	1445	1	5	0	2	1	9	6.228	0.692	3.460	0.000	1.384
1729	NZ	S1A025	M	F	30-39	3	MAO	NA	NA	ACD	0.482	1079	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1634	NZ	S1A025	P	F	40+	4	OTH	NA	SEC	SLM	0.482	994	0	0	0	1	0	1	1.006	0.000	0.000	0.000	1.006
1637	NZ	S1A026	A	F	20-29	1	OTH	NA	NA	EDU	1.862	672	1	0	0	0	0	1	1.488	1.488	0.000	0.000	0.000
1710	NZ	S1A026	T	M	20-29	1	OTH	NA	NA	EDU	1.862	1476	0	3	0	0	0	3	2.033	0.000	2.033	0.000	0.000
1626	NZ	S1A027	L	F	20-29	1	OTH	NA	UNI	SLM	0	952	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1589	NZ	S1A027	M	M	40+	4	OTH	NA	NA	ACD	0	1093	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1565	NZ	S1A027	N	F	40+	4	OTH	NA	UNI	ACD	0	453	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1633	NZ	S1A028	A	F	40+	4	OTH	NA	UNI	CLM	0	1355	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1695	NZ	S1A028	B	F	40+	4	OTH	NA	SEC	SLM	0	1215	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1627	NZ	S1A029	C	F	20-29	1	OTH	NA	NA	EDU	0	995	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1746	NZ	S1A029	D	M	20-29	1	OTH	NA	SEC	SLM	0	1035	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1697	NZ	S1A030	E	F	30-39	3	MAO	NA	NA	CLM	2.74	922	1	0	0	0	0	1	1.085	1.085	0.000	0.000	0.000
1623	NZ	S1A030	K	F	20-29	1	NA	NA	NA	EDU	2.74	1268	3	2	0	0	0	5	3.943	2.366	1.577	0.000	0.000
1667	NZ	S1A031	A	F	20-29	1	OTH	NA	NA	EDU	2.492	1150	0	2	1	2	0	5	4.348	0.000	1.739	0.870	1.739
1649	NZ	S1A031	B	F	20-29	1	OTH	NA	NA	EDU	2.492	1258	1	0	0	0	0	1	0.795	0.795	0.000	0.000	0.000
1600	NZ	S1A032	D	M	40+	4	OTH	NA	NA	EDU	0.975	1215	1	0	0	0	0	1	0.823	0.823	0.000	0.000	0.000
1534	NZ	S1A032	S	F	30-39	3	OTH	NA	SEC	SLM	0.975	837	1	0	0	0	0	1	1.195	1.195	0.000	0.000	0.000
1696	NZ	S1A033	A	M	16-19	1	OTH	NA	SEC	SLM	4.398	1338	0	3	0	1	0	4	2.990	0.000	2.242	0.000	0.747
1602	NZ	S1A033	B	M	40+	4	OTH	NA	UNI	EDU	4.398	1163	2	4	1	0	0	7	6.019	1.720	3.439	0.860	0.000

1605	NZ	S1A034	F	F	16-19	1	OTH	NA	NA	EDU	8.238	718	2	0	0	0	0	2	2.786	2.786	0.000	0.000	0.000
1583	NZ	S1A034	M	M	16-19	1	OTH	NA	UNI	CLM	8.238	1467	5	8	0	3	0	16	10.907	3.408	5.453	0.000	2.045
1592	NZ	S1A035	D	M	30-39	3	OTH	NA	UNI	CLM	0.447	646	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1628	NZ	S1A035	R	M	30-39	3	OTH	NA	NA	CLM	0.447	1591	0	0	0	1	0	1	0.629	0.000	0.000	0.000	0.629
1597	NZ	S1A036	A	F	20-29	1	OTH	NA	SEC	CLM	6.277	1756	5	5	1	2	0	13	7.403	2.847	2.847	0.569	1.139
1621	NZ	S1A036	B	F	20-29	1	OTH	NA	UNI	SLM	6.277	1271	2	3	0	1	0	6	4.721	1.574	2.360	0.000	0.787
1677	NZ	S1A037	J	M	20-29	1	OTH	NA	SEC	CLM	0.986	862	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1668	NZ	S1A037	M	M	20-29	1	OTH	NA	UNI	CLM	0.986	1166	0	2	0	0	0	2	1.715	0.000	1.715	0.000	0.000
1745	NZ	S1A038	M	M	20-29	1	NA	NA	SEC	SLM	0.632	1270	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1675	NZ	S1A038	P	M	20-29	1	OTH	NA	SEC	CLM	0.632	1893	1	0	0	1	0	2	1.057	0.528	0.000	0.000	0.528
1614	NZ	S1A039	Q	F	16-19	1	OTH	NA	NA	EDU	4.418	1181	2	3	0	1	0	6	5.080	1.693	2.540	0.000	0.847
1615	NZ	S1A039	T	F	16-19	1	OTH	NA	NA	EDU	4.418	856	2	0	0	1	0	3	3.505	2.336	0.000	0.000	1.168
1616	NZ	S1A040	A	F	16-19	1	OTH	NA	UNI	SLM	9.126	1197	3	3	0	0	0	6	5.013	2.506	2.506	0.000	0.000
1617	NZ	S1A040	V	F	16-19	1	OTH	NA	UNI	SLM	9.126	885	8	4	0	1	0	13	14.689	9.040	4.520	0.000	1.130
1671	NZ	S1A041	J	M	20-29	1	NA	NA	NA	EDU	7.201	1733	7	3	0	4	0	14	8.078	4.039	1.731	0.000	2.308
1650	NZ	S1A041	V	F	20-29	1	NA	NA	NA	EDU	7.201	350	1	0	0	0	0	1	2.857	2.857	0.000	0.000	0.000
1692	NZ	S1A042	B	F	20-29	1	OTH	NA	NA	EDU	7.433	1125	3	1	0	0	0	4	3.556	2.667	0.889	0.000	0.000
1739	NZ	S1A042	J	F	20-29	1	NA	NA	UNI	SLM	7.433	893	1	8	0	2	0	11	12.318	1.120	8.959	0.000	2.240
1599	NZ	S1A043	L	F	20-29	1	OTH	NA	UNI	CLM	2.837	986	0	2	0	0	0	2	2.028	0.000	2.028	0.000	0.000
1622	NZ	S1A043	S	F	16-19	1	OTH	NA	UNI	EDU	2.837	1129	1	1	0	2	0	4	3.543	0.886	0.886	0.000	1.771
1743	NZ	S1A044	J	M	16-19	1	OTH	NA	UNI	SLM	2.098	624	0	0	0	1	0	1	1.603	0.000	0.000	0.000	1.603
1625	NZ	S1A044	K	F	16-19	1	OTH	NA	NA	NA	2.098	681	1	0	0	1	0	2	2.937	1.468	0.000	0.000	1.468
1596	NZ	S1A044	S	M	16-19	1	OTH	NA	NA	NA	2.098	749	0	1	0	0	0	1	1.335	0.000	1.335	0.000	0.000
1585	NZ	S1A045	B	F	40+	4	OTH	NA	SEC	SLM	1.351	561	1	0	0	0	0	1	1.783	1.783	0.000	0.000	0.000
1629	NZ	S1A045	F	F	16-19	1	OTH	NA	NA	EDU	1.351	367	2	0	0	0	0	2	5.450	5.450	0.000	0.000	0.000

1630	NZ	S1A045	M	M	40+	4	OTH	NA	UNI	ACD	1.351	1113	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1550	NZ	S1A046	M	F	20-29	1	OTH	NA	SEC	SLM	1.955	925	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1658	NZ	S1A046	V	F	20-29	1	OTH	NA	UNI	CLM	1.955	2655	4	2	0	1	0	7	2.637	1.507	0.753	0.000	0.377
1744	NZ	S1A047	A	M	20-29	1	OTH	NA	UNI	SLM	5.076	654	5	1	0	2	0	8	12.232	7.645	1.529	0.000	3.058
1584	NZ	S1A047	B	M	20-29	1	OTH	NA	UNI	CLM	5.076	975	0	2	0	3	0	5	5.128	0.000	2.051	0.000	3.077
1646	NZ	S1A047	C	M	20-29	1	OTH	NA	NA	EDU	5.076	329	0	1	0	0	0	1	3.040	0.000	3.040	0.000	0.000
1606	NZ	S1A047	D	M	20-29	1	OTH	NA	NA	EDU	5.076	262	0	1	1	0	0	2	7.634	0.000	3.817	3.817	0.000
1741	NZ	S1A048	A	F	20-29	1	OTH	NA	SEC	SLM	0.449	1209	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1638	NZ	S1A048	B	F	16-19	1	OTH	NA	NA	EDU	0.449	1017	1	0	0	0	0	1	0.983	0.983	0.000	0.000	0.000
1639	NZ	S1A049	M	M	40+	4	OTH	NA	SEC	CLM	0.816	1642	1	1	0	0	0	2	1.218	0.609	0.609	0.000	0.000
1761	NZ	S1A049	T	M	20-29	1	OTH	NA	UNI	ACD	0.816	808	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1659	NZ	S1A050	A	F	40+	4	OTH	NA	UNI	ACD	1.484	1076	0	0	0	1	0	1	0.929	0.000	0.000	0.000	0.929
1700	NZ	S1A050	B	F	30-39	3	OTH	NA	SEC	CLM	1.484	945	1	0	0	1	0	2	2.116	1.058	0.000	0.000	1.058
1665	NZ	S1A051	A	F	20-29	1	OTH	NA	UNI	SLM	2.344	1057	1	0	0	0	0	1	0.946	0.946	0.000	0.000	0.000
1699	NZ	S1A051	G	F	20-29	1	OTH	NA	NA	EDU	2.344	1076	2	0	0	2	0	4	3.717	1.859	0.000	0.000	1.859
1730	NZ	S1A052	A	M	30-39	3	OTH	NA	UNI	SLM	1.685	407	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1731	NZ	S1A052	B	M	30-39	3	OTH	NA	UNI	SLM	1.685	1368	1	1	0	0	0	2	1.462	0.731	0.731	0.000	0.000
1647	NZ	S1A052	C	M	40+	4	OTH	NA	UNI	SLM	1.685	412	0	1	0	0	0	1	2.427	0.000	2.427	0.000	0.000
1742	NZ	S1A053	C	F	20-29	1	NA	NA	NA	EDU	0.623	674	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1603	NZ	S1A053	K	F	20-29	1	OTH	NA	NA	EDU	0.623	1036	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1688	NZ	S1A053	L	F	20-29	1	OTH	NA	NA	EDU	0.623	569	0	0	0	1	0	1	1.757	0.000	0.000	0.000	1.757
1641	NZ	S1A054	D	M	20-29	1	MAO	NA	NA	EDU	0	1276	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1655	NZ	S1A054	H	F	16-19	1	OTH	NA	UNI	CLM	0	106	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1757	NZ	S1A054	M	M	20-29	1	MAO	NA	NA	SLM	0	288	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1609	NZ	S1A054	P	F	20-29	1	OTH	NA	UNI	SLM	0	313	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1678	NZ	S1A055	A	F	20-29	1	OTH	NA	UNI	CLM	6.376	921	5	1	0	3	0	9	9.772	5.429	1.086	0.000	3.257
1719	NZ	S1A055	F	F	20-29	1	OTH	NA	UNI	ACD	6.376	1118	2	1	0	1	0	4	3.578	1.789	0.894	0.000	0.894
1737	NZ	S1A056	C	M	16-19	1	OTH	NA	NA	NA	0.991	639	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1661	NZ	S1A056	D	M	20-29	1	OTH	NA	SEC	SLM	0.991	1380	1	1	0	0	0	2	1.449	0.725	0.725	0.000	0.000
1608	NZ	S1A057	F	F	20-29	1	NA	NA	SEC	EDU	3.247	1154	1	5	0	0	0	6	5.199	0.867	4.333	0.000	0.000
1763	NZ	S1A057	M	M	30-39	3	OTH	NA	UNI	EDU	3.247	1002	0	1	0	0	0	1	0.998	0.000	0.998	0.000	0.000
1549	NZ	S1A058	F	F	20-29	1	OTH	NA	UNI	CLM	4.866	792	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1694	NZ	S1A058	T	F	20-29	1	OTH	NA	UNI	EDU	4.866	1263	4	2	0	4	0	10	7.918	3.167	1.584	0.000	3.167
1618	NZ	S1A059	B	F	20-29	1	OTH	NA	NA	EDU	6.907	953	0	1	0	1	0	2	2.099	0.000	1.049	0.000	1.049
1557	NZ	S1A059	F	F	20-29	1	OTH	NA	UNI	ACD	6.907	1074	6	1	3	2	0	12	11.173	5.587	0.931	2.793	1.862
1669	NZ	S1A060	K	M	40+	4	MAO	NA	UNI	CLM	0	743	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1569	NZ	S1A060	T	M	30-39	3	MAO	NA	NA	CLM	0	1002	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1666	NZ	S1A060	W	F	30-39	3	OTH	NA	UNI	ACD	0	286	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1539	NZ	S1A061	F	F	40+	4	OTH	NA	UNI	ACD	2.81	891	4	0	0	0	0	4	4.489	4.489	0.000	0.000	0.000
1701	NZ	S1A061	G	F	30-39	3	OTH	NA	SEC	CLM	2.81	1244	1	0	0	1	0	2	1.608	0.804	0.000	0.000	0.804
1707	NZ	S1A062	C	F	40+	4	OTH	NA	UNI	ACD	0	676	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1527	NZ	S1A062	M	F	30-39	3	OTH	NA	UNI	ACD	0	2084	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1676	NZ	S1A063	K	M	40+	4	MAO	NA	UNI	CLM	0	1145	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1578	NZ	S1A063	T	M	30-39	3	MAO	NA	NA	CLM	0	731	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1548	NZ	S1A063	W	F	30-39	3	OTH	NA	UNI	ACD	0	197	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1679	NZ	S1A064	D	F	16-19	1	OTH	NA	SEC	SLM	4.458	1370	2	2	0	1	0	5	3.650	1.460	1.460	0.000	0.730
1748	NZ	S1A064	F	F	20-29	1	OTH	NA	UNI	ACD	4.458	649	1	2	0	1	0	4	6.163	1.541	3.082	0.000	1.541
1680	NZ	S1A065	O	M	40+	4	OTH	NA	UNI	SLM	2.495	1093	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1735	NZ	S1A065	W	M	30-39	3	OTH	NA	UNI	SLM	2.495	911	1	4	0	0	0	5	5.488	1.098	4.391	0.000	0.000
1657	NZ	S1A066	F	F	20-29	1	MAO	NA	UNI	CLM	0.987	876	0	1	0	0	0	1	1.142	0.000	1.142	0.000	0.000

1568	NZ	S1A066	M	M	30-39	3	OTH	NA	SEC	CLM	0.987	1151	0	1	0	0	0	1	0.869	0.000	0.869	0.000	0.000
1662	NZ	S1A067	A	F	30-39	3	OTH	NA	UNI	EDU	2.865	887	3	1	0	1	0	5	5.637	3.382	1.127	0.000	1.127
1703	NZ	S1A067	S	F	40+	4	MAO	NA	NA	CLM	2.865	1207	1	0	0	0	0	1	0.829	0.829	0.000	0.000	0.000
1664	NZ	S1A068	G	F	20-29	1	OTH	NA	UNI	CLM	2.199	394	0	1	0	0	0	1	2.538	0.000	2.538	0.000	0.000
1571	NZ	S1A068	I	M	30-39	3	OTH	NA	NA	NA	2.199	1880	2	2	0	0	0	4	2.128	1.064	1.064	0.000	0.000
1698	NZ	S1A069	H	F	20-29	1	OTH	NA	UNI	CLM	1.649	602	0	1	0	0	0	1	1.661	0.000	1.661	0.000	0.000
1540	NZ	S1A069	J	F	40+	4	OTH	NA	UNI	ACD	1.649	355	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1691	NZ	S1A069	R	M	40+	4	OTH	NA	UNI	ACD	1.649	298	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1750	NZ	S1A069	S	F	20-29	1	OTH	NA	UNI	CLM	1.649	363	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1686	NZ	S1A069	W	M	20-29	1	OTH	NA	UNI	ACD	1.649	850	0	1	0	1	0	2	2.353	0.000	1.176	0.000	1.176
1752	NZ	S1A070	A	F	20-29	1	OTH	NA	UNI	SLM	4.192	1042	1	1	0	2	0	4	3.839	0.960	0.960	0.000	1.919
1685	NZ	S1A070	B	M	20-29	1	OTH	NA	UNI	ACD	4.192	1105	1	1	0	3	0	5	4.525	0.905	0.905	0.000	2.715
1747	NZ	S1A071	D	F	20-29	1	OTH	NA	NA	NA	0.747	1195	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1687	NZ	S1A071	K	F	30-39	3	OTH	NA	NA	NA	0.747	1483	1	0	0	1	0	2	1.349	0.674	0.000	0.000	0.674
1642	NZ	S1A072	B	M	20-29	1	OTH	NA	UNI	SLM	0	295	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1702	NZ	S1A072	C	F	20-29	1	OTH	NA	UNI	ACD	0	690	1	0	0	0	0	1	1.449	1.449	0.000	0.000	0.000
1604	NZ	S1A072	J	F	20-29	1	OTH	NA	UNI	CLM	0	869	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1760	NZ	S1A072	P	M	20-29	1	OTH	NA	UNI	CLM	0	312	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1727	NZ	S1A073	D	M	20-29	1	OTH	NA	UNI	SLM	1.076	1123	3	0	0	0	0	3	2.671	2.671	0.000	0.000	0.000
1545	NZ	S1A073	J	F	20-29	1	OTH	NA	NA	EDU	1.076	334	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1681	NZ	S1A073	P	M	20-29	1	OTH	NA	UNI	CLM	1.076	595	0	1	0	0	0	1	1.681	0.000	1.681	0.000	0.000
1581	NZ	S1A074	A	M	20-29	1	OTH	NA	NA	EDU	5.23	646	0	0	0	1	0	1	1.548	0.000	0.000	0.000	1.548
1559	NZ	S1A074	F	F	20-29	1	OTH	NA	NA	EDU	5.23	299	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1620	NZ	S1A074	J	F	20-29	1	OTH	NA	UNI	CLM	5.23	250	0	1	0	0	0	1	4.000	0.000	4.000	0.000	0.000
1672	NZ	S1A074	L	M	20-29	1	OTH	NA	UNI	ACD	5.23	244	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1579	NZ	S1A074	M	M	20-29	1	OTH	NA	UNI	ACD	5.23	712	2	1	0	2	0	5	7.022	2.809	1.404	0.000	2.809
1704	NZ	S1A075	E	F	40+	4	OTH	NA	UNI	ACD	2.248	1736	2	1	0	1	0	4	2.304	1.152	0.576	0.000	0.576
1715	NZ	S1A075	W	F	40+	4	OTH	NA	UNI	CLM	2.248	933	0	0	0	2	0	2	2.144	0.000	0.000	0.000	2.144
1713	NZ	S1A076	D	M	40+	4	OTH	NA	NA	NA	0	924	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1712	NZ	S1A076	E	M	40+	4	OTH	NA	NA	CLM	0	1283	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1673	NZ	S1A077	A	M	20-29	1	OTH	NA	NA	NA	3.447	1031	0	2	0	0	0	2	1.940	0.000	1.940	0.000	0.000
1756	NZ	S1A077	T	F	20-29	1	OTH	NA	UNI	EDU	3.447	1000	4	1	0	0	0	5	5.000	4.000	1.000	0.000	0.000
1716	NZ	S1A078	K	M	40+	4	OTH	NA	UNI	ACD	0	1380	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1717	NZ	S1A078	M	M	40+	4	OTH	NA	UNI	ACD	0	913	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1734	NZ	S1A079	G	M	20-29	1	NA	NA	UNI	ACD	0.582	417	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1554	NZ	S1A079	P	M	30-39	3	OTH	NA	UNI	ACD	0.582	393	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1708	NZ	S1A079	R	F	20-29	1	MAO	NA	UNI	ACD	0.582	1324	1	0	0	0	0	1	0.755	0.755	0.000	0.000	0.000
1721	NZ	S1A080	N	M	40+	4	MAO	NA	UNI	ACD	0.594	2321	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1722	NZ	S1A080	R	M	40+	4	MAO	NA	NA	NA	0.594	1048	0	1	0	1	0	2	1.908	0.000	0.954	0.000	0.954
1725	NZ	S1A081	A	M	40+	4	MAO	NA	UNI	ACD	0.448	1328	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1724	NZ	S1A081	B	M	40+	4	MAO	NA	UNI	ACD	0.448	903	1	0	0	0	0	1	1.107	1.107	0.000	0.000	0.000
1551	NZ	S1A082	J	F	40+	4	MAO	NA	UNI	CLM	0.926	1483	0	2	0	0	0	2	1.349	0.000	1.349	0.000	0.000
1546	NZ	S1A082	V	F	40+	4	OTH	NA	NA	CLM	0.926	676	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1754	NZ	S1A083	A	F	20-29	1	MAO	NA	UNI	SLM	3.759	1000	1	3	1	4	0	9	9.000	1.000	3.000	1.000	4.000
1733	NZ	S1A083	B	F	30-39	3	MAO	NA	UNI	EDU	3.759	2458	2	1	0	1	0	4	1.627	0.814	0.407	0.000	0.407
1636	NZ	S1A084	A	F	40+	4	MAO	NA	UNI	ACD	0.258	1549	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1594	NZ	S1A084	J	F	40+	4	MAO	NA	NA	NA	0.258	2332	1	0	0	0	0	1	0.429	0.429	0.000	0.000	0.000
1683	NZ	S1A085	K	F	30-39	3	MAO	NA	UNI	CLM	0.985	936	1	0	0	0	0	1	1.068	1.068	0.000	0.000	0.000
1755	NZ	S1A085	R	F	20-29	1	NA	NA	UNI	CLM	0.985	1094	1	0	0	0	0	1	0.914	0.914	0.000	0.000	0.000
1640	NZ	S1A086	A	F	40+	4	MAO	NA	UNI	ACD	0.567	1816	0	0	0	1	0	1	0.551	0.000	0.000	0.000	0.551

1690	NZ	S1A086	B	F	40+	4	MAO	NA	NA	NA	0.567	1712	0	1	0	0	0	1	0.584	0.000	0.584	0.000	0.000
1590	NZ	S1A087	A	M	20-29	1	MAO	NA	UNI	ACD	4.599	1118	3	2	0	0	0	5	4.472	2.683	1.789	0.000	0.000
1591	NZ	S1A087	B	M	20-29	1	MAO	NA	NA	EDU	4.599	1709	0	7	1	0	0	8	4.681	0.000	4.096	0.585	0.000
1528	NZ	S1A088	A	F	20-29	1	MAO	NA	UNI	ACD	3.37	1160	2	1	1	0	0	4	3.448	1.724	0.862	0.862	0.000
1561	NZ	S1A088	B	F	20-29	1	NA	NA	UNI	CLM	3.37	1511	2	0	0	3	0	5	3.309	1.324	0.000	0.000	1.985
1732	NZ	S1A089	A	M	16-19	1	MAO	NA	NA	EDU	3.228	1041	1	1	0	1	0	3	2.882	0.961	0.961	0.000	0.961
1553	NZ	S1A089	B	M	16-19	1	MAO	NA	NA	EDU	3.228	1437	5	0	0	0	0	5	3.479	3.479	0.000	0.000	0.000
1654	NZ	S1A090	L	F	30-39	3	MAO	NA	UNI	CLM	0.344	1545	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1705	NZ	S1A090	Q	F	40+	4	MAO	NA	UNI	CLM	0.344	1362	1	0	0	0	0	1	0.734	0.734	0.000	0.000	0.000
1611	NZ	S1A091	A	F	30-39	3	OTH	NA	SEC	EDU	0	1110	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1612	NZ	S1A091	B	F	40+	4	OTH	NA	NA	NA	0	912	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1644	NZ	S1A092	A	F	20-29	1	OTH	NA	UNI	SLM	3.248	1341	2	2	0	2	0	6	4.474	1.491	1.491	0.000	1.491
1538	NZ	S1A092	B	F	20-29	1	OTH	NA	NA	EDU	3.248	814	1	0	0	0	0	1	1.229	1.229	0.000	0.000	0.000
1684	NZ	S1A093	P	M	16-19	1	OTH	NA	UNI	EDU	3.305	1792	3	4	0	0	0	7	3.906	1.674	2.232	0.000	0.000
1570	NZ	S1A093	S	M	16-19	1	OTH	NA	UNI	EDU	3.305	326	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1751	NZ	S1A094	A	F	16-19	1	OTH	NA	UNI	CLM	2.801	1648	2	1	0	2	0	5	3.034	1.214	0.607	0.000	1.214
1558	NZ	S1A094	B	F	20-29	1	OTH	NA	NA	EDU	2.801	494	0	1	0	0	0	1	2.024	0.000	2.024	0.000	0.000
1753	NZ	S1A095	B	F	16-19	1	OTH	NA	UNI	CLM	7.45	687	2	2	0	0	0	4	5.822	2.911	2.911	0.000	0.000
1598	NZ	S1A095	M	F	20-29	1	OTH	NA	UNI	ACD	7.45	1595	8	1	0	4	0	13	8.150	5.016	0.627	0.000	2.508
1556	NZ	S1A096	A	F	20-29	1	OTH	NA	UNI	EDU	5.611	1582	3	4	0	4	0	11	6.953	1.896	2.528	0.000	2.528
1643	NZ	S1A096	B	F	20-29	1	OTH	NA	UNI	CLM	5.611	913	1	1	1	0	0	3	3.286	1.095	1.095	1.095	0.000
1758	NZ	S1A097	G	M	40+	4	MAO	NA	NA	NA	0.442	1196	1	0	0	0	0	1	0.836	0.836	0.000	0.000	0.000
1613	NZ	S1A097	N	F	30-39	3	OTH	NA	NA	NA	0.442	1066	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1652	NZ	S1A098	D	F	30-39	3	OTH	NA	NA	CLM	0	813	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1718	NZ	S1A098	M	F	40+	4	OTH	NA	NA	NA	0	1417	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1547	NZ	S1A099	J	F	40+	4	OTH	NA	UNI	CLM	0	1000	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1635	NZ	S1A099	P	F	40+	4	OTH	NA	UNI	ACD	0	1295	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1572	NZ	S1A100	M	F	40+	4	OTH	NA	NA	ACD	0.372	1063	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1706	NZ	S1A100	O	F	40+	4	OTH	NA	UNI	ACD	0.372	1624	0	1	0	0	0	1	0.616	0.000	0.616	0.000	0.000

Table 145: Final data set of PhiE.

ID	VAR	FILE	SPEAKER	SEX	age.orig	AGE	ETH	DATE	EDU	OCC	L1	PAI	words	ini.n	med.n	fin.m	non.n	na.n	all.n	all	ini	med	fin	non
1765	PHI	S1A-001	A	M	21-30	2	FIL	2	UNI	EDU	TAG	2.131	962	1	0	0	0	0	1	1.040	1.040	0.000	0.000	0.000
1766	PHI	S1A-001	B	M	21-30	2	FIL	2	UNI	EDU	NA	2.131	1384	1	3	0	0	0	4	2.890	0.723	2.168	0.000	0.000
1768	PHI	S1A-002	A	M	21-30	2	OTH	2	UNI	EDU	NA	0.464	1045	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1769	PHI	S1A-002	B	F	31-40	3	FIL	2	UNI	ACD	TAG	0.464	1111	1	0	0	0	0	1	0.900	0.900	0.000	0.000	0.000
1771	PHI	S1A-003	A	F	41+	4	FIL	2	UNI	ACD	TAG	0.000	585	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1772	PHI	S1A-003	B	M	41+	4	FIL	2	UNI	ACD	OTH	0.000	1312	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1774	PHI	S1A-004	A	F	31-40	3	FIL	3	UNI	CLM	NA	0.972	2567	0	2	0	1	0	3	1.169	0.000	0.779	0.000	0.390
1775	PHI	S1A-004	B	F	31-40	3	FIL	3	UNI	CLM	NA	0.972	520	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1776	PHI	S1A-005	A	F	21-30	2	FIL	1	UNI	EDU	TAG	0.967	284	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1777	PHI	S1A-005	B	F	31-40	3	FIL	1	UNI	ACD	TAG	0.967	1784	0	0	0	2	0	2	1.121	0.000	0.000	0.000	1.121
1778	PHI	S1A-006	A	F	31-40	3	FIL	3	UNI	EDU	TAG	2.872	1173	1	2	0	2	1	6	5.115	0.853	1.705	0.000	1.705
1779	PHI	S1A-006	B	F	31-40	3	FIL	3	UNI	ACD	NA	2.872	1961	1	1	0	1	0	3	1.530	0.510	0.510	0.000	0.510
1780	PHI	S1A-007	A	F	16-20	1	FIL	2	UNI	EDU	NA	8.118	745	1	0	0	2	0	3	4.027	1.342	0.000	0.000	2.685
1781	PHI	S1A-007	B	F	16-20	1	FIL	2	UNI	EDU	NA	8.118	1349	5	6	0	3	0	14	10.378	3.706	4.448	0.000	2.224
1783	PHI	S1A-008	A	F	41+	4	FIL	2	UNI	ACD	TAG	0.465	183	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1784	PHI	S1A-008	B	M	31-40	3	FIL	2	UNI	ACD	OTH	0.465	1966	1	0	0	0	0	1	0.509	0.509	0.000	0.000	0.000
1786	PHI	S1A-009	A	F	21-30	2	FIL	1	UNI	EDU	TAG	0.000	183	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1787	PHI	S1A-009	B	F	31-40	3	FIL	1	UNI	NA	TAG	0.000	1966	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1789	PHI	S1A-010	A	F	21-30	2	FIL	1	UNI	EDU	TAG	1.890	496	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1790	PHI	S1A-010	B	F	21-30	2	FIL	1	UNI	ACD	TAG	1.890	2149	4	0	0	1	0	5	2.327	1.861	0.000	0.000	0.465
1792	PHI	S1A-011	A	M	16-20	1	FIL	1	UNI	EDU	NA	0.471	1905	1	0	0	0	0	1	0.525	0.525	0.000	0.000	0.000
1793	PHI	S1A-011	B	F	16-20	1	FIL	1	UNI	EDU	NA	0.471	218	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1795	PHI	S1A-012	A	F	16-20	1	FIL	2	UNI	EDU	TAG	3.737	915	6	5	0	0	0	11	12.022	6.557	5.464	0.000	0.000

1796	PHI	S1A-012	B	F	16-20	1	FIL	2	UNI	EDU	ENG	3.737	598	1	0	0	0	0	1	1.672	1.672	0.000	0.000	0.000
1797	PHI	S1A-012	C	F	16-20	1	FIL	2	UNI	EDU	TAG	3.737	740	0	2	0	0	2	4	5.405	0.000	2.703	0.000	0.000
1799	PHI	S1A-013	A	F	16-20	1	FIL	2	UNI	EDU	TAG	1.902	1144	1	3	0	0	0	4	3.497	0.874	2.622	0.000	0.000
1800	PHI	S1A-013	B	F	16-20	1	FIL	2	UNI	EDU	TAG	1.902	959	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1802	PHI	S1A-014	A	F	NA	NA	FIL	3	NA	NA	NA	0.000	719	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1803	PHI	S1A-014	B	M	NA	NA	FIL	3	NA	NA	NA	0.000	1725	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1805	PHI	S1A-015	A	F	21-30	2	FIL	2	UNI	CLM	NA	3.716	1904	4	3	0	1	0	8	4.202	2.101	1.576	0.000	0.525
1806	PHI	S1A-015	B	F	21-30	2	FIL	2	UNI	CLM	NA	3.716	249	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1808	PHI	S1A-016	A	F	16-20	1	FIL	2	UNI	EDU	TAG	1.416	1178	1	1	0	0	0	2	1.698	0.849	0.849	0.000	0.000
1809	PHI	S1A-016	B	F	16-20	1	FIL	2	UNI	EDU	OTH	1.416	940	1	0	0	0	0	1	1.064	1.064	0.000	0.000	0.000
1811	PHI	S1A-017	A	M	16-20	1	FIL	2	UNI	EDU	TAG	5.116	855	0	1	0	0	0	1	1.170	0.000	1.170	0.000	0.000
1812	PHI	S1A-017	B	F	16-20	1	FIL	2	UNI	EDU	TAG	5.116	1295	0	8	0	2	0	10	7.722	0.000	6.178	0.000	1.544
1814	PHI	S1A-018	A	F	21-30	2	FIL	2	UNI	EDU	TAG	1.829	736	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1815	PHI	S1A-018	B	F	16-20	1	FIL	2	UNI	EDU	TAG	1.829	1451	0	3	0	0	1	4	2.757	0.000	2.068	0.000	0.000
1817	PHI	S1A-019	A	M	21-30	2	FIL	2	UNI	EDU	NA	2.413	1133	0	1	0	1	0	2	1.765	0.000	0.883	0.000	0.883
1818	PHI	S1A-019	B	F	16-20	1	FIL	2	UNI	EDU	NA	2.413	939	2	1	0	0	0	3	3.195	2.130	1.065	0.000	0.000
1819	PHI	S1A-020	A	F	21-30	2	FIL	2	UNI	EDU	TAG	0.950	988	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1820	PHI	S1A-020	B	F	21-30	2	FIL	2	UNI	ACD	TAG	0.950	1118	1	1	0	0	0	2	1.789	0.894	0.894	0.000	0.000
1822	PHI	S1A-021	A	M	16-20	1	FIL	1	UNI	EDU	NA	0.000	1596	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1823	PHI	S1A-021	B	F	16-20	1	FIL	1	UNI	EDU	NA	0.000	310	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1824	PHI	S1A-021	C	F	16-20	1	FIL	1	UNI	EDU	NA	0.000	155	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1826	PHI	S1A-022	A	F	16-20	1	FIL	2	UNI	EDU	OTH	2.329	877	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1827	PHI	S1A-022	B	F	16-20	1	FIL	2	UNI	EDU	TAG	2.329	1270	3	1	0	1	0	5	3.937	2.362	0.787	0.000	0.787
1828	PHI	S1A-023	A	F	16-20	1	FIL	2	UNI	EDU	TAG	0.459	1111	0	1	0	0	0	1	0.900	0.000	0.900	0.000	0.000
1829	PHI	S1A-023	B	F	16-20	1	FIL	2	UNI	EDU	TAG	0.459	1070	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1831	PHI	S1A-024	A	M	31-40	3	FIL	2	UNI	ACD	TAG	0.000	1238	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1832	PHI	S1A-024	B	F	21-30	2	FIL	2	UNI	ACD	TAG	0.000	845	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1834	PHI	S1A-025_1	A	M	21-30	2	FIL	2	UNI	CLM	ENG	2.270	923	1	2	0	1	0	4	4.334	1.083	2.167	0.000	1.083
1835	PHI	S1A-025_1	B	F	21-30	2	OTH	2	UNI	EDU	ENG	2.270	839	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1837	PHI	S1A-025_2	A	F	21-30	2	FIL	2	UNI	EDU	TAG	2.101	318	0	1	0	0	0	1	3.145	0.000	3.145	0.000	0.000
1838	PHI	S1A-025_2	B	M	21-30	2	FIL	2	UNI	EDU	TAG	2.101	158	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1840	PHI	S1A-026	A	M	21-30	2	FIL	2	UNI	CLM	NA	4.693	665	1	2	0	3	0	6	9.023	1.504	3.008	0.000	4.511
1841	PHI	S1A-026	B	M	21-30	2	FIL	2	UNI	CLM	NA	4.693	1466	1	1	0	2	0	4	2.729	0.682	0.682	0.000	1.364
1843	PHI	S1A-027	A	F	16-20	1	FIL	2	UNI	EDU	NA	6.951	1168	3	2	0	1	0	6	5.137	2.568	1.712	0.000	0.856
1844	PHI	S1A-027	B	F	21-30	2	FIL	2	UNI	EDU	NA	6.951	990	5	2	0	2	0	9	9.091	5.051	2.020	0.000	2.020
1845	PHI	S1A-028	A	M	21-30	2	FIL	2	UNI	SLM	NA	0.958	1438	1	0	0	0	0	1	0.695	0.695	0.000	0.000	0.000
1846	PHI	S1A-028	B	M	21-30	2	FIL	2	UNI	CLM	NA	0.958	650	1	0	0	0	0	1	1.538	1.538	0.000	0.000	0.000
1848	PHI	S1A-029	A	M	16-20	1	FIL	2	UNI	EDU	NA	7.597	540	0	2	0	2	0	4	7.407	0.000	3.704	0.000	3.704
1849	PHI	S1A-029	B	M	16-20	1	FIL	2	UNI	EDU	NA	7.597	1566	6	3	1	2	0	12	7.663	3.831	1.916	0.639	1.277
1850	PHI	S1A-030	A	F	16-20	1	FIL	2	UNI	EDU	NA	4.198	1395	3	4	0	1	0	8	5.735	2.151	2.867	0.000	0.717
1851	PHI	S1A-030	B	F	16-20	1	FIL	2	UNI	EDU	NA	4.198	749	0	0	0	1	0	1	1.335	0.000	0.000	0.000	1.335
1854	PHI	S1A-031	A	F	16-20	1	FIL	2	UNI	EDU	NA	0.482	668	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1855	PHI	S1A-031	B	F	16-20	1	FIL	2	UNI	EDU	NA	0.482	1407	0	0	0	1	0	1	0.711	0.000	0.000	0.000	0.711
1857	PHI	S1A-032	A	F	31-40	3	FIL	2	UNI	ACD	NA	0.450	852	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1858	PHI	S1A-032	B	F	31-40	3	FIL	2	UNI	EDU	TAG	0.450	1368	1	0	0	0	0	1	0.731	0.731	0.000	0.000	0.000
1860	PHI	S1A-033	A	F	16-20	1	FIL	2	UNI	EDU	NA	5.231	697	1	0	0	0	0	1	1.435	1.435	0.000	0.000	0.000
1861	PHI	S1A-033	B	F	16-20	1	FIL	2	UNI	EDU	NA	5.231	1406	1	7	0	2	0	10	7.112	0.711	4.979	0.000	1.422
1863	PHI	S1A-034	A	M	16-20	1	FIL	2	UNI	EDU	NA	4.350	973	1	4	0	0	0	5	5.139	1.028	4.111	0.000	0.000
1864	PHI	S1A-034	B	F	16-20	1	FIL	2	UNI	EDU	NA	4.350	1096	1	3	0	0	0	4	3.650	0.912	2.737	0.000	0.000
1865	PHI	S1A-035	A	F	16-20	1	FIL	2	UNI	EDU	NA	2.943	1317	0	3	0	2	0	5	3.797	0.000	2.278	0.000	1.519

1866	PHI	S1A-035	B	F	16-20	1	FIL	2	UNI	EDU	NA	2.943	722	0	1	0	0	0	1	1.385	0.000	1.385	0.000	0.000
1868	PHI	S1A-036	A	F	41+	4	FIL	3	UNI	ACD	TAG	0.000	622	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1869	PHI	S1A-036	B	F	41+	4	FIL	3	UNI	ACD	TAG	0.000	1326	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1870	PHI	S1A-036	C	F	41+	4	FIL	3	UNI	ACD	OTH	0.000	900	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1871	PHI	S1A-037	A	F	16-20	1	FIL	2	UNI	EDU	NA	0.000	1016	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1872	PHI	S1A-037	B	F	16-20	1	FIL	2	UNI	EDU	NA	0.000	1081	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1873	PHI	S1A-038	A	F	16-20	1	FIL	2	UNI	EDU	NA	9.078	961	0	4	0	4	0	8	8.325	0.000	4.162	0.000	4.162
1874	PHI	S1A-038	B	F	16-20	1	FIL	2	UNI	EDU	NA	9.078	1132	2	8	0	1	0	11	9.717	1.767	7.067	0.000	0.883
1876	PHI	S1A-039	A	F	16-20	1	FIL	2	UNI	EDU	NA	1.040	1180	0	2	0	1	0	3	2.542	0.000	1.695	0.000	0.847
1877	PHI	S1A-039	B	F	16-20	1	FIL	2	UNI	EDU	NA	1.040	825	0	1	0	0	0	1	1.212	0.000	1.212	0.000	0.000
1878	PHI	S1A-039	C	F	16-20	1	FIL	2	UNI	EDU	NA	1.040	137	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1879	PHI	S1A-040	A	F	21-30	2	FIL	2	UNI	EDU	TAG	3.685	1018	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1880	PHI	S1A-040	B	F	21-30	2	FIL	2	UNI	EDU	TAG	3.685	1153	2	1	1	4	0	8	6.938	1.735	0.867	0.867	3.469
1882	PHI	S1A-041	A	F	16-20	1	FIL	2	UNI	EDU	NA	1.856	1104	1	0	0	1	0	2	1.812	0.906	0.000	0.000	0.906
1883	PHI	S1A-041	B	M	21-30	2	FIL	2	UNI	CLM	NA	1.856	1051	2	0	0	0	0	2	1.903	1.903	0.000	0.000	0.000
1885	PHI	S1A-042	A	F	16-20	1	FIL	2	UNI	EDU	TAG	2.101	829	1	1	0	0	0	2	2.413	1.206	1.206	0.000	0.000
1886	PHI	S1A-042	B	F	16-20	1	FIL	2	UNI	EDU	TAG	2.101	1297	1	0	0	2	0	3	2.313	0.771	0.000	0.000	1.542
1887	PHI	S1A-042	NA	F	16-20	1	FIL	2	UNI	EDU	TAG	2.101	131	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1888	PHI	S1A-043	A	F	16-20	1	FIL	2	UNI	EDU	NA	2.841	1142	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1889	PHI	S1A-043	B	F	16-20	1	FIL	2	UNI	EDU	NA	2.841	970	3	2	1	0	0	6	6.186	3.093	2.062	1.031	0.000
1891	PHI	S1A-044	A	F	21-30	2	FIL	2	UNI	ACD	TAG	1.368	690	0	1	0	0	0	1	1.449	0.000	1.449	0.000	0.000
1892	PHI	S1A-044	B	M	31-40	3	FIL	2	UNI	ACD	TAG	1.368	1503	0	0	0	2	0	2	1.331	0.000	0.000	0.000	1.331
1893	PHI	S1A-045	A	F	16-20	1	FIL	1	UNI	EDU	NA	0.000	900	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1894	PHI	S1A-045	B	F	21-30	2	FIL	1	UNI	EDU	NA	0.000	1224	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1896	PHI	S1A-046	A	F	16-20	1	FIL	2	UNI	EDU	NA	1.924	487	0	1	0	1	0	2	4.107	0.000	2.053	0.000	2.053

1897	PHI	S1A-046	B	F	16-20	1	FIL	2	UNI	EDU	NA	1.924	1592	0	2	0	0	0	2	1.256	0.000	1.256	0.000	0.000
1898	PHI	S1A-047	A	F	16-20	1	FIL	2	UNI	EDU	NA	7.741	634	0	2	0	0	0	2	3.155	0.000	3.155	0.000	0.000
1899	PHI	S1A-047	B	F	16-20	1	FIL	2	UNI	EDU	NA	7.741	1562	2	12	1	0	0	15	9.603	1.280	7.682	0.640	0.000
1904	PHI	S1A-049	A	F	16-20	1	FIL	2	UNI	EDU	TAG	4.737	1031	2	7	0	1	0	10	9.699	1.940	6.790	0.000	0.970
1905	PHI	S1A-049	B	F	16-20	1	FIL	2	UNI	EDU	TAG	4.737	1080	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1907	PHI	S1A-050	A	M	41+	4	FIL	2	UNI	CLM	TAG	0.000	929	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1908	PHI	S1A-050	B	M	41+	4	FIL	2	UNI	CLM	TAG	0.000	864	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1909	PHI	S1A-050	C	F	41+	4	FIL	2	UNI	CLM	OTH	0.000	365	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1910	PHI	S1A-051	A	M	41+	4	FIL	1	UNI	ACD	TAG	0.000	973	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1911	PHI	S1A-051	B	M	31-40	3	FIL	1	UNI	ACD	NA	0.000	955	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1913	PHI	S1A-052	A	F	31-40	3	FIL	2	UNI	ACD	TAG	1.507	868	1	0	0	0	0	1	1.152	1.152	0.000	0.000	0.000
1914	PHI	S1A-052	B	M	31-40	3	FIL	2	UNI	ACD	TAG	1.507	1123	0	2	0	0	0	2	1.781	0.000	1.781	0.000	0.000
1916	PHI	S1A-053	A	F	21-30	2	FIL	2	UNI	ACD	TAG	0.000	1061	2	0	0	0	0	2	1.885	1.885	0.000	0.000	0.000
1917	PHI	S1A-053	B	F	21-30	2	FIL	2	UNI	ACD	OTH	0.000	306	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1918	PHI	S1A-053	C	F	21-30	2	FIL	2	UNI	ACD	TAG	0.000	848	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1920	PHI	S1A-054	B	M	41+	4	FIL	2	UNI	CLM	TAG	0.991	949	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1921	PHI	S1A-054	C	F	41+	4	FIL	2	UNI	CLM	TAG	0.991	629	1	0	0	0	0	1	1.590	1.590	0.000	0.000	0.000
1922	PHI	S1A-054	D	M	41+	4	FIL	2	UNI	CLM	TAG	0.991	380	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1924	PHI	S1A-055	A	F	16-20	1	FIL	1	UNI	EDU	NA	3.610	642	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1925	PHI	S1A-055	B	F	16-20	1	FIL	1	UNI	EDU	NA	3.610	494	3	0	0	0	0	3	6.073	6.073	0.000	0.000	0.000
1926	PHI	S1A-055	C	M	16-20	1	FIL	1	UNI	EDU	NA	3.610	337	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1927	PHI	S1A-056	A	F	21-30	2	FIL	2	UNI	ACD	TAG	0.748	636	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1928	PHI	S1A-056	B	F	21-30	2	FIL	2	UNI	ACD	ENG+	0.748	745	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1929	PHI	S1A-056	C	F	41+	4	FIL	2	UNI	ACD	TAG	0.748	592	1	0	0	0	0	1	1.689	1.689	0.000	0.000	0.000
1930	PHI	S1A-057	A	F	21-30	2	FIL	1	UNI	EDU	NA	0.473	1458	0	1	0	0	0	1	0.686	0.000	0.686	0.000	0.000

1931	PHI	S1A-057	B	F	16-20	1	FIL	1	UNI	EDU	NA	0.473	657	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1933	PHI	S1A-058	A	F	16-20	1	FIL	2	UNI	EDU	TAG	0.941	1217	1	1	0	0	0	2	1.643	0.822	0.822	0.000	0.000
1934	PHI	S1A-058	B	F	16-20	1	FIL	2	UNI	EDU	TAG	0.941	908	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1936	PHI	S1A-059	A	F	31-40	3	FIL	2	UNI	ACD	OTH	1.833	1179	3	1	0	0	0	4	3.393	2.545	0.848	0.000	0.000
1937	PHI	S1A-059	B	F	16-20	1	FIL	2	UNI	EDU	TAG	1.833	1003	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1939	PHI	S1A-060	A	M	21-30	2	FIL	2	UNI	ACD	TAG	4.342	1115	0	3	0	1	0	4	3.587	0.000	2.691	0.000	0.897
1940	PHI	S1A-060	B	F	21-30	2	FIL	2	UNI	ACD	TAG	4.342	1188	1	4	1	0	0	6	5.051	0.842	3.367	0.842	0.000
1942	PHI	S1A-061	A	F	21-30	2	FIL	2	UNI	ACD	TAG	0.987	1400	0	0	0	1	0	1	0.714	0.000	0.000	0.000	0.714
1943	PHI	S1A-061	B	F	21-30	2	FIL	2	UNI	ACD	OTH	0.987	627	1	0	0	0	0	1	1.595	1.595	0.000	0.000	0.000
1946	PHI	S1A-062	A	F	21-30	2	FIL	2	UNI	ACD	OTH	5.902	598	1	0	0	0	0	1	1.672	1.672	0.000	0.000	0.000
1947	PHI	S1A-062	B	M	21-30	2	FIL	2	UNI	ACD	TAG	5.902	811	1	0	0	0	0	1	1.233	1.233	0.000	0.000	0.000
1948	PHI	S1A-062	C	F	21-30	2	FIL	2	UNI	ACD	TAG	5.902	714	2	5	0	1	0	8	11.204	2.801	7.003	0.000	1.401
1950	PHI	S1A-063	A	M	21-30	2	FIL	2	UNI	EDU	TAG	0.969	736	0	1	0	0	0	1	1.359	0.000	1.359	0.000	0.000
1951	PHI	S1A-063	B	M	21-30	2	FIL	2	UNI	NA	TAG	0.969	1328	0	0	0	1	0	1	0.753	0.000	0.000	0.000	0.753
1953	PHI	S1A-064	A	F	21-30	2	FIL	2	UNI	ACD	OTH	6.692	1145	9	0	0	0	0	9	7.860	7.860	0.000	0.000	0.000
1954	PHI	S1A-064	B	M	21-30	2	FIL	2	UNI	ACD	OTH	6.692	947	2	2	0	1	0	5	5.280	2.112	2.112	0.000	1.056
1956	PHI	S1A-065	A	M	16-20	1	FIL	2	UNI	EDU	TAG	0.000	1144	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1957	PHI	S1A-065	B	M	16-20	1	FIL	2	UNI	EDU	TAG	0.000	996	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1958	PHI	S1A-066	A	M	21-30	2	FIL	2	UNI	EDU	TAG	0.959	1424	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1959	PHI	S1A-066	B	F	21-30	2	FIL	2	UNI	CLM	TAG	0.959	662	1	0	1	0	0	2	3.021	1.511	0.000	1.511	0.000
1960	PHI	S1A-067	A	M	21-30	2	FIL	2	UNI	EDU	NA	6.277	1396	1	6	0	2	0	9	6.447	0.716	4.298	0.000	1.433
1961	PHI	S1A-067	B	M	21-30	2	FIL	2	UNI	EDU	NA	6.277	675	2	0	0	1	1	4	5.926	2.963	0.000	0.000	1.481
1962	PHI	S1A-068	A	M	21-30	2	FIL	2	UNI	NA	TAG	3.622	1651	5	1	0	2	0	8	4.846	3.028	0.606	0.000	1.211
1963	PHI	S1A-068	B	F	16-20	1	FIL	2	UNI	NA	TAG	3.622	558	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1964	PHI	S1A-069	A	F	21-30	2	FIL	2	UNI	ACD	TAG	2.683	752	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

1965	PHI	S1A-069	B	F	21-30	2	FIL	2	UNI	ACD	OTH	2.683	253	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1966	PHI	S1A-069	C	F	21-30	2	FIL	2	UNI	ACD	TAG	2.683	195	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1967	PHI	S1A-069	D	M	21-30	2	FIL	2	UNI	ACD	TAG	2.683	923	0	1	0	2	0	3	3.250	0.000	1.083	0.000	2.167
1969	PHI	S1A-070	A	F	21-30	2	FIL	2	UNI	ACD	TAG	5.706	747	1	0	0	0	0	1	1.339	1.339	0.000	0.000	0.000
1970	PHI	S1A-070	B	F	21-30	2	FIL	2	UNI	ACD	TAG	5.706	1356	4	6	1	0	0	11	8.112	2.950	4.425	0.737	0.000
1974	PHI	S1A-071	B	F	21-30	2	FIL	1	UNI	EDU	ENG+	2.554	134	1	1	0	0	0	2	14.925	7.463	7.463	0.000	0.000
1975	PHI	S1A-071	C	F	21-30	2	FIL	1	UNI	ACD	TAG	2.554	1229	3	1	0	0	0	4	3.255	2.441	0.814	0.000	0.000
1976	PHI	S1A-071	D	M	21-30	2	FIL	1	UNI	ACD	TAG	2.554	729	1	0	0	0	0	1	1.372	1.372	0.000	0.000	0.000
1978	PHI	S1A-072	A	M	21-30	2	FIL	2	UNI	ACD	TAG	0.000	1085	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1979	PHI	S1A-072	B	F	21-30	2	FIL	2	UNI	CLM	TAG	0.000	1166	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1981	PHI	S1A-073	A	F	21-30	2	FIL	2	UNI	CLM	OTH	2.644	608	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1982	PHI	S1A-073	B	F	31-40	3	FIL	2	UNI	CLM	TAG	2.644	574	0	1	0	0	0	1	1.742	0.000	1.742	0.000	0.000
1983	PHI	S1A-073	C	F	31-40	3	FIL	2	UNI	CLM	TAG	2.644	939	3	0	0	0	0	3	3.195	3.195	0.000	0.000	0.000
1984	PHI	S1A-074	A	M	16-20	1	FIL	2	UNI	EDU	TAG	0.000	1043	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1985	PHI	S1A-074	B	M	21-30	2	FIL	2	UNI	EDU	TAG	0.000	1073	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1989	PHI	S1A-075	A	F	16-20	1	FIL	2	UNI	EDU	OTH	4.597	1239	4	3	0	0	1	8	6.457	3.228	2.421	0.000	0.000
1990	PHI	S1A-075	B	F	16-20	1	FIL	2	UNI	EDU	OTH	4.597	1154	0	0	0	2	1	3	2.600	0.000	0.000	0.000	1.733
1991	PHI	S1A-076	A	F	21-30	2	FIL	2	UNI	EDU	NA	0.472	1293	0	1	0	0	0	1	0.773	0.000	0.773	0.000	0.000
1992	PHI	S1A-076	B	F	16-20	1	FIL	2	UNI	EDU	NA	0.472	825	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1994	PHI	S1A-077	A	F	16-20	1	FIL	2	UNI	EDU	NA	0.465	893	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1995	PHI	S1A-077	B	M	21-30	2	FIL	2	UNI	EDU	NA	0.465	1259	1	0	0	0	0	1	0.794	0.794	0.000	0.000	0.000
1997	PHI	S1A-078	A	F	16-20	1	FIL	2	UNI	EDU	NA	5.698	800	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
1998	PHI	S1A-078	B	F	16-20	1	FIL	2	UNI	EDU	NA	5.698	1306	0	12	0	0	0	12	9.188	0.000	9.188	0.000	0.000
2000	PHI	S1A-079	A	F	16-20	1	FIL	2	UNI	EDU	NA	0.473	881	0	0	1	0	0	1	1.135	0.000	0.000	1.135	0.000
2001	PHI	S1A-079	B	F	16-20	1	FIL	2	UNI	EDU	NA	0.473	1234	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

2002	PHI	S1A-080	A	F	16-20	1	FIL	2	UNI	EDU	NA	1.997	1440	2	0	0	1	1	4	2.778	1.389	0.000	0.000	0.694
2003	PHI	S1A-080	B	F	16-20	1	FIL	2	UNI	EDU	NA	1.997	563	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2004	PHI	S1A-081	A	F	16-20	1	FIL	2	UNI	EDU	TAG	6.108	843	2	1	0	0	0	3	3.559	2.372	1.186	0.000	0.000
2005	PHI	S1A-081	B	F	NA	NA	FIL	2	UNI	EDU	OTH	6.108	958	5	1	0	2	0	8	8.351	5.219	1.044	0.000	2.088
2006	PHI	S1A-082	A	M	21-30	2	FIL	2	UNI	ACD	TAG	1.575	928	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2007	PHI	S1A-082	B	M	21-30	2	FIL	2	UNI	ACD	TAG	1.575	764	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2008	PHI	S1A-082	C	F	31-40	3	FIL	2	UNI	ACD	TAG	1.575	506	2	0	0	0	0	2	3.953	3.953	0.000	0.000	0.000
2009	PHI	S1A-083	A	F	21-30	2	FIL	2	UNI	ACD	TAG	4.669	1080	0	1	0	0	0	1	0.926	0.000	0.926	0.000	0.000
2010	PHI	S1A-083	B	F	21-30	2	FIL	2	UNI	ACD	ENG+	4.669	460	1	0	0	0	1	2	4.348	2.174	0.000	0.000	0.000
2011	PHI	S1A-083	C	F	21-30	2	FIL	2	UNI	ACD	TAG	4.669	611	0	3	0	0	0	3	4.910	0.000	4.910	0.000	0.000
2013	PHI	S1A-084	A	M	31-40	3	FIL	2	UNI	CLM	TAG	1.775	1530	2	1	0	1	0	4	2.614	1.307	0.654	0.000	0.654
2014	PHI	S1A-084	B	F	21-30	2	FIL	2	UNI	ACD	TAG	1.775	723	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2015	PHI	S1A-085	A	F	21-30	2	FIL	2	UNI	ACD	TAG	8.158	912	1	8	1	2	0	12	13.158	1.096	8.772	1.096	2.193
2016	PHI	S1A-085	B	F	31-40	3	FIL	2	UNI	CLM	TAG	8.158	1417	2	3	0	2	0	7	4.940	1.411	2.117	0.000	1.411
2017	PHI	S1A-086	A	F	16-20	1	FIL	2	UNI	EDU	TAG	1.349	1313	0	1	0	0	0	1	0.762	0.000	0.762	0.000	0.000
2018	PHI	S1A-086	B	M	16-20	1	FIL	2	UNI	EDU	TAG	1.349	911	1	0	1	0	0	2	2.195	1.098	0.000	1.098	0.000
2019	PHI	S1A-087	A	F	21-30	2	FIL	2	UNI	ACD	OTH	0.412	1225	0	1	0	0	0	1	0.816	0.000	0.816	0.000	0.000
2020	PHI	S1A-087	B	M	31-40	3	FIL	2	UNI	ACD	OTH	0.412	1200	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2022	PHI	S1A-088	A	M	31-40	3	FIL	1	UNI	EDU	ENG+	2.205	1411	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2023	PHI	S1A-088	B	M	31-40	3	FIL	1	UNI	ACD	ENG+	2.205	736	1	0	0	0	0	1	1.359	1.359	0.000	0.000	0.000
2024	PHI	S1A-088	C	M	21-30	2	FIL	1	UNI	ACD	ENG+	2.205	171	0	0	0	0	1	1	5.848	0.000	0.000	0.000	0.000
2029	PHI	S1A-091	A	M	21-30	2	FIL	2	UNI	CLM	TAG	0.488	798	1	0	0	0	0	1	1.253	1.253	0.000	0.000	0.000
2030	PHI	S1A-091	B	F	21-30	2	FIL	2	UNI	ACD	TAG	0.488	1251	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000
2032	PHI	S1A-092	A	F	31-40	3	FIL	2	NA	NA	OTH	0.392	1260	0	0	0	1	0	1	0.794	0.000	0.000	0.000	0.794
2033	PHI	S1A-092	B	M	41+	4	FIL	2	NA	NA	OTH	0.392	1288	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000

11.1 Eidesstattliche Versicherung

Hierdurch versichere ich an Eides Statt, dass ich die Arbeit selbständig angefertigt, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die den herangezogenen Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

_____

Martin Schweinberger, Hamburg, den 10.6.2014