

THE DEVELOPMENT OF PRAGMATIC MARKERS IN CANADIAN ENGLISH

by

Derek Denis

A thesis submitted in conformity with the requirements  
for the degree of Doctor of Philosophy  
Graduate Department of Linguistics  
University of Toronto

© Copyright 2015 by Derek Denis

# Abstract

The Development of Pragmatic Markers in Canadian English

Derek Denis

Doctor of Philosophy

Graduate Department of Linguistics

University of Toronto

2015

This thesis examines the mechanisms of linguistic change involved in the actualization and subsequent development of pragmatic markers (PMs). Using the variationist method, I test the predictions of analysts who have argued that the development of PMs unfolds according to grammaticalization theory (*inter alia* Hopper and Traugott 1995). In doing so, I address two desiderata suggested in the literature for a better understanding of how PMs change: 1) an examination of both real- and apparent-time data (Pichler and Levey 2011), and 2) an examination of multiple features (Tagliamonte and Denis 2010).

The data come from several collections of recorded Ontario English. I examine the Toronto English Archive (Tagliamonte 2006b), a contemporary sociolinguistic corpus, and two collections of oral history interviews recorded in 1975 and 1984, the Belleville Oral History Project and the Farm Work and Farm Life Since 1890 Oral History Collection. Together, these represent an apparent-time span of over one hundred years. The features are general extenders and epistemic parentheticals both of which have been argued to show evidence of ongoing grammaticalization (e.g., Cheshire 2007; Thompson and Mulac 1991).

The analysis of the general extenders system finds a lack of evidence for phonetic reduction, decategorialization, and semantic-pragmatic shift of the innovative variant *and stuff*—each of these changes is a critical component of grammaticalization

according to grammaticalization theory. I further show that although the epistemic parenthetical system of Ontario English reorganized through the twentieth century, the changes involved were not the result of gradual grammaticalization, as evidenced by a lack of ongoing fusion, a lack of change in syntactic mobility, and a lack of ongoing semantic bleaching.

While the two case studies provide evidence against the idea that PMs develop according to grammaticalization theory, there is no denying that grammaticalization, *the phenomenon*, has taken place: lexical elements have become PMs. I conclude that the evidence suggests the grammaticalization of pragmatic markers as abrupt reanalysis of lexical material, from one syntactic category to some other syntactic category (see Roberts and Roussou 2003).

DEDICATION

*This thesis is dedicated to my great champion, critic, and companion, A.M.*

## Acknowledgements

I'd first like to acknowledge my supervisor Sali Tagliamonte and the members of my advisory committee Jack Chambers and Elizabeth Cowper. All three were integral in shaping the direction and argumentation of this thesis. I hope the influence that each of you has had on me and this project shines through! At the end of the writing process, I also benefitted greatly from my defense committee members Naomi Nagy, Aaron Dinkin, and my external member, Jenny Cheshire. Many of their comments have been incorporated into the final product here, while others I will ruminate on for years!

In addition to my committee, there are a number of people who I am grateful to have had a chance to discuss this work with. In particular, I have had lengthy discussions with Alex Motut, Ruth Maddeaux, Christopher Spahr, Radu Craioveanu, Ross Godfrey, Emily Clare, Dan Milway, Matt Hunt Gardner, Marisa Brook, Youri Zab-bal, and Alex D'Arcy.<sup>1</sup> Analyses in this thesis were presented for various audiences. I'd like to thank those at DiPVaC2014 including Heike Pichler, Maddie Shellgren, Stephen Levey, Cathleen Waters, Claire Childs, and Ashley Hesson; those at various CVC conferences including Shana Poplack, Gerard van Herk, Charles Boberg, James Walker, Nathalie Dion, Rick Grimm, and Joe Roy; and audiences at various NWAVs, LSAs, ADSs, and CLAs.

There are also several people who I want to acknowledge for their unwavering support throughout grad school. First, I want to acknowledge the LGCU. I really think us grad students have something special in the depo that is unlike any department in the world. We are *far* more than colleagues. We are friends; we are family; we are always there for each other. There are people who I first met in the lounge who I consider to be among the best friends I've ever had.<sup>2</sup> The hardest thing about leaving U of T will be that I'll no longer get to see you all on a daily basis, but the

---

<sup>1</sup>I've also had extended discussions with Jam the cat, but she never really has much to say.

<sup>2</sup>Sorry guys, I'm not going to name names, you know who you are.

fact that some of my aforementioned friends have come and gone, yet we remain close, makes me smile! I hope you'll let me come to pub in the future! I've also benefitted greatly from all the faculty members in the department. Having been here through my undergrad years I had the opportunity to be taught by almost every person. Additionally, Mary Hsu and Jill Given-King (and formerly Bill Forrest) have been the behind-the-scenes team that allowed the department to run. They are not acknowledged enough for their work!

My mom and dad have been nothing but supportive through my journey through grad school. One's parents are an integral part of who one is, what one does, and how one approaches life. Everyday I'm grateful to have received from mine a fraction of my mom's incredible work ethic and a fraction of my dad's creativity. My brother has always been an inspiration to me. He listened to punk, so I listened to punk. He did an undergrad at U of T, so I did too. He became a sociologist, so I became a sociolinguist. He went to grad school, so I went to grad school.

Lastly, perhaps the two events that have most drastically shaped my life to this point happened within metres of each other, albeit with a few brutalist concrete walls and almost five years separating them. The first almost never happened, and the second most likely would have never happened without the first: 1) meeting Sali and becoming an RA in my second year of undergrad at U of T; 2) meeting Alex in my MA year and falling in love. It's scary to think how different my life might be if one of those two events didn't take place! Certainly, this dissertation would not have been written!

Sali has been my teacher, my co-author, my guide, and my friend from the beginning of my academic career. She is my most important academic influence and all of my work is built on her foundations. I hope that this thesis has made her proud and that I will continue to make her proud, just as she will continue to inspire me!

Alex, I'm at a loss for words. Nothing I write here will come close to sufficiently

summarizing how important you are to me! We began this wild grad school journey together and now we're near the end; but that is so secondary. I wasn't me until I was with you and our real journey has barely started! I dedicate this thesis to you; you resonate through every page herein just like you do in every aspect of my life!

## Permissions declaration

All examples from Niagara and Eastern Ontario are from the *Farm Work and Farm Life in Ontario Since 1890* oral history project records (Archive of Ontario reference #: RG 16-200). Copyright is primarily held by the Crown. Permission to publish examples has been granted by the Archive of Ontario.

Permission to use the Belleville Oral History Project was granted by Professor Sali A. Tagliamonte (University of Toronto) through arrangements with the Hastings County Historical Society in May 2007.

Permission to use the Toronto English Archive was granted by Professor Sali A. Tagliamonte (University of Toronto).



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Grammaticalization . . . . .	4
1.1.1	Grammaticalization: Phenomenon and Theory . . . . .	4
1.1.2	Grammaticalization of Pragmatic Markers . . . . .	7
1.1.3	Criticisms of Grammaticalization Theory . . . . .	9
1.2	Overview of the Argument . . . . .	21
<b>2</b>	<b>General Methodology and the Earlier Ontario English data</b>	<b>24</b>
2.1	Pragmatic Markers, Variation, and Change . . . . .	25
2.1.1	The Variationist Method . . . . .	25
2.1.2	Linguistic Variation at the Extra-Sentential Level . . . . .	34
2.1.3	The Variationist Method and Grammaticalization . . . . .	38
2.2	The Earlier Ontario English Data . . . . .	44
2.2.1	Constitution of the Earlier Ontario English data . . . . .	47
2.2.2	Construction/Compilation . . . . .	53
2.2.3	A note on possible confounds: Time, urban-ness, and social class	62
2.2.4	Oral Histories as Historical Variationist Data . . . . .	63
2.3	A Note on Tools . . . . .	73
2.3.1	Logistic Regression Modelling in R as Variable Rule Analysis . .	73
2.3.2	Presentation of statistical results . . . . .	75

<b>3</b>	<b>The Development of General Extenders</b>	<b>76</b>
3.1	Introduction . . . . .	76
3.2	General Extenders and Grammaticalization . . . . .	78
3.2.1	Sociolinguistic Approaches to GEs . . . . .	84
3.3	General Extenders in EOE . . . . .	98
3.3.1	Overall Distribution . . . . .	98
3.3.2	The Long-Term Trajectory . . . . .	103
3.4	The Long-Term Trajectory of the Mechanisms of Grammaticalization . . . . .	106
3.4.1	Phonetic Reduction . . . . .	107
3.4.2	Decategorialization . . . . .	121
3.4.3	Semantic Bleaching . . . . .	133
3.4.4	Pragmatic Shift . . . . .	139
3.5	The Development of GEs: Multiple, Independent Changes . . . . .	147
<b>4</b>	<b>The Development of Epistemic Parentheticals</b>	<b>151</b>
4.1	Introduction . . . . .	151
4.2	Background . . . . .	152
4.3	Variable Context . . . . .	168
4.4	Tracking Changes Over the 20 <sup>th</sup> Century . . . . .	172
4.4.1	Overall Distribution in Ontario Across Time . . . . .	172
4.4.2	Intervening Material: Negation, Modals, and Adverbials . . . . .	176
4.4.3	Syntactic Position . . . . .	181
4.4.4	Epistemic/Doxastic Strength . . . . .	190
4.5	The Development of EPs: Lexical Replacement, Specialization . . . . .	209
<b>5</b>	<b>Conclusion</b>	<b>212</b>
5.1	Introduction . . . . .	212
5.2	The Non-Gradual Development of Pragmatic Markers . . . . .	212

5.2.1	Excursus on Variation, Change, and Competing Grammars . . .	214
5.2.2	Gradualness, Discreteness, and the Constant Rate Effect . . . . .	219
5.2.3	Reconciling the Constant Rate Effect with Pragmatic Change . .	224
5.2.4	A Schematic for the Development of Pragmatic Markers as (Abrupt) Reanalysis . . . . .	227
5.3	Implications . . . . .	233
5.3.1	Implications for Grammaticalization Theory . . . . .	233
5.3.2	Implications for Variationist Work on Pragmatic Variables . . . .	235
5.4	Next Steps . . . . .	235
	<b>Appendices</b>	<b>240</b>
	<b>A GEs Appendix</b>	<b>240</b>
	<b>Bibliography</b>	<b>245</b>

# List of Tables

1.1	Grammaticalization of <i>going to</i> . . . . .	7
2.1	Thompson and Mulac's (1991) Table 7 . . . . .	42
2.2	Thompson and Mulac's (1991) Table 7, revised . . . . .	43
2.3	Tagliamonte 2012, Table 1.1 . . . . .	43
2.4	Tagliamonte 2012, Table 1.2 . . . . .	43
2.5	Distribution of interviews in FWFL by region, speaker sex and age. . .	49
2.6	Distribution of interviews in <i>Belleville 1975</i> by speaker sex and age. . .	50
2.7	Demographics of the Belleville 1975 speakers. . . . .	56
2.8	Demographics of the Eastern Ontario speakers. . . . .	58
2.9	Demographics of the Niagara speakers. . . . .	60
2.10	Temporal coverage of TEA and EOE. . . . .	62
2.11	Comparison of FWFL interview schedule to standard sociolinguistic interview topics. . . . .	68
3.1	Typical GE templatic structure . . . . .	81
3.2	Tests of grammaticalization through apparent time. . . . .	92
3.3	Diagnostics of grammaticalization of <i>stuff</i> type GEs through apparent time in Toronto English . . . . .	94
3.4	Grammaticalization in York English . . . . .	96

3.5	Normalized frequency of adjunctive and disjunctive GEs in EOE, TEA, and Cheshire 2007. . . . .	101
3.6	Top 6 GE types in EOE communities . . . . .	102
3.7	Top 6 GE types in TEA age groups . . . . .	102
3.8	Prototypical GEs, short and long forms . . . . .	110
3.9	Test of phonetic reduction in real time . . . . .	111
3.10	Mixed-effects logistic regression testing phonetic reduction . . . . .	118
3.11	Mixed-effects logistic regression testing decategorialization. . . . .	129
3.12	Analysis of deviance, $\chi^2$ test for model reported in Table 3.11. . . . .	129
3.13	Test of semantic bleaching in real-time . . . . .	138
3.14	Co-occurring DMs . . . . .	141
3.15	Pichler and Levey's taxonomy of semantic-pragmatic change in GEs . .	144
3.16	Summary of the mechanisms of grammaticalization in Ontario English	149
4.1	Rodríguez Louro and Harris' (2013) predictions for the grammatical-ization of EPs . . . . .	166
4.2	Distribution of EPs in EOE and TEA . . . . .	174
4.3	Frequency of negated EPs . . . . .	177
4.4	Frequency of modal EPs . . . . .	179
4.5	Frequency of EPs with intervening adverbials . . . . .	180
4.6	Mixed-effects logistic regression testing the effect of syntactic position of <i>I think</i> over time. . . . .	187
4.7	Analysis of deviance, $\chi^2$ test for model reported in Table 4.6. . . . .	187
4.8	Mixed-effects logistic regression testing the effect of syntactic position of <i>I guess</i> over time. . . . .	189
4.9	Analysis of deviance, $\chi^2$ test for model reported in Table 4.8. . . . .	189
4.10	Distributional analysis of EPs by subject of the complement in EOE and TEA . . . . .	197

4.11	Mixed-effects logistic regression testing the effect of syntactic position of <i>I guess</i> over time. . . . .	197
4.12	Analysis of deviance, $\chi^2$ test for model reported in Table 4.11. . . . .	198
4.13	Mixed-effects logistic regression testing the effect of syntactic position of <i>I guess</i> over time. . . . .	199
4.14	Analysis of deviance, $\chi^2$ test for model reported in Table 4.13. . . . .	200
4.15	Mixed-effects logistic regression testing the effect of syntactic position of <i>I suppose</i> in EOE. . . . .	202
4.16	Analysis of deviance, $\chi^2$ test for model reported in Table 4.15. . . . .	202
4.17	Summary of complement clause subject results. . . . .	202
4.18	Negation and complement clause subject. . . . .	205
5.1	Summary of results. . . . .	225
A.1	Complete list of adjunctive GE forms, by raw frequency in EOE . . . . .	240
A.2	Complete list of adjunctive GE forms, by raw frequency in EOE, <i>cont.</i> . . . .	241
A.3	Complete list of disjunctive GE forms, by raw frequency in EOE . . . . .	242
A.4	Complete list of connectorless GE forms, by raw frequency in EOE . . . . .	243
A.5	Overall distribution of all general extender types. Frequency of all GEs (N). . . . .	244

# List of Figures

1.1	Bailey's (1973) wave model . . . . .	16
2.1	The three regions of the EOE . . . . .	54
2.2	Birthplaces of Belleville 1975 speakers . . . . .	56
2.3	Birthplaces of Eastern Ontario speakers . . . . .	59
2.4	Birthplaces of Niagara speakers . . . . .	61
2.5	Histogram of speakers across 20 <sup>th</sup> century . . . . .	62
3.1	Typology of GEs . . . . .	83
3.2	Meta-analysis of GE frequency . . . . .	85
3.3	Tagliamonte and Denis' (2010) Figure 8 . . . . .	93
3.4	Innovation of <i>stuff</i> forms . . . . .	95
3.5	Proportion of main variants of adjunctive general extenders over apparent time . . . . .	104
3.6	Proportion of main variants of disjunctive general extenders over apparent time . . . . .	105
3.7	Mean individual differences of the normalized frequency of long and short <i>stuff</i> , <i>thing</i> , <i>something</i> and <i>everything</i> type GEs in Belleville, Eastern Ontario and Niagara . . . . .	110
3.8	Proportion of long <i>stuff</i> , <i>thing</i> , <i>something</i> , <i>everything</i> , and <i>so on</i> GEs (vs. short forms) through apparent time. N = 2044. . . . .	115

3.9	Probability of <i>like that</i> comparative (vs. $\emptyset$ ) by GE type through apparent time . . . . .	120
3.10	GE referents over time . . . . .	125
3.11	GE variants by referent over time . . . . .	127
3.12	Stylized models of change . . . . .	128
3.13	Probability of <i>stuff</i> GE by referents over time . . . . .	130
3.14	Co-occurrence of discourse markers over apparent time . . . . .	143
4.1	Distribution of EPs in EOE and TEA . . . . .	173
4.2	EPs apparent time in EOE . . . . .	175
4.3	Distribution of EP variants by polarity . . . . .	178
4.4	Conditional inference tree of (semi-)independent EPs . . . . .	184
4.5	Distribution of EP variants by syntactic position . . . . .	185
4.6	Probability of <i>I think</i> by syntactic position over time . . . . .	188
4.7	Probability of <i>I guess</i> by syntactic position over time . . . . .	189
4.8	Distribution of EPs by subject of complement over time . . . . .	195
4.9	Probability of <i>I think</i> by subject of complement over time . . . . .	199
4.10	Probability of <i>I guess</i> by subject of complement over time . . . . .	201
4.11	Probability of <i>I suppose</i> by subject of complement in EOE . . . . .	203
5.1	Kroch (1989) Figure 7 . . . . .	219



# Chapter 1

## Introduction

A number of variationist studies published in the early part of the twenty-first century have presented compelling counter-evidence to the idea that grammaticalization theory can be applied to the development of pragmatic markers. In particular, Tagliamonte and Denis (2010) and Pichler and Levey (2011) conclude that there is a lack of evidence supporting the idea that the changes in the general extenders (GEs) variable system are consistent with gradual, ongoing grammaticalization. That is, specific changes theoretically associated with grammaticalization—phonetic reduction, decategorialization, and semantic-pragmatic expansion—are not increasing when viewed in apparent time. However, these papers do not definitively reject the role of the grammaticalization process, both noting that there are traces of these associated changes within the system. These traces are argued to be possible vestiges of earlier changes associated with grammaticalization that may have since become arrested (cf. Hopper and Traugott 1993:95). Thus:

“[I]t remains to be determined whether the synchronic stable patterns of GE variability [...] are in fact the product of grammaticalization processes that may have been operative at an earlier stage of the language, predating the time-span in our corpus.”

(Pichler and Levey 2011:462)

As such, Pichler and Levey (2011:462) plead for an “appropriate real-time benchmark” to unravel the unknown history of these changes. In the same spirit, Tagliamonte and Denis (2010:362) argue that comparative work on a broader inventory of pragmatic markers is necessary before making any sweeping conclusions.

This thesis addresses both these desiderata. My goal is to provide new insights into the mechanisms of linguistic change involved in the actuation and subsequent development of pragmatic markers. I do this first by extending the apparent-time span of Tagliamonte and Denis’ (2010) work on the Toronto English Archive (TEA) by considering new data from Ontario English. The data used in this thesis include the *Belleville 1975* collection of oral histories housed at the University of Toronto *Language Variation and Change Lab* (Hastings County Historical Society 1975)<sup>1</sup> and the *Farm Life and Farm Work Since 1890* oral histories (Archive of Ontario 1987),<sup>2</sup> a collection of nearly three hundred hours of oral history interviews conducted with elderly individuals in the 1980s in five regions of Ontario. In this thesis, I refer to these two collections together as the *Earlier Ontario English* data or EOE for short. The collections were recorded in the mid-seventies and mid-eighties respectively, and thus, the EOE provides an “appropriate real-time benchmark” to the TEA (recorded between 2002 and 2004) for studying the development of pragmatic markers. Second, in addition to GEs as in (1), which several variationists have examined (*inter alia* Cheshire 2007, Tagliamonte and Denis 2010, Pichler and Levey 2011, Wagner, Hesson, and Little 2014), I discuss epistemic parentheticals (EPs) as in (2), another set of pragmatic markers that, like GEs, can be conceived of as a variable system, although it has received less attention by variationists (though see Rodríguez Louro and Harris

---

<sup>1</sup>Permission to use the Belleville Oral History Project was obtained by Professor Sali A. Tagliamonte (University of Toronto) through arrangements with the Hastings County Historical Society in May 2007.

<sup>2</sup>Permission was granted from the Archive of Ontario to use the Farm Work and Farm Life Project. Copyright is held by the Canadian Crown and permission to publish excerpts has been granted by the Archive of Ontario

2013).<sup>3</sup>

- (1) a. Well yes. They had snack-bars you know *and stuff like that*.  
(EON/M/1898)
- b. When they had teas *and that* I served and I always made cakes for them  
*and that kind of thing*  
(BLV/F/1884)
- (2) a. So the raspberries were Cuthberts, cause *I think* that's all they grew then.  
(NIA/F/1904)
- b. *I guess* in my own tiny mind I figured this is how some snazzy cows must  
have lived.  
(TOR/M/1951)

These new diachronic and empirical insights position this thesis to critically assess the role of grammaticalization with respect to discourse-pragmatic variation and change and the developments of pragmatic markers.<sup>4</sup> I argue herein that the results of Tagliamonte and Denis (2010) and Pichler and Levey (2011) are not a remnant of bygone grammaticalization as hypothesized, but rather that the development of both GEs and EPs across the whole of the twentieth century did not involve gradual grammaticalization. Rather, both sets of pragmatic markers exhibit evidence for abrupt, not gradual, grammaticalization, lexical replacement, and competition between variants typically associated with morphosyntactic change.

---

<sup>3</sup>Metadata is included in parentheses after examples from corpora consulted. The first two or three letters indicate the region the speaker is from. These include NIA for Niagara Region, EON for Eastern Ontario, BLV for Belleville, TOR for Toronto, and YRK for York, England. Followed by speaker sex (M or F) and year of birth.

<sup>4</sup>In this sense, this thesis is inspired by Campbell and Janda's (2001:94) fostering of "a greater appreciation for empirical approaches" to grammaticalization.

In this introduction, I review the literature on grammaticalization, focusing on its past application to the understanding of the development of pragmatic markers. I also consider several critiques of grammaticalization theory. Many of the points made in these critiques form the basis of the arguments to be made here. In the last section of this chapter I provide an overview of the structure of the thesis and structure of the argumentation.

## 1.1 Grammaticalization

### 1.1.1 Grammaticalization: Phenomenon and Theory

The term grammaticalization has been used to describe at least two language phenomena and as a label for a specific framework that seeks to understand how linguistic material grammaticalizes (Hopper and Traugott 1993:1–2).<sup>5</sup> It is thus necessary to provide some precise definitions before continuing.

In one sense, grammaticalization has been used in linguistics to refer to the phenomenon of morphological or morphosyntactic material being used to overtly express grammatical meanings in a language's grammar. For example, we can describe English as a language that has a *grammaticalized* plural because the grammar of English contains a morpheme that marks a noun as being a plurality of entities. Other languages, for example Inuktitut, have a morpheme that marks dual number in addition to one marking the plural (Spalding 1998:110, cited in Compton 2011). In this sense, Inuktitut has grammaticalized dual and plural. This sense of the term grammaticalization is not the sense which I use in this thesis. Rather, I use the term grammaticalization to refer to a type of language change.

A vast body of literature has been produced over the last hundred years that dis-

---

<sup>5</sup>Grammaticalization is sometimes referred to as grammaticization and grammatization. Where these terms occur in quotes from the literature but strictly refer to the same notion, I will write 'grammaticalization' in an effort to avoid any potential confusion.

cusses the process of change that has been labeled grammaticalization, beginning with Meillet (1912/1926). In a survey of over thirty influential papers and monographs on the topic, Campbell and Janda (2001:107) narrow the “minimal”, “core” notion of GRAMMATICALIZATION to the phenomenon in which some linguistic element changes into some more grammatical element. This generally takes the form of an ‘autonomous’ lexeme or lexical collocation developing a grammatical function, or a grammatical element taking on a further grammatical role.

The phenomenon ‘whereby items become more grammatical through time’ is most often discussed by researchers working within a framework also referred to as ‘grammaticalization’ (Hopper and Traugott 1993:1–2). Following Campbell and Janda (2001:94), I will refer to this framework as GRAMMATICALIZATION THEORY. Grammaticalization theorists have proposed that grammaticalization (the phenomenon) exhibits a number of properties, principles, and parameters and that there are a number of mechanisms or associated changes that take place during the course of grammaticalization. For example, one proposed property of grammaticalization is that the change from lexical to grammatical occurs along a serial cline or pathway, as in (3) from Hopper and Traugott (1993:7) or more generally as in (4) (see also Heine, Claudi and Hünemeyer 1991, Bybee, Perkins, and Pagliuca 1994, Lass 1997, Hopper 1998).

(3) content item > grammatical word > clitic > inflectional affix

(4) lexical > grammatical > more grammatical

Many authors surveyed by Campbell and Janda (2001) have argued that grammaticalization is also a UNIDIRECTIONAL and irreversible process. That is, the change in (3) can only proceed rightward, and never leftward (*inter alia* Traugott 1988, Croft 1990, Bybee et al. 1994, Pagliuca 1994, Lass 1997, Haspelmath 1998). Furthermore, the

pathways of grammaticalization are argued to represent a GRADUAL process taking hundreds of years to complete and as such, the locus of attention is on the development of morphemes over time. Lastly, key to grammaticalization theory are the interrelated, associated changes or mechanisms that take place over the course of grammaticalization. Most prominent in the literature are the REDUCTION or attrition of a form's morpho-phonological shape and concomitant loss or BLEACHING of its semantic meaning, each mentioned in eleven definitions of grammaticalization surveyed by Campbell and Janda (2001). Some authors, particularly Traugott (1995:3), reject this "impoverishment" view, arguing that SEMANTIC CHANGE (rather than strictly bleaching) and PRAGMATIC STRENGTHENING or expansion are more common. That is, when a form grammaticalizes it gains new semantic meanings and/or pragmatic functions. This typically results in a form having a layering of meanings: an older, original meaning persists in tandem with the newer meaning (Heine 1997:6; Hopper 1991:22; Hopper and Traugott 1993:124). In addition to these phonological and semantic changes is DECATEGORI(AL)IZATION, or the loss of a defined morphosyntactic context for a morpheme (Heine 2003:579).<sup>6</sup>

There are many examples of grammaticalization in languages from around the world discussed in the literature that are argued to exhibit unidirectionality, gradualness, and evidence of the mechanisms of grammaticalization. A classic example is the development of the English *going to* future marker from a verb of motion. Along the way from (lexical) motion verb to (grammatical) aspectual marker, *going to* underwent the changes listed in Table 1.1, and grammaticalization theorists argue that all those changes, only together, constitute grammaticalization (e.g., Lehmann 1982:v; Heine 2003:579; Traugott 2003:643–4; Diewald 2010:19–20).<sup>7</sup>

---

<sup>6</sup>Increasing SUBJECTIFICATION is another tendency that is often associated with grammaticalization. As a form grammaticalizes, Traugott and König (1991:209) suggest that the meaning of the form becomes "increasingly situated in the speaker's subjective belief-state/attitude toward the situation." See (5).

<sup>7</sup>Some grammaticalization theorists summarize some of these mechanisms in terms of context expansion (Himmelmann 2004:31). That is, a form expands with respect to its syntactic position and

Table 1.1: Grammaticalization of *going to*.

Mechanism	Result
Phonetic reduction	From [gowɪŋ tu] to [gʌnə] (or even just [ə] as in <i>I'ma</i> )
Decategorialization	From V to Aux/ <i>v</i>
Semantic change	From motion to future temporal reference
Pragmatic expansion	From future temporal reference of motion verbs to future temporal reference generally

### 1.1.2 Grammaticalization of Pragmatic Markers

The grammaticalization of lexical morphemes into grammatical material as exemplified by *going to* is a prototype for grammaticalization theory. However, several linguists have generally maintained that the same grammaticalization theory provides an explanatory account of how pragmatic markers develop, and include many discourse-pragmatic changes under the same umbrella as changes characterized as lexical to grammatical (Brinton 1996, 2008; Thompson and Mulac 1991; Traugott 1995). Others identify changes from lexical to pragmatic, or pragmatic to more pragmatic as the related but independent process of “pragmaticalization” (Erman and Kotsinas 1993; Aijmer 1997).<sup>8</sup> However, most of the literature agrees that this distinction is only a matter of terminology (Cheshire 2007, Günthner & Mutz 2004, Diewald 2011). The only necessary adjustment to make to the core notion of grammaticalization is to expand what is meant by ‘more grammatical’ to include the expansion of pragmatic function. In other words, so long as the domain of pragmatics is included within the broader scope of what the grammar of a language includes, if a form expands from having a strictly lexical meaning to expressing a discourse/pragmatic/conversational function (i.e., becoming less lexical and more functional), it can be said to be grammaticalizing (Brinton 2008:52; Diewald 2011).

Like its morphosyntactic counterpart, grammaticalization theorists have demar-

---

its semantic-pragmatic meaning.

<sup>8</sup>Others argue that the development of pragmatic markers follows the process of lexicalization (e.g., Fischer 2007).

cated a gradual, unidirectional cline to the development of pragmatic markers, a pathway such as in (5) (Traugott 1982:256).

- (5) propositional/referential meaning > (textual/discourse meaning > expressive/interpersonal meaning)<sup>9</sup>

That is, in the course of the development of a pragmatic marker, one can observe that a form first has strictly referential meaning or truth-conditional meaning independent of the conversation/text in which it is embedded. At the next stage a form develops a textual or discourse meaning, relative to the text or discourse (e.g., marking turn-taking or a shift in topic). Lastly, expressive/interpersonal meanings are those meanings that express a speaker's attitude or feelings toward a proposition. For example, the aptly named "expressives" like English *fuckin'* and German *bloß* as in (6) do not contribute to the propositional meaning of the utterance, but signify the speaker's negative attitude toward the proposition itself.

- (6) a. I have to mow the *fuckin'* lawn.<sup>10</sup>

(Potts 2005:60)

- b. Wo liegt *bloß* meine Brille?  
Where lies *BLOSS* my glasses  
'Where (*on earth*) did put my glasses? (I have already looked everywhere)'

(Bayer and Obenauer 2011:468)

Among several examples of a change from propositional to textual to expressive, Diewald (2011:383) discusses the case of the German modal particle *eben*.<sup>11</sup> In the

<sup>9</sup>I put parentheses around the last two stages as more recently, the directionality of these stages in development of pragmatic markers has been relaxed (e.g., see Brinton 1996:275). The (short) path can thus be summarized as 'truth-conditional > non-truth-conditional' (Traugott and Dasher 2002:40).

<sup>10</sup>As Potts (2005:60) puts it: "The speaker probably bears no ill-will towards lawns, or his lawn. Rather the proposition that he must mow the lawn is what he seeks to disparage."

<sup>11</sup>The English equivalent of *eben* (i.e., *just*) seems to function similarly. Whether or not it underwent the same serial development is a historical question.



propositional stage, *eben* expresses a local-spatial relation of ‘even, plain, smooth.’ In the second stage, the form expresses the more textual meaning of temporal relation ‘right now’. In the third stage, *eben* expresses the fact that the speaker “held this opinion before and [...] hold[s] it and say[s] it now”, as in (7) from Diewald (2011:380).

- (7) Das ist eben keine gute Idee.  
 That is EBEN not-a good idea.  
 [DD: That just isn’t a good idea.]

Brinton (2008:52–53), summarizing her previous work and Traugott’s (Brinton 1996; Brinton and Traugott 2005; Traugott 1995, 2003), argues that the development of pragmatic markers is also subject to the mechanisms of grammaticalization. Pragmatic markers decategorialize from “more major to more minor word class membership” (i.e., from nouns and/or verbs to adverbials/conjunctions/particles). They undergo semantic change and pragmatic expansion as they change from having referential meanings to serving pragmatic functions (as discussed immediately above). And, while they do not typically undergo phonetic reduction, they often become fused (e.g., *kind of*, *sort of* > *kinda*, *sorta*).

### 1.1.3 Criticisms of Grammaticalization Theory

Critics of grammaticalization theory observe three major problems for the framework. First, there are multiple counter-examples to the unidirectional hypothesis, suggesting that grammaticalization is not irreversible. Related to this is that the “morpheme-centred” view of grammaticalizationists has led to assuming falsely that grammaticalization is a continuous, gradual process (Janda 2001:283). Lastly, whether or not the process of grammaticalization is a unique diachronic process, independent of other, already well-established types of linguistic change, is questionable. Many

of these criticisms have been summarized in Campbell and Janda (2001).<sup>12</sup>

### Grammaticalization is Not Unidirectional or Irreversible

As discussed above, a major finding in the grammaticalization literature is the strong tendency for morphosyntactic change to proceed from lexical to grammatical and not from grammatical to lexical. Some grammaticalization theorists align themselves with Haspelmath (1998:319) who observes that “it is an undeniable empirical fact that such changes [from grammatical to lexical] do not occur.” However, as Janda (2001:292) points out, more than eighty counter-examples have been found to the unidirectionality/irreversibility of grammaticalization.

Among English examples is the case of the preposition/particle *off* developing into a lexical verb *to off*, meaning ‘to kill’, perhaps originally from the particle verb *to kill off*.<sup>13</sup> Another example comes from the case of the German noun *After*, which today means ‘anus’ and is “perceived as offensive outside of medical jargon” (Janda 2001:300). This form, cognate with English preposition *after*, developed from the Old High German preposition/adverb *afar* ‘behind, after’ eventually to the euphemistic noun *aftero* ‘buttocks’. The grammatical prepositional/adverb form was lost by Modern German, leaving only the lexical noun; its meaning further specializing.

In addition to numerous other examples in the literature of grammatical elements developing into lexical elements, the case of *yada, yada, yada* is arguably an example of a pragmatic marker developing a contentful, lexical meaning. *Yada, yada, yada* is a pragmatic marker functioning in the same way as the more common variant *blah, blah, blah* that “indicates ‘there is more, but the details are irrelevant here’” (Overstreet 1999:156). Indeed, this pragmatic marker was the theme of an episode

<sup>12</sup>See also Diertani (2011) for another critical view of grammaticalization theory.

<sup>13</sup>While Janda (2001:300) states that *to kill off* is the likely source of *to off*, J.K. Chambers (p.c.) has suggested that *to bump off* is a more likely source: *to bump off* and *to off* both mean ‘to murder’ and *to kill off* means ‘to bring about the death of a fictional character’.

of the television show *Seinfeld*<sup>14</sup>, in which the main characters discuss the usefulness of *yada, yada, yada* for avoidance of “incriminating details about the ‘more’ that it indicates” (Overstreet 1999:156). In one scene, repeated in (8), George is concerned that his girlfriend Marcy’s use of *yada, yada, yada* covered up her having slept with an ex-boyfriend. Elaine then admits to having used the pragmatic marker to cover up a one night stand.

- (8) **George:** Listen to this. Marcy comes up and she tells me her ex-boyfriend was  
 over late last night, and “**yada yada yada**, I’m really tired today.” You  
 don’t think she **yada yada’d** sex. 1  
 2  
 3
- Elaine:** (Raising hand) I’ve **yada yada’d** sex. 4
- George:** Really? 5
- Elaine:** Yeah. I met this lawyer, we went out to dinner, I had the lobster bisque,  
 we went back to my place, **yada yada yada**, I never heard from him again. 6  
 7
- Jerry:** But you **yada yada’d** over the best part. 8
- Elaine:** No, I mentioned the bisque. 9

In addition to the pragmatic usage of *yada, yada, yada* in lines 2 and 7, George, Elaine and Jerry all use *yada yada* as a lexical verb, complete with tense inflection, in lines 3, 4, and 9 respectively. In this excerpt, *to yada yada* has the meaning of ‘to cover up details through discursive means, particularly by using *yada, yada, yada*’. Later in the episode, *yada yada* is also used nominally, as in (9).

- (9) **George:** All right, enough! Enough! From now on, no more **yada yadas**. Just  
 give me the full story. 1  
 2

<sup>14</sup>Mehlman, P. J. Franklyn (Writers), and A. Ackerman (Director). (1997). The Yada Yada [Television series episode]. In A. Ackerman (Producer), *Seinfeld*. Hollywood, California: Sony Pictures Television.

This subsequent nominal development is also important, as it proceeds in the opposite direction typical of English zero derivation, from nouns to verbs (as in *file* to *to file*).

Janda (2001:292–294) discusses several responses to these counter-examples. Some observe that certain counter-examples are only counter-examples of some versions of the unidirectionality hypothesis or that the analyses that describe the counter-examples are inadequate (Lehmann 1982; Haspelmath 1998). Others modify or specify what exactly is meant by unidirectionality. For example, Herring (1991:253) narrows the unidirectional nature of grammaticalization to the observation that speakers tend to use concrete concepts to refer to abstract notions, and thus, grammaticalization is unidirectional with respect to abstraction. Still others suggest that since grammaticalization is defined as a change from lexical to grammatical, it is inherently unidirectional (Lehmann 1982).<sup>15</sup> In the same spirit, Heine (2003:582–583) argues that many of these counter-examples are either i) idiosyncratic, ii) the result of some other social, psychological, or cultural force, such as hypercorrection, or iii) involve “extreme language contact” and the formation of creoles. Furthermore, Heine (2003:582) observes that there has been no complete reversal of a grammaticalized form, from lexical to grammatical back to lexical.

The unidirectional hypothesis does not play any vital role in the argumentation in this thesis and I remain agnostic as to its validity. However, as Janda (2001:293) observes, there is a major problem with the kind of responses above. The hypothesis is not that there is a predominant direction of grammaticalization but rather that grammaticalization is unidirectional, from lexical to grammatical. The former hypothesis can tolerate a statistical minority of counter-examples to the theorized predominant direction but the latter must take every potential counter-example seriously.

---

<sup>15</sup>This of course is a tautology (Janda 2001:294). Such approach to unidirectionality is simply a description of the theoretical grammaticalization process and not a hypothesis about language change.

### Grammaticalization is Not Gradual or Continuous

One key proposal of grammaticalization theory is that changes along the clines in (3) through (5) are gradual.<sup>16</sup> In fact, the idea is explicit in Bybee, Perkins, and Pagliuca's (1994) definition of grammaticalization:

[G]rammatical morphemes develop gradually out of lexical morphemes or combinations of lexical morphemes [...] Included are changes in lexical morphemes by which some few of them become more frequent and general in meaning, gradually shifting to grammatical status, and developing further after grammatical status has been attained.

(Bybee, Perkins, and Pagliuca 1994:4–5)

One of the clearest discussion of gradualness and grammaticalization is in Traugott and Trousdale's (2010) recent volume:

Gradualness refers to the fact that most change involves (a series of) micro-changes, an issue which is sometimes overlooked in considerations of more general patterns of language change. As Brinton and Traugott (2005:150) observe, although change is sometimes understood (or at least formulated) as A>B, studies of gradualness in linguistic change attempt to uncover "the tiny local steps between A and B that the arrow '>' encompasses."

(Traugott and Trousdale 2010:23)

Another criticism of grammaticalization is that language change, insofar as grammaticalization theorists are concerned, cannot be gradual or continuous. That is, lexical morphemes do not "gradually shift[...] to grammatical status" as Bybee, Perkins, and Pagliuca (1994:4) argue. Janda (2001) discusses at length the inherently cross-generational discontinuity of language transmission. As Kiparsky (1968:175, quoted in Janda 2001:272) puts it "a language is not some gradually and imperceptibly

<sup>16</sup>Hopper and Traugott (1993:59–60) discuss Fries (1940) quantitative study of the grammaticalization of VO word order in the history of English which they argue "highlight the gradualness of the spread of changes." An decrease of OV order at the expense of VO order from 52.5 percent in the eleventh century to 40 percent in the fourteenth century to 1.87 percent in the fifteenth century is indeed a gradual change but it is unclear how competition between OV and VO word order is a case of grammaticalization, and not syntactic (parameter) change.

changing object which smoothly floats through time and space, as historical linguistics based on philological material all too easily suggests...[; r]ather, the transmission of language is discontinuous and a language is recreated by each child on the basis of the speech data it hears." Joseph (2004:63) summarizes this issue eloquently: "as linguists trying to understand change as something that speakers do (not something that happens to a language), we should not take a perspective on language change which a speaker cannot take (thus no trans-generational 'diachronic processes' – cf. Janda 2001)." In essence, morphemes and structures do not have a continuous lifespan of their own, independent of the mental grammars of its speakers, and thus cannot change gradually.

Janda (2001) does offer a way out for grammaticalization theorists: Labov's (1994) model of *GENERATIONAL CHANGE* (later elaborated in Labov 2001, 2012) may account for the apparent gradualness of linguistic change by means of discrete, directional decreases in lexical meaning over successive generations. However, to the best of my knowledge, there has been no substantial discussion in the literature about what the metric of "lexical" or "grammatical" is or should be.<sup>17</sup> Labov's (1994) model is used by sociolinguists who examine changes in linguistic variables and developments of variants within speech communities. Discrete changes from variant X to variant Y over time appear to be gradual because each successive generation can be shown to have a higher proportion of Y than X until finally Y is the only variant available. However, since grammaticalization theorists tend to focus on individual morphological forms and not linguistic variables, it is unclear how the naturally quantitative concepts of gradualness and continuity apply to non-gradient phenomena such as the change of a lexical element to functioning as a grammatical element (cf. Traugott

---

<sup>17</sup>In fact, the view of many grammaticalization theorists (e.g., Bybee 1988; Hopper 1998) is that "grammar and lexicon are not neatly separated or compartmentalized" (Torres Cacoullos and Walker 2009a:35).

and Trousdale 2010).<sup>18</sup> Brinton and Traugott's (2005:99) definition of grammaticalization comes close to providing for gradualness in such a change. Grammaticalization is:

the change whereby in certain linguistic contexts speakers use parts of a construction with a grammatical function. Over time the resulting grammatical item may become more grammatical by acquiring more grammatical functions and expanding its host-classes.

(Brinton and Traugott 2005:99)

In other words, to become more grammatical is either to gain more grammatical functions, or to serve a grammatical function in a wider range of linguistic contexts.

An explanation of the development of pragmatic markers from grammaticalization theory requires evidence of gradualness. For example, following Brinton and Traugott (2005:99), as pragmatic markers develop, there should be evidence that they acquire additional functions, or spread to a wider range of linguistic contexts sequentially, rather than wholesale. Imagine some highly grammaticalized pragmatic marker with multiple functions (or functioning in multiple contexts) in some language state  $B$  that developed out of some lexical material in language state  $A$ . If this development involved gradual grammaticalization, there ought to be 'micro'-stages between  $A$  and  $B$  in which the pragmatic marker served only a subset of functions (or occurred in only a subset of contexts). Gradual development via 'micro'-stages is not unlike Bailey's (1973) wave model of change as schematized in Figure 1.1.

According to the wave model, as a new linguistic rule (or form) replaces an older one, it does so by gradually expanding its operation to more and more environments

---

<sup>18</sup>A trend in grammaticalization theory is to examine constructional grammaticalization. Here, the focus is not on individual linguistic elements but rather "set[s] of forms which display similar properties, and which have developed in a particular set of ways over time" (Trousdale 2008:59). In many ways, such "a network of constructions" is not unlike a linguistic variable (Trousdale 2008:58). Hoffmann (2004:195) similarly observes that grammaticalization might "rely much less on the nature and context-dependent use of individual content words than previously assumed" and rather could "result in the establishment of constructional schemas whose slots can be filled with suitable lexical items."

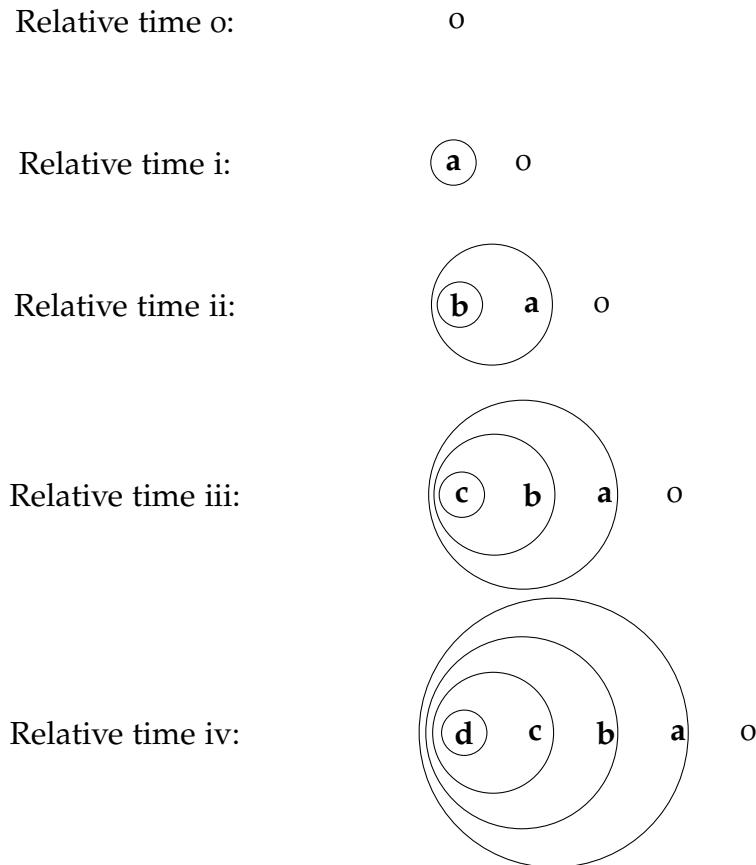


Figure 1.1: Bailey's (1973:68) wave model: "The simplest form of the wave model. (The letters represent successively later, or lighter-weighted, environments in which the rule operates.)"

(or contexts). As Bailey (1973:67) puts it "the isolect<sup>19</sup> generated by the operation of [a rule] in environment **a** [...] is prior to the one generated by the operation of the rule in environment **b** [...]; this is prior to the isolect generated by the rule in environment **c** [...]; and this is in turn prior to the isolect generated by the rule in environment **d**."

Bailey (1973:82) assumes that the frequency of an innovation and the rate of change (from one rule or form to another) in different contexts correlates with the relative time that the rule (or form) began to apply in that context: "what is quantitatively less is slower and later; what is more is earlier and faster." Kroch (1989) provides counter-evidence to this claim, arguing that "change proceeds at the same rate

<sup>19</sup>"Isolects are varieties of a language that differ only in a minimal way, say by the presence or weighting of a single feature in a rule, or by a minimal difference in rule ordering" (Bailey 1973:11).



in all contexts, and that, as far as one can tell, disfavoring contexts acquire new forms no later than favoring ones, though at lower initial frequencies.” This *CONSTANT RATE EFFECT* is specifically the case “if a single grammatical parameter is involved in a change and the mix of the two opposed settings is slowly changing over time in a given speech community” (Kroch 2001:720). In other words, the rate of change in different contexts is expected to be the same when the underlying change is abrupt and catastrophic since the choice between two competing grammatical options is what is changing and the contextual effects (i.e., favouring/disfavouring environments) are external to such competition.

Thus, the rate of change can be used as a diagnostic of gradualness (or abruptness). In the case of pragmatic markers, and in particular those that are undergoing increases in usage frequency relative to other competing forms (as is the case with both GEs and EPs), if the slope of the rise in frequency of innovative pragmatic markers is found to occur at different rates in different contexts, this can be taken as evidence for the gradualness of grammaticalization. On the other hand, even in the face of differential frequencies/probabilities of innovating pragmatic markers in different contexts, if change is found to be happening at a constant rate in each context, this can be taken as evidence that the underlying change from language state *A* to language state *B* (i.e., from lexical material to pragmatic marker) occurred abruptly, rather than gradually, context by context.

It is relevant to this thesis that the development of pragmatic markers could more logically be a gradual process than morphemes traditionally discussed in the grammaticalization literature. Whereas a morpheme is either lexical or grammatical in any given context, pragmatic markers are observed to be multi-functional and simultaneously polysemous (Aijmer 2002:3; Cheshire 2007:183). Single pragmatic markers are argued to be able to express several pragmatic functions at once including, for example, marking social solidarity and functioning in turn-taking. However, in Denis

and Tagliamonte (2014), we test this gradualness hypothesis for pragmatic markers and argue that the rise of the utterance final tag *right* in Toronto English does not exhibit gradual expansion of pragmatic functions. We operationalize the number of discourse contexts in which *right* appears as a measure of pragmatic expansion and although it appears that *right* began in Toronto English by occurring only in a restricted set of discourse contexts and then expanding to more (as Brinton and Traugott 2005 might predict), this apparent expansion is simply an artifact of the frequency with which *right* is used relative to other pragmatic markers that serve the same functions. When *right* was new and infrequent, it only occurred in our data in a few discourse contexts (those that highly favour its usage) precisely because its overall frequency was low. *Right* was available for use in other discourse contexts from its earliest stages, but its potential breadth is not observable due to its status as a minority variant. However, by the time *right* was the predominant variant, the pragmatic marker is observed to be used in all the possible discourse contexts in which its previous competitors were used. This argument is supported by a zero-inflated poisson regression model that indicates that the number of discourse contexts in which an individual used *right* is a function only of the number of times they used *right* and crucially not the speaker's age (used as a proxy for time). In other words, there was one abrupt change—the innovation of *right* as an utterance final tag—and not a gradual expansion.

Throughout this thesis, I will quantitatively test whether the development of pragmatic markers is gradual in a variety of ways.

### **Grammaticalization is Not an Independent Process**

The last criticism is that grammaticalization is not a process, but rather a result of other already established processes. Campbell (2001), Joseph (2001, 2004), Newmeyer (2001), and Roberts and Roussou (2003) all argue that the phenomenon of grammati-

calization is derivative of reanalysis, analogy, metaphor, metonym and regular sound change. As Joseph (2004:61) puts it:

I do not for a moment deny that there is a 'phenomenon' of grammaticalization; as I have argued elsewhere (Joseph 2001, 2003), however, I see it as a result, not a process, an epiphenomenon perhaps. I would say that we could just do with grammaticalization as a result, a product, and reserve the designation 'process' or 'mechanism' for the traditionally recognized sound change, analogy, reanalysis, and metaphorical extension; moreover, since social dimensions are taken into consideration (e.g., in contact situations, whether between dialects or languages), then borrowing (and all this last entails, like calquing) and hypercorrection also need to be added in, the latter being a powerful process/mechanism in change due to dialect contact.

(Joseph 2004:61)

I leave the reader with the above citations, but I will return to this criticism in more depth throughout.

Along this same line of reasoning, a close look at Hopper's (1991:22) influential principles of grammaticalization reveals that only one or two principles apply strictly to grammaticalization phenomena, despite others frequently being used as evidence for grammaticalization. Hopper's (1991:22) principles are:

- (10) a. *Layering*. Within a broad functional domain, new layers are continually emerging. As this happens, the older layers are not necessarily discarded, but may remain to coexist with and interact with the new layers.
- b. *Divergence*. When a lexical form undergoes grammaticalization to a clitic or affix, the original lexical form may remain as an autonomous element and undergo the same changes as ordinary lexical items.
- c. *Specialization*. Within a functional domain, at one stage a variety of forms with different semantic nuances may be possible; as grammaticalization takes place, this variety of formal choices narrows and the smaller number of forms selected assume more general grammatical meanings.

- d. *Persistence*. When a form undergoes grammaticalization from a lexical to a grammatical function, so long as it is grammatically viable some traces of its original lexical meanings tend to adhere to it, and details of its lexical history may be reflected in constraints on its grammatical distribution.
- e. *De-categorialization*. Forms undergoing grammaticalization tend to lose or neutralize the morphological markers and syntactic privileges characteristic of the full categories Noun and Verb, and to assume attributes characteristic of secondary categories such as Adjective, Participle, Preposition, etc.

*Layering* is apparent in all cases of sociolinguistic/intra-speaker change whether or not grammaticalization is involved.<sup>20</sup> As Weinreich, Labov, and Herzog (1968) observe “[n]ot all variability and heterogeneity in language structure involves change; but all change involves variability and heterogeneity.”<sup>21</sup> *Specialization* is predicted to occur in any change, where one option wins out over other obsolescing options (Kroch 1994). While *persistence* of an original lexical meaning certainly is observed to occur as a form grammaticalizes, it will also necessarily be the case if a form is not grammaticalizing in any way. *De-categorialization*, while clearly a process that happens during grammaticalization, by Hopper’s description only seems to restate the definition of grammaticalization itself: a change from more lexical categories (nouns and verbs) to more grammatical categories (adjective, preposition). The process of *divergence* does seem to be unique to grammaticalization—also by definition.

Hopper (1991:21) admits this himself: these principles “also characterize aspects of change in general, and are not distinctive for grammaticalization.” He goes on

---

<sup>20</sup>Traugott (1995:21) makes the same observation: “Layering or coexistence of variants is a characteristic of all change.”

<sup>21</sup>Layering has been interpreted in two ways: 1) layering of polysemies of a single form (i.e., grammaticalized and non-grammaticalized forms) 2) layering of homonymy/variants of a variable.

to say that they should be viewed as potential “diagnostics of grammatical forms and constructions out of already available material, and also of different degrees of grammaticalization, where grammaticalization has already recognizably proceeded.” While this is true, we must be careful not to use these principles as the *only* evidence that something is undergoing grammaticalization.

## 1.2 Overview of the Argument

In this thesis, I test two major criteria that must be met in order to argue that grammaticalization theory is the best model for the development of pragmatic markers. First, because grammaticalization is claimed to be a gradual process, as a pragmatic marker increases in frequency, a concomitant gradual expansion of functions or linguistic contexts in which the marker can appear is expected. Second, as a pragmatic marker develops, increasing in frequency, we should expect to find evidence of the mechanisms of grammaticalization. Phonetic reduction, decategorialization, semantic change, and pragmatic expansion should all increase as a form’s frequency increases.<sup>22</sup> Furthermore, because grammaticalization is “of a composite nature” (Diewald 2010:19) and only these mechanisms occurring *together* can constitute grammaticalization, I will also argue that evidence for only a subset is insufficient to conclude that the development of pragmatic markers proceeds as predicted by grammaticalization theory.

In Chapter 3, I examine the innovative GE *and stuff*. In the earliest part of my data, *and stuff* is incipient. Thus, I am able to track its development from inception: that is from the moment it was initially used as a GE in this community. As the form increases in frequency across the twentieth-century, becoming the predominant variant

---

<sup>22</sup>Given the multigenerational timespan of my data, this should be observable whether increases in frequency trigger grammaticalization (Bybee 2003) or grammaticalization triggers increases in frequency (Mair 2004).

of the GE variable system, I examine the diagnostics of the mechanisms of grammaticalization that Cheshire (2007) uses in her study of GEs. However, here, these diagnostics are tested across real and apparent time from inception. I am thus able to confront Pichler and Levey's (2011) concern that early work on the grammaticalization of GEs (e.g., Tagliamonte and Denis 2010) may have captured previous, arrested grammaticalization, given the possible observation of vestiges of grammaticalization mechanisms. What I find is that apparent phonetic reduction of *and stuff* is the result of an independent process of morphological clipping that effects *all* GEs at a constant rate, not just the innovative one as it, ostensibly, gradually grammaticalizes. Likewise, I find no evidence for decategorialization. Rather, *and stuff* rises in frequency at a constant rate regardless of its antecedent. There is some suggestion in my data that the youngest speakers use *and stuff* semantically bleached of its set-referential meaning but there is no evidence it has pragmatically shifted. Taken together, there is evidence of only one out of four possible mechanisms over the course of development of *and stuff*. Thus, it seems that it did not grammaticalize as conceived of by grammaticalization theory.

Several grammaticalization theorists have made quantitatively testable hypotheses about the grammaticalization of EPs. In Chapter 4, I examine the EP system, focussing on *I think* and *I guess*, and put these hypotheses to the test. While Thompson and Mulac (1991) (among others) argue that EPs grammaticalized from lexical main clause subjects and verbs, Kaltenböck (2013) and Kärkkäinen (2003) argue that *I think* is further grammaticalizing, being bleached of its meaning as an epistemic marker and coming to be used with a strictly discourse-structuring function. As such, the form is predicted to become more frequent in clause-initial position where it performs this more advanced function. While *I think* increases in frequency relative to other EPs, it does so at the same rate in both clause-initial and non-clause-initial position. Were *I think* being semantically bleached, its frequency should increase at

a faster rate in clause-initial position. Rodríguez Louro and Harris (2013) and Torres Cacoullos and Walker (2009a) argue that as *I think* grammaticalizes *I* and *think* become increasingly fused, such that *I think* operates as a single unit. However, the presence and absence of intervening material within *I think* in my data is stable across time. Thompson and Mulac (1991) argue that the fact that *I think* seems to implicate a stronger commitment to the truth of its complement proposition than *I guess* does is a case of *persistence* of the lexical meanings of the predicates of these EPs. Indeed, the epistemic/doxastic strength of these two variants, as operationalized by the subject of the complement clause, persists across time. However, as I suggest above, *persistence* is also the null hypothesis for situations in which no grammaticalization is occurring. Furthermore, while *I think* increases across time, the size and direction of this effect is fixed, indicating a constant rate of change in each context, and thus, together with the constant rate of change in clause-initial and non-clause-initial position, suggests that the development of *I think* is characterized by abrupt change, not gradualness.

These conclusions and the analyses they are based on rest on the methodological foundation of variationist sociolinguistics. Furthermore, what allows me to test the hypotheses of previous researchers is data with a rich temporal range. In the next chapter, I discuss both the methodology of variationist sociolinguistics and go into detail about the corpora from which my data come.

## Chapter 2

# General Methodology and the Earlier Ontario English Data

*The fact of language change is a given; it is too obvious to be recorded or even listed among the assumptions of our research. Yet, this fact alone – the existence of language change – is among the most stubborn and difficult to assimilate when we try to come to grips with the nature of language in general as it is reflected in the history of language.*

*(Labov 1994:9)*

This chapter has two broad goals. First, I describe the VARIATIONIST METHOD, the general methodological approach that this thesis employs to investigate the development of pragmatic markers. Variationists who have studied pragmatic markers have encountered a number of difficulties that require us to problematize the nature of the core abstract unit of variationist work, the LINGUISTIC VARIABLE. These issues and potential resolutions are discussed in what follows. The second goal is to introduce the data that contributes a real-time component to the understanding of the development of pragmatic markers in Canadian English. The contemporary data is the Toronto English Archive (TEA), a corpus of sociolinguistic interviews recorded with native residents of Toronto, aged nine to ninety-two between 2003 and 2006 (Taglia-



monte 2006b). The real-time benchmark comes from two oral history projects, one housed at the University of Toronto Language Variation and Change Lab (*Belleville 1975*), and the other, recently compiled for variationist work, is housed at the Archive of Ontario (*Farm Work and Farm Life Since 1890*). In this thesis, these collections together will be referred to as the Earlier Ontario English data (EOE).

## 2.1 Pragmatic Markers, Variation, and Change

### 2.1.1 The Variationist Method

For half a century now, variationist sociolinguistics has concerned itself with three facts (Tagliamonte 2006a:5–7): 1) languages vary in an orderly way; 2) languages change; 3) and language conveys more than the truth-conditional semantics of the morphosyntactic composition of lexical items.<sup>1</sup> To examine these three facts, it is necessary to study language in use (Tagliamonte 2006a:8–9); in particular, to study the natural vernacular or “everyday speech” of a speech community in an accountable, scientific, and quantitative way (Labov 1966a; Sankoff 1973/1980:54).

The VARIATIONIST METHOD, the general methodology employed in this thesis, is just such an accountable, scientific, and quantitative method for the study of linguistic variation (Chambers 2009; Labov 1972; Poplack and Tagliamonte 2001; Tagliamonte 2006a, 2012). Common to all variationist work is 1) the use of appropriate data, 2) Labov’s (1972:72) Principle of Accountability, and 3) a method for modelling the choice process (Poplack and Tagliamonte 2001). I will address each of these in turn.

First, data are required to be sufficient with respect to “quality, quantity, and representativity” (Poplack and Tagliamonte 2001:88). Qualitatively, the data must be

---

<sup>1</sup>With respect to this last point, by hearing only a short amount of speech from a given individual, one can determine personal, stylistic, social, sociocultural and sociological information about the speaker and her speech community (Chambers 2009:2–10).

useful for researchers. Good sound quality of recordings and accurate, detailed, and consistent transcription are necessary for accountable empirical research of language in use (Beal, Corrigan, and Moisl 2007), especially when dealing with variation in the discourse-pragmatic domain (Pichler 2010). Furthermore, data from speakers must be of sufficient length (Poplack and Meechan 1998:129). Although the occurrence of variation is common within a paragraph of text or thirty-second soundbite, much more data is required to establish patterns through statistical means. Lastly, data must be representative of the speech community under investigation. This means that researchers interested in gender or socio-economic differences must include representation from across the gender spectrum or from individuals representing a range of social classes. As Labov (1966a:9) observes “the speech of many individuals appears as studded with oscillations and contradictions, and it is only when it is placed against the over-all framework of social and stylistic variation of the speech community that we can discern the regular structural pattern that governs this behavio[u]r.”

Second, data analysis must conform to the PRINCIPLE OF ACCOUNTABILITY (Labov 1972:72). To study language variation, one cannot look at the distribution of a single form in isolation.<sup>2</sup> The Principle of Accountability requires that researchers consider the whole pool of possible variants, not just the form of interest. That is, we must include in our analyses every context in which a variant could have occurred, whether it did or not.<sup>3</sup> As such, the main target of investigation of variationist sociolinguistics is the linguistic variable (Tagliamonte 2012:3). The LINGUISTIC VARIABLE can be simply

---

<sup>2</sup>While this is true of all scientific investigation, it is unfortunately something we are neurologically biased against. Modern neuroscience observes our tendency to focus on *illusory correlations*—the fact that “[o]ur forebrains have evolved to notice co-occurrences of events but not a lack of occurrence” (Levitin 2014:253). Furthermore, Levitin (2014:255) discusses the concept of *denominator neglect*, our neurological bias toward focussing on the raw number of occurrences without reference to the size of the denominator.

<sup>3</sup>Conforming to the Principle of Accountability is often the main methodological difference between variationist sociolinguistics and corpus linguistics (Tagliamonte 2012:19). This is a crucial point since much of the work on the development of pragmatic markers comes from a corpus linguistics tradition and many observations about grammaticalization rely on the results of corpus investigations that simply look at the forms without considering their place in the broader system of variation. This point will be discussed more below.

defined as an abstract set of two or more ways of saying the same thing. This basic definition has evolved over the fifty-year history of variationist sociolinguistics.<sup>4</sup> The linguistic variable was first described by Labov (1966a):

Whereas the linguistic variant is a particular item—a morph or a phone—the variable is a class of variants which are ordered along a continuous dimension and whose position is determined by an independent linguistic or extra-linguistic variable.

(Labov 1966a:13)

Tagliamonte (2012:4–5) provides a more detailed definition of the linguistic variable that incorporates developments subsequent to Labov’s (1966) first description (in particular, see Sankoff [1973/1980] and Sankoff [1988] regarding points 4 and 5).

(1) The Linguistic Variable (Tagliamonte 2012:4–5):

1. two or more ways of saying the same thing;
2. an abstraction;
3. made up of variants;
4. comprising a linguistically defined set of some type:
  - a phoneme
  - a lexical item
  - a structural category
  - a natural class of units
  - a syntactic relationship
  - the permutation or placement of items
5. although its delineation can be at any level of the grammar, the variants of the variable must have a structurally defined relationship in the grammar;

---

<sup>4</sup>See Wolfram (1991) for an early discussion of the subtle development of the linguistic variable.

6. they must also co-vary, correlating with patterns of social and/or linguistic phenomena.

With respect to points 4 and 5, the linguistic variable has evolved from a phonological concept that uses the semantic equivalence of morphemes (or “the requirement of sameness” [Lavandera 1978:174]) as the metric of “saying the same thing”, to a concept that is applicable to all domains of language (Sankoff 1973/1980; Tagliamonte 2012:15). For example, Tagliamonte and Denis (2010) use the metric of functional-structural equivalence in defining the variable context of general extenders (see §2.1.2 and below for further discussion).

Prior to Labov’s (1963, 1966a, 1966b) foundational work, instances of two or more ways of saying the same thing in a language were considered to be instances of free variation (Chambers 2009:13). As Labov (1966a:2) observes, earlier studies of “massive free variation” that exist in speech communities, although valuable in their phonetic and phonological analyses, fell short in their discussion and explanation of what we know now to be linguistic variables. Labov (1966a) quotes Hubbell (1950) with respect to the variable pronunciation of /r/ in New York City:

The pronunciation of a very large number of New Yorkers exhibits a pattern in these words that might most accurately be described as the complete absence of any pattern. Such speakers sometimes pronounce /r/ before a consonant or a pause and sometimes omit it, in a thoroughly haphazard fashion.

(Hubbell 1950:48)

Labov (1966a) rebuts:

Is it possible that such a large part of the speech system of New Yorkers is the product of chance factors? The idea goes against the grain of our conception of language as the most highly structured type of human behaviour.

(Labov 1966a:3)

Labov's question foreshadows what would become Weinreich, Labov and Herzog's (1968) axiom that amid massive heterogeneity in language, there is order. That is, although linguistic variables are not categorical, and therefore not deterministic, they are also not random or free. The solution to the analytical "standstill" of variation was solved by introduction of a different kind of data, data that was both quantitative and that represented the vernacular of the speech community (Labov 1966a:2). Armed with this new kind of data, quantitative questions about variation could be asked, such as 'what factors govern the occurrence of one variant of a variable over another' (Labov 1966a:3)? And indeed, Chambers (2009:18), in his introductory volume on sociolinguistics states that the "primarily empirical task" of variationist sociolinguists is to identify the social, stylistic, cognitive and linguistic factors that correlate with the choice between these two (or more) ways of saying the same thing. The foundational empirical observation of variationist sociolinguistics is that in all cases of variability, the realization of one variant over another is subject to probability given an array of independent factors.

Given the multifactorial nature of the variant choice process, variationist sociolinguistics ideally involves statistical modelling. Although distributional results often reveal most of a variable's story, multiple regression techniques are able to tease apart complex interactions and take into account statistical regularities and irregularities by considering all conditioning factors simultaneously. For most of its history, variationist sociolinguistics has relied on the Variable Rule Analysis program (VRA). For

all intents and purposes VRA is equivalent to binomial logistic regression (Bayley 2002:124).<sup>5,6</sup>

Today, sociolinguists divide the history of variationist work into three stages, or ‘waves’ (Eckert 2012). Each wave has privileged a different set of questions with respect to language and society. In the first wave (the foundational work of sociolinguistics), linguistic variation was correlated with sociological categories: sex, social class, occupation, income, ethnicity, and age (e.g., Labov 1966b; Wolfram 1969; Cedergren 1973; Sankoff and Sankoff 1973; Trudgill 1974b). In the second wave, locally-constructed categories, determined on the basis of ethnographic analysis of social networks and communities of practice were the focus of intensive research (Milroy 1980, Cheshire 1982, Eckert 1989).<sup>7</sup> The third wave has moved away from static categories and recognizes the dynamic use of linguistic variation by speakers to express, maintain, and manipulate their social identities. The focus has been on linguistic variation as stylistic practice (Bucholtz 1999, Kiesling 1998, Cutler 1999, Zhang 2005).

Developing in tandem with these well-recognized three waves has been a separate line of research which focuses on linguistic variation as part of the human language faculty. Since its start, generative linguistics has asked questions about competence rather than performance (Chomsky 1965). The object of study is the mental grammar consisting of the abstract rules and representations that constrain a

---

<sup>5</sup>It is my understanding that forty years ago when VRA was first written, there were no easily accessible computer programs that could perform binomial logistic regression. David Sankoff wrote VARBRUL and specifically tailored it to the needs of sociolinguists who needed to model binary choice variables and multiplex internal and external conditioning factors. Subsequent computational advances make logistic regression easily available and while other fields, having more recently begun to implement this technique, use general statistical platforms such as R, sociolinguists have stuck with what they already know. However, in the last five years, more and more sociolinguists have moved toward using general platforms, particularly R (Johnson 2009; Tagliamonte and Baayen 2012).

<sup>6</sup>In this thesis, I by and large make use of binomial logistic regression. Where necessary, mixed-effects models (also called hierarchical or multilevel models) will be implemented in order to account for individual speaker fluctuations.

<sup>7</sup>Some of the earlier, first wave, studies included discussion of locally constructed categories. For example, Labov (1963) explained the lowering of /ay/ in the Martha’s Vineyard speech community in terms of a speaker’s attitude toward the community, while Trudgill (1972) argues that middle-class Norvician men exhibited a high degree of working-class features as an expression of virility that working-class men embody.

language. In order to ask questions about language competence, generative linguists must ask questions about an idealized speaker “unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his[/her] knowledge of the language in actual performance” (Chomsky 1965:3). Studying language outside of social/performative context is a contentious issue for non-generativists, however it is an undeniable fact that the scholastic enterprise that adopted Chomsky’s approach has expanded our knowledge of language diversity and universality tremendously. In essence, “inquiring into the idealization” works (Chomsky 1980:25).<sup>8</sup>

As Chambers (2009:27) observes, critics of Chomsky take issue with his early proclamation that “observed use of language ... surely cannot constitute the actual subject matter of linguistics” (Chomsky 1965:4). This quote must be put into historical context however. At the time, variationist sociolinguistics had hardly begun. There was not yet an empirically sound methodology for studying language in use until Labov (1966a:4) argued that variation is “an essential part of general linguistics.” Chambers (2009:28) notes however that Chomsky (1980) “tacitly revises” the place of language variation in linguistics proper.<sup>9</sup>

The core difference between variationists and generativists is that they ask fundamentally different questions about language. Variationists ask questions about performance while generativists ask questions about competence. However, around the turn of the twenty-first century, variationist methods have been used to investigate competence (*inter alia* Adger 2006; Adger and Smith 2005; Anttila 2002; Boersma and Hayes 2001; Cheshire 2005a; Henry 1995; Kroch 1989; Nevins and Parrott 2010). These papers take intra-speaker variation, the kind familiar to sociolinguists, as a

---

<sup>8</sup>For a recent review of major discoveries of generative linguistics (including the hierarchical structure of language and locality effects) see Pesetsky (2013).

<sup>9</sup>Chomsky (1980:24–25) acknowledges that matters of linguistic variation and sociolinguistics are simply beyond the scope of what the ideal speaker methodology is intended to study.

proxy for inter-speaker/cross-linguistic variation.<sup>10</sup> For example, the cross-linguistic variation exhibited between Icelandic and Norwegian in (2) and (3) is constrained by the same mental-grammar mechanisms as what determines the intra-speaker variation in (4) from Adger (2007).

- (2) a. *Ég var hér.*  
I was here.  
(Icelandic)
- b. *Jeg var her.*  
I was here.  
(Norwegian)
- (3) a. *Við vorum hér.*  
We were here.  
(Icelandic)
- b. *Vi var her.*  
We were here.  
(Norwegian)
- (4) a. There was one night we *were* lyin at anchor.  
There was one night we were lying at anchor  
(Buckie English; Adger 2007:512)
- b. We played on 'at beach til we *was* tired, sailin boaties, bilin  
We played on that beach until we were tired, sailing boats, boiling  
whelks.  
whelks.  
(Buckie English; Adger 2007:512)

Icelandic, like English, marks a number distinction on the past tense of the copula. Agreement on the copula with a first person, singular subject is realized as *var*, while agreement with a first person, plural subject is realized as *vorum*. In Norwegian, the entire paradigm is level and the copula is realized as *var* regardless of the mor-

<sup>10</sup>Although Recapitulation Theory is generally considered defunct in evolutionary biology (Bleichschmidt 1977:32), it seems that with respect to linguistic variation ontogeny recapitulates phylogeny.



phosyntactic properties of the subject.

In Buckie English, there is an analogical contrast such that individual speakers vary with respect to the realization of the past tense of the copula *be*. The same speaker might say *we were* (with Icelandic-like morphological contrast) at one time but *we was* (with Norwegian-like levelling) at another time. I will return to this discussion in chapter 5.

Addressing questions about how the human language faculty is structured so as to allow for such variation within a single speaker is another flourishing direction for variationist sociolinguistics. When we combine this approach with the independent factor of time, the mechanisms of language change can be observed, described and potentially explained (Labov 1966a; Sankoff and Thibault 1981; Weinreich, Labov, and Herzog 1968). If we examine language data from different points in time, we can track changes in REAL TIME. However, diachronic data is not always necessary for investigating change. Researchers have noted that individuals' grammars tend to stabilize in adolescence (e.g., Chambers 1992; Gardner, Denis, Brook, and Tagliamonte 2013; Labov 2001; Tagliamonte and D'Arcy 2009; Tagliamonte and Molfenter 2007). The result of this stabilization is that the grammar of a fifty-year-old in 2014 is in many respects equivalent to the grammar of a twenty-year-old in 1984. Thus, change can also be investigated if variationists correlate language variation with a speaker's age. This is referred to as APPARENT TIME data (Bailey, Wikle, Tillery & Sand 1991). This thesis makes use of data that constitute both real and apparent time data. The Toronto English Archive (TEA) was collected between 2003 and 2005 and includes speakers born between the 1920s and the 1990s, giving a seventy year apparent-time range (Tagliamonte 2006b). The *Farm Work and Farm Life Since 1890* collection contains components collected in 1984, while the *Belleville 1975* collection was collected in 1975.<sup>11</sup> The latter two corpora contain interviews with speakers born

---

<sup>11</sup>These corpora are discussed in detail in §2.2.

between 1879 and 1919. Thus, the comparison of data in TEA and EOE provides a real-time comparison with an apparent-time span of over one hundred and twenty years.<sup>12</sup> This thesis will capitalize on this unique attribute by making judicious use of age and year of birth to investigate the development of pragmatic markers.

### 2.1.2 Linguistic Variation at the Extra-Sentential Level

The variationist method has been applied to variables that function at the extra-sentential level, e.g., narrative structure, information structure, pragmatic markers, (for two recent reviews, see Macaulay 2013; Pichler 2010). At different points in the history of variationist sociolinguistics, there have been peaks of interest in such variation and subsequent advances in methodological practices (Cheshire 2005b, 2007; Cheshire, Kerswill, and Williams 2005; D’Arcy 2005; Dines 1980; Dubois 1992; Erman 1995; Laberge 1978; Lavandera 1978; Lemieux, Fontaine, and Sankoff 1987; Macaulay 2002; Meyerhoff 1994; Sankoff, Thibault, and Bérubé 1978; Sankoff, Thibault, Nagy, Blondeau, Fonollosa and Gagnon 1997; Stubbe and Holmes 1995; Vincent 1992; Vincent and Sankoff 1992). While Pichler (2010:582) critiques previous projects for a “lack of coherent set of methodological principles,” those projects that have most successfully examined extra-sentential variation have used variationist methodology. To be more precise, variationist work on discourse-pragmatic variables is successful, precisely because it follows variationist methods that have been advanced since the seventies. While Pichler (2010) advocates that researchers take seriously issues of corpus construction, circumscription of the variable context, quantification of data, and qualitative analysis “to ensure reliability, generalizability and comparability” when working with discourse-pragmatic variation, it is critical to remember that these are methodological issues that have been taken seriously in all variationist work—at any

---

<sup>12</sup>This time span includes all six broad generations of the twentieth century: The Lost Generation (1883–1900), The Greatest Generation (1901–1924), The Silent Generation (1925–1945), The Baby Boom Generation (1945–1964), Generation X (1964–1981), Generation Y (1981–1999).

level of the grammar.

First, Pichler (2010) shows that many previous studies of discourse-pragmatic variation vary widely in the type of corpora examined. Different corpora can easily represent different styles and contexts of speech. Since pragmatic markers are reported to vary widely with respect to style, context, topic, and interlocutors, researchers must be cautious of comparative analysis across corpora that have not been built under identical circumstances (Pichler 2010:584–585). Pichler (2010:585), echoing earlier variationists (see Beal, Corrigan, and Moisl 2007 for example) recommends that all corpora construction follow the same design principles. Of course, this is not always possible and thus, corpus builders should include appropriate textual meta-data about their corpora including social, stylistic, topical, attitudinal and discourse factors. In §3.2, I make the case that both components of the EOE are appropriately comparable to TEA.

Perhaps the most important reason that relatively little work has been done on pragmatic variation within the variationist tradition is that pragmatic variables do not easily fit into the mold of what a linguistic variable is. Researchers must contend with what is meant by “two or more ways of saying the same thing” with respect to pragmatics. This is the “first and perhaps foremost challenge to be confronted” when working in the discourse-pragmatic domain (Tagliamonte 2012:269). Phonological and morphological variables can rely on semantic equivalence. For example, the presence or absence of the *-s* morpheme in Detroit African American Vernacular English are two semantically equivalent ways of expressing plural number (Wolfram 1969). However, in her critique of applying variationist methods to levels above and beyond phonology, Lavandera (1978:171) observes that “it is inadequate at the current state of sociolinguistic research to extend to other levels of analysis of variation the notion of sociolinguistic variable originally developed on the basis of phonological data.” For example, the variation between active and passive constructions exam-

ined by Weiner and Labov (1983) is not a case of semantic sameness. One variant has an agent; the other does not. After a discussion of cases such as these, Lavandera (1978:181) proposes “to relax the condition that the referential meaning must be the same for all the alternants and substitute for it a condition of functional comparability.” Another early study of discourse-pragmatic variation, Lemieux, Fontaine, and Sankoff’s (1987:382) study of the discourse-pragmatic marker *tout* and its variable realizations, recognized the theoretical and practical problems of semantic equivalence. As Lavandera (1978) did, Lemieux *et al.* (1987:382) argue that since the variants share “une fonction commune au niveau du discours” (“a common function at the level of discourse”–DD), this “justifie leur regroupement sous une même variable” (“justifies grouping them under a single variable”–DD).

Indeed, syntactic variation can rely on functional equivalence. Although *have to* and *must* have subtle truth-conditional and structural differences, the core function of both is to express deontic modality in many varieties of English (Tagliamonte and D’Arcy 2007b). This metric of functional equivalence was also used by early studies on general extenders (e.g., Dubois 1992).

However, Sankoff, Thibault and Bérubé (1978), in their discussion of how to delimit lexical and semantic variation, relax even functional equivalence. Among their case studies of variables in Montreal French is an examination of “a cluster of very general terms, meaning ‘thing’ or ‘something’” (Sankoff *et al.* 1978:24). The majority forms, *chose* and *affaire*, are mostly substitutable for one another, but each has its “differential utility in different contexts (i.e., for different functions)” (Sankoff *et al.* 1978:39). Indeed, a major characteristic of pragmatic markers is their multifunctionality. Thus, Pichler (2010), following the example set by Cheshire (2007), Pichler (2008), Tagliamonte and Denis (2010), and Pichler and Levey (2011), recommends that at least some discourse-pragmatic variables be based on structural equivalence

rather than semantic or functional equivalence.<sup>13</sup> Pichler (2010:590) argues that structural definitions of the linguistic variable provide the analyst with a way of dealing with “diachronic meaning changes and synchronic polyvalence whilst still ensuring that ‘*the variants are in some way the same, have something in common*’ (Dines 1980:19, italics in original).”<sup>14</sup> Since tracking, describing, and explaining the potential development of new pragmatic functions of GEs and EPs is a key question of this thesis, I circumscribe the variable context of these variables following this procedure (see §3.2 and §4.3 below).

A third area of methodological concern for variationists is how to quantify variants. Linguistic variables, by definition, are a closed set. However, as Pichler (2010:593) observes, some pragmatic markers are not easily thought of as a closed set (i.e., a variable system) because, despite the best efforts of analysis of structural equivalence, the identification of all potential variants of a variable is not easy.<sup>15</sup> An alternative method for quantifying pragmatic variables is using normalized frequencies of forms for individuals (or groups of individuals), usually with respect to a word count measure.<sup>16</sup> However, this method can be problematic for cross-corpora comparisons. Pichler (2010:595) illustrates the profound effect that different transcription protocols can have on the number of words in a corpus. Her data, from north-east England, has as many as 276 707 words if false starts, filled pauses, and minimal responses are included, cliticized morphemes are separated, and compounds are hyphenated (all

---

<sup>13</sup>Pichler (2013:31) revises this term to “derivational equivalence”.

<sup>14</sup>This approach is consistent with grammaticalization theorists who tend to consider the grammaticalization of a structural unit (e.g., *be going to*) rather than a single lexical item. For example, Himmelmann’s (2004) approach to grammaticalization considers not a form in isolation (the “element-based view”), but rather the form within its syntagmatic context.

<sup>15</sup>For example, if GEs function as punctors, (i.e., to bracket units of discourse and not add any new semantic-pragmatic information [Vincent and Sankoff 1992]) they may overlap with the functions of utterance-final tags such as *right* or *eh* (Denis and Tagliamonte 2014a). Since both appear utterance finally, it is possible to analyze both variants under the same umbrella. This thesis separates the two features for practical and theoretical reasons. In addition to clearly constituting subsets, my questions are about the development of particular variants through time.

<sup>16</sup>Other idiosyncratic measures such as tokens per line (Dubois 1992) or tokens per minute of speech have been used as well (Meyerhoff 1994).

transcription conventions of various corpora) and as few as 240 187 words if these conventions are not followed. As I will discuss in the next session, the TEA and EOE were transcribed using identical transcription protocols, so I do not anticipate this affecting the results of this thesis.<sup>17</sup> In the next two chapters I will argue that both GEs and EPs can be conceived of as a variable system and thus can be analyzed in terms of proportions and probabilities of variants (relative to other variants) just like other linguistic variables.<sup>18</sup>

Lastly, because of the key role of multifunctionality, Pichler (2010) recommends that researchers use qualitative methods to assess every token to code for pragmatic function, and include this in modelling. Pichler and Levey (2011) follow this method for their analysis of GEs, though Tagliamonte and Denis (2010) do not. In §3.4.4 I discuss the possibility of such an analysis for GEs but conclude that a fine-grained approach to coding for pragmatic function requires an *a priori* theory about the direction of development which may artificially impose a sequential view. That said, stressing Pichler's (2010) methodological concern more generally, only through qualitative analysis can such a decision be made.

### 2.1.3 The Variationist Method and Grammaticalization

The variationist approach to grammaticalization and change can be contrasted with most other methodological approaches with respect to two features: temporal scope of the data and the Principle of Accountability.

#### Temporal Scope of Data

Language change in action can be observed by looking at the immediate past, particularly at successive generations by way of the apparent time construct (Bailey *et*

---

<sup>17</sup>See Beal, Corrigan, and Moisl (2007) and Poplack (2007) on the importance of inter- and intra-corpora consistency.

<sup>18</sup>See also D'Arcy (2005) for a discussion of treating discourse *like* as a variable system.

*al.* 1991, Sankoff 2005). This is unlike most work on grammaticalization that takes either a synchronic or historical perspective. Researchers working from a synchronic perspective intuit grammaticalization from patterns in synchronic varieties, arguing for what a previous language state must have been (cf. Erman 1995, Thompson and Mulac 1991), while historical linguists consider diachronic correspondences that are sometimes separated by hundreds of years. However, as Heine (2002:83) points out:

Unfortunately, most processes of grammaticalization that have been studied so far are conventionalized and buried in history – to the extent that much of what happened on the way from A to B is no longer historically clearly recoverable.

Furthermore, diachronic correspondences can obfuscate the actual changes. As Andersen (1972:12, as quoted in Janda 2001:268) critiques:

[A] diachronic correspondence can be used to summarize a ... change ... but since it merely defines a relation between equivalent units at different times, it does not by itself reveal whether it summarizes a single change or a series of changes. Nor does a diachronic correspondence state how the change(s) took place which gave rise to the correspondence ... [I]t should be clear that the study of diachronic correspondences cannot substitute for the study of actual ... changes.

The variationist approach considers the sequence of parent-to-child transmission of change in progress and thus provides researchers with a perspective of the trajectory of changes and the development of innovative variants within a variable system. Given Labov's (1972:275) Uniformitarian Principle that "the forces operating to produce linguistic change today are of the same kind and order of magnitude as those which operated in the past", this methodology allows unique access to the mechanisms of change, as systems develop. Almost ten years prior to Tagliamonte and Denis (2010), Janda (2001:318) pleads for work on grammaticalization within the variationist method:

What we are most sorely missing, then, are sociolinguistically oriented studies which would compare the ways in which elements apparently

undergoing grammaticalization are used by speakers vis-à-vis those on whom they model their behavio[u]r. Until investigations of apparent grammaticalization in progress today begin to outnumber the current plethora of work on alleged instances in medieval, ancient, and prehistoric times, the latter will continue to be subject to [...] criticism (explicit or implied) that [has] been—and/or should be—level[l]ed against non-uniformitarian methods.<sup>19</sup>

Ten years later Tagliamonte (2012) notes that:

The type of large-scale survey data that is often used by sociolinguists – representative community-based samples of spoken vernacular language data – can provide a picture of varying stages in the grammaticalization process if the data come from a broad enough age sample (say 9–90, i.e. 80 years) and providing that whatever change is under investigation is happening in such a way as to be captured by this time span.

(Tagliamonte 2012:88)

A number of variationists have done just this, examining grammaticalization in the context of sociolinguistic corpora (Tagliamonte 2006a:77; Poplack and Tagliamonte 2001; Schwenter and Torres Cacoullos 2008). The tools and methods of variationist analysis have been put to the task of diagnosing grammaticalization in a number of ways. For example, semantic bleaching has been tested by examining the magnitude of effect of semantic constraints (Tagliamonte 2006a:77). That is, at an earlier stage, we expect a strong effect for semantic constraints on variants that are purportedly undergoing grammaticalization, while at a later stage, we expect weakening and eventual neutralization of these effects. An example of this is Poplack and Tagliamonte's (2001:230–1) analysis of the *be going to* future marker; the non-significance of the lexical content effect (i.e., whether *be going to* co-occurred with a verb of motion or not) in Ottawa English was taken as an indication that this variety had "proceeded farther along the grammaticalization path" than other varieties they examined, in which the effect was significant. Tagliamonte (2012:91) provides a step-by-step guide for implementing such a study:

---

<sup>19</sup>Coincidentally, Poplack and Tagliamonte (2001), a monograph that examines grammaticalization just as suggested by Janda, was published that same year.



- Gather information about the lexical source and early usage patterns.
- Project the development pathway from early to later stages e.g. *going to*: human → animate → inanimate.
- Operationalize factor groups historically implicated in the change.
- Include each factor group in a statistical analysis.
- Partition the data to reflect early to later stages of development based on external (sociogeographic or socioeconomic ecology) e.g. conservative dialect/variety → innovative dialect/variety. Compare and contrast across partitions using comparative sociolinguistic methods.

In other cases, the reverse result has been argued as diagnostic of Hopper's (1991) notion of persistence (and thus grammaticalization). Where semantic constraints, particularly those that can be related to the source, lexical meaning of the grammaticalizing form, retain their magnitude and direction over time, this has been taken as indicating persistence (Aaron 2010:9; Poplack and Malvar 2007; Poplack and Turpin 1999; Tagliamonte 2002, 2012; Torres Cacoullos and Walker 2009b). For example, Torres Cacoullos and Walker (2009b:331) suggest that if the *be going to* future marker was consistently found to be favoured with motion verbs, this semantic harmony could be taken as evidence for persistence.<sup>20</sup>

However, because the diagnostics of bleaching and persistence are mirror images, it can be potentially problematic to use the presence of either as an indication that grammaticalization is underway. This is particularly problematic when there is no time dimension (or proxy of time such as a cline of dialect conservativeness) in the study. Torres Cacoullos and Walker (2009b:332) recognize the dialectic nature of these principles in their study of the English future temporal reference system. The fact that bleaching and persistence are "twin grammaticalization principles" can lead to superfluous results without clear hypotheses (Torres Cacoullos and Walker 2009b:332). Thus, they hypothesize that persistence would be the case for the newer variant *be*

---

<sup>20</sup>Torres Cacoullos and Walker (2009b:331) immediately note that in other studies (including Poplack and Tagliamonte 2001) *be going to* either disfavours motion verbs or is not constrained by verb type, and thus, this could be taken as evidence for bleaching.

*going to*, while a higher degree of bleaching was expected for the older variant *will*. It is important to emphasize that such a hypothesis is only possible when studying grammaticalizing forms as part of a broader sociolinguistic variable system (i.e., the linguistic variable) and can only be tested using the methodological and statistical tools of variationist sociolinguistics.

### Principle of Accountability

An important methodological difference between grammaticalization theory and variationist theory, is that when grammaticalization theorists take a quantitative approach, this usually takes the form of ‘sideways’ calculations.<sup>21</sup> An example of such sideways calculation is Thompson and Mulac’s (1991:323) approach to the frequency of the *I think* EP.<sup>22</sup> Their Table 7 is replicated here in Table 2.1.

Table 2.1: Table 7 from Thompson and Mulac (1991:323).

	<i>-that</i>	<i>+that</i>	Total
<i>I think</i>	599 (92%)	54 (8%)	653 (100%)
Other	513 (81%)	121 (19%)	634 (100%)

Such calculations are ‘sideways’ in the sense that proportions are calculated from side to side, not top to bottom.<sup>23</sup> That is, the denominator of the fraction from which a proportion is calculated is the total frequency of a variant, not the total frequency of the context. While the focus of investigation of most quantitative studies of grammaticalization is the linguistic form (unto itself) and how frequently the form occurs in different contexts, variationists are concerned with how variants are embedded within the broader grammar (cf. Weinreich et al.’s 1967 *Embedding Problem*). While the pattern in Thompson and Mulac’s (1991:323) table turns out to give us the same story if proportions are calculated from top to bottom (i.e., *I think* is favoured in the

<sup>21</sup>Thanks to A. Dinkin for the term “*sideways* calculation.”

<sup>22</sup>Note that Thompson and Mulac (1991) are concerned with the constraints on complementizer deletion, rather than on the realization of EPs. So, for their purposes, this calculation is not ‘sideways’.

<sup>23</sup>Of course, this depends on the orientation of the table.

–*that* contexts and disfavoured in the +*that* contexts), as shown in Table 2.2, this is not necessarily always the case.

Table 2.2: Table 7 from Thompson and Mulac (1991:323), revised.

	– <i>that</i>	+ <i>that</i>
<i>I think</i>	599 (54%)	54 (31%)
Other	513 (46%)	121 (69%)
Total	1112 (100%)	175 (100%)

As Tagliamonte (2012:19) argues, analysts should not “simply count the number of times something occurs because this does not tell you very much unless you know how many times that something occurs in the body of material.” The importance of Labov’s *Principle of Accountability* is clearly demonstrated by Tagliamonte (2012:20) with the quotative *be like* and the rate at which it introduces different types of quotative content. Compare Tables 2.3 and 2.4.

Table 2.3: Count of all quotative types with *be like* as a quotative (Tagliamonte 2012:Table 1.1).

	%	N
Internal dialogue	24.1	186
Direct speech	69.0	533
Sound/gesture	2.8	22
Hypothetical	3.8	30
Writing	0.3	2
Total number of <i>be like</i>		773

Table 2.4: Distribution of *be like* according to type of quotative, as a proportion of all quotatives. (Tagliamonte 2012:Table 1.2).

	%	N
Internal dialogue	69.4	186/258
Direct speech	55.8	533/955
Sound/gesture	55.0	22/40
Hypothetical	39.0	30/77
Writing	11.8	2/17
Total number of <i>be like</i>		773
Total number of quotatives		1357

When we examine *be like* in isolation from all other quotatives as in Table 2.3, it appears as though we have a linguistic device that is used primarily to introduce direct speech. The form is also sometimes used to introduce internal dialogue and only very rarely introduces other kinds of quotative content. However, when we consider *be like* as a variant of a variable quotative system by counting not just *be like* but all quotative tokens, as in Table 2.4, we see “diametrically opposed results” (Tagliamonte 2012:20). Now it appears that the primary function of *be like* is to introduce internal dialogue: 69.4 per cent of quotatives that introduce internal dialogue were *be like* quotatives. Even more striking is the difference in frequency between the two tables for quotatives introducing sounds and gestures (2.8 to 55.0 per cent). The fact that in Table 2.3 *be like* appears to be used mostly to introduce direct speech is simply an artifact of an independent property of quotative use generally: quotatives typically introduce direct speech. In fact, if we take the denominators for the N column in Table 2.4 and divide each by the total number of quotatives, we see that the distribution for all quotatives is not unlike the distribution for *be like* in Table 2.3 (20, 70, 3, 6, 1 per cent respectively). This is what is meant by “accountability”. Only by examining variants *accountably* are we able to determine how particular variants function in particular contexts.

I now move on to a detailed discussion of the data utilized in this thesis and the issues surrounding diachronic analysis and data comparability.

## 2.2 The Earlier Ontario English Data

Data used for variationist investigation come from a wide variety of sources including letters (van Herk and Poplack 2003; Dollinger 2008; Nevalainen, Raumolin-Brunberg, and Mannila 2011), diaries (Dollinger 2008), proceedings of parliament (Jankowski 2013), magazines (Jankowski 2013), computer-mediated communication

(Tagliamonte and Denis 2008; Buchstaller, Rickford, Traugott, Wasow, and Zwicky 2010), television (Tagliamonte and Roberts 2005; Dion and Poplack 2005) and radio (Heffernan 2007). However, vernacular speech is the data *sine qua non* in the field of variationist sociolinguistics. When other types of data (such as those above, and particularly written data) are examined, there is often explicit acknowledgement of how the data relates to or reflects the spoken vernacular:

[M]ost written records of interest in this context **represent a speech act**: either a genuine, historical one that took place at a specific time and place, recorded but indirectly in writing, or a perhaps fictional but necessarily characteristic one, rendering speech forms that a typical member of a given speech community might have uttered with some degree of likelihood, representative of the everyday communication in this community.

(Schneider 2002:67)

The idea is that [the Ottawa Repository of Early African American Correspondences] **approximates the vernacular.**"

(Fuller Medina 2012)

In the construction of text-based, historical corpora, dialogue in fiction and drama has been considered by Biber *et al.* (1993: 3–5) to be "**reflections of casual face-to-face conversation.**"

(Jankowski 2005:2)

The patterns of language use that we have documented in [Instant Messaging] do not support the view that users are 'simplifying' language to make communication easier or more efficient; in fact, users are less concerned with efficiency than with expressivity, seeking to make the language **both look as well as 'sound' more like informal talk.**

(Jones and Schieffelin 2009:109)

Capturing and recording spoken vernacular data is no simple task. How does one "observe the way people use language when they are not being observed" (Labov 1972:61)? This OBSERVER'S PARADOX has been a "fundamental problem" of the variationist method since Labov's earliest work (Chambers 2009:19). To solve this problem, the SOCIOLINGUISTIC INTERVIEW was developed. This technique, first employed

by Labov (1963), asks speakers about particular topics of discussion “which recreate strong emotions [they have] felt in the past, or involve [them] in other contexts” while at the same time allowing speakers to lead the conversation (Labov 1972:209). This method can reliably elicit natural, unmonitored, vernacular speech.

For the historically-oriented researcher, a different paradox exists. If we wish to investigate the vernacular of the past, we tend to be limited to written records, but written records tend to be limited to non-vernacular registers. Despite best efforts, written representations of speech are simply not the real vernacular. Poplack and St-Amand (2007:729) “add a cautionary note to those already expressed by scholars attempting to reconstruct spoken vernaculars from written texts. [...] There is no substitute for real-time [spoken] data in the diachronic study of linguistic change.”

A partial solution to this DIACHRONIC OBSERVER’S PARADOX lies with oral histories. In the narrow, modern sense, oral histories are “recorded interviews with individuals about the past” used to glean “historical, cultural, and sociological data from first-person accounts” (Canadian Oral History Association Online).<sup>24</sup> The modern oral history methodology arose from multiple fields in the late 1930s, when historians, sociologists, journalists, publishers and the military all realized that ‘normal folk’ had valuable, interesting, and important knowledge of events, society, and culture that should be recorded and archived (Sharpless 2006). These first-hand narrative accounts reflect unique individual perspectives and give a voice to the people who experienced history when that history was their present. As W. T. Couch (1939:ix,xi), an early champion of the oral history method, observed: “with all our talk about democracy, it seems not inappropriate to let the people speak for themselves.”

Since the goal of oral history recordings is to capture genuine, authentic first-hand accounts, the data is comparable to sociolinguistic interviews.<sup>25</sup> One example of oral

---

<sup>24</sup>Before the advent of writing, all history could be considered “oral history”. This thesis considers the narrow definition only.

<sup>25</sup>In some senses, sociolinguistic interviews are oral histories. Recent fieldwork carried out by Sali Tagliamonte and the University of Toronto Language Variation and Change Lab in rural Ontario has

histories as sociolinguistic data is the *Récit du Français Québécois d'Autrefois*, a corpus composed of recordings that closely “resemble spontaneous conversation” of Non-Mobile, Older Rural Québécoise, collected by folklorists Luc Lacourcière and Carmen Roy in the 1940s and 1950s (Poplack and St-Amand 2007). Poplack and St-Amand (2007) systematically show how the oral history-like data in this corpus exhibits the inherent variability of the speech community recorded and is therefore comparable to modern, sociolinguistic corpora. In fact, the use of oral histories has a long history in variationist sociolinguistics. For example, the Ex-Slave Recordings have played a central role in the African American Vernacular origins debate—one of the most controversial and contentious topics in the history of the field. Beginning with a volume in the early nineties (Bailey, Maynor, and Cukor-Avila 1991), this collection of oral histories has provided data for a multitude of research projects that sought to situate contemporary AAVE in a diachronic, real-time context (e.g., Mufwene, Rickford, Bailey and Baugh 1998; Poplack 2000; Poplack and Tagliamonte 2001). The primary importance of oral histories for variationists is that they provide just such a historical perspective.<sup>26</sup> This thesis directly takes advantage of this aspect of oral histories in the investigation of an earlier state of the pragmatic markers of Ontario English. I now turn to a description of the constitution of the Earlier Ontario English data.

### 2.2.1 Constitution of the Earlier Ontario English data

The EOE is composed of two parts: the *Farm Work and Farm Life Since 1890* (FWFL) Oral History Collection (Archives of Ontario RG 16-200) and the *Belleville 1975* Oral

---

captured a large number of narratives recorded during the sociolinguistic interviews. These stories about local history and culture have been compiled and given back to the communities involved (Tagliamonte, Chen, Chin, and Maddeaux 2011; Tagliamonte 2010–2013).

<sup>26</sup>Oral histories have become increasingly important for linguists working with First Nations languages in Canada as well (Keren Rice p.c.). In the most unfortunate cases, some now extinct languages only exist (aurally) on tape.

History Collection (Hastings County Historical Society 1975). I will discuss these collections in turn.

The *Farm Work and Farm Life Since 1890* oral history project was collected in the mid-1980s by Alan Brookes of the University of Guelph History Department with the support of the Ontario Ministry of Agriculture and Food in commemoration of the two hundredth anniversary of European settlement in Ontario in an effort to document the lifestyle and working conditions of early 20<sup>th</sup> century farmers. Between 1984 and 1987, 154 interviews were conducted with 155 speakers. In particular, the researchers interviewed elderly Ontarians who were born and raised on farms in five regions of Ontario. Speakers were from various townships in five regions: Dufferin County, Eastern Ontario (now United Counties of Stormont, Dundas, and Glengarry and Prescott and Russell United Counties), Essex County, Niagara Region (the former Lincoln and Wentworth Counties, which are now part of Regional Municipality of Niagara and City of Hamilton respectively) and Northern Region (Temiskaming District).

For each region, one student research assistant conducted the interviews. The interviews were related to farm life and farm work. The Archives of Ontario's description of these records include the following topics: types of farms and crops, farm houses and barns, chores and farm routines, school and childhood memories, family life, leisure time, and adult involvement in the community. Interviews began with the interviewer stating something along the lines of (5).

- (5) I'd like to ask you a number of question and chat a little bit about what it was like to grow up in rural Ontario in the early years of the century.

(Interviewer, Niagara)

The original audio-cassettes and textual records associated with the project were



transferred to the Archives of Ontario from the Ministry of Agriculture and Food and are currently stored at the Archives.<sup>27</sup> After stumbling upon this collection while browsing the audio collections at the Archives of Ontario in late 2011, I had these audio-cassettes digitized and began transcribing them following the standard transcription protocol utilized at the University of Toronto Language Variation and Change (UofTLVC) Lab (Tagliamonte 2006a).

Table 2.5 presents the distribution of interviewees by sex and birth year in the five regions. In four of the five regions, there are roughly twice as many male speakers as female speakers,<sup>28</sup> though in the Northern region there is one more female than male.

Table 2.5: Distribution of interviews in FWFL by region, speaker sex and age.

Region	Sex	1890	1900	1910	1920	Total
Dufferin	Male	4	8	5	1	18
	Female	2	6	1	0	9
Eastern	Male	2	8	9	1	20
	Female	1	2	5	0	8
Essex	Male	1	15	2	0	18
	Female	4	6	3	0	13
Niagara	Male	3	12	9	0	24
	Female	1	7	4	0	12
Northern	Male	1	10	5	0	16
	Female	4	9	4	0	17
Total		23	83	47	2	155

The *Belleville 1975* Oral History collection was acquired by the University of Toronto Language Variation and Change Lab, under the supervision of Sali Tagliamonte,

<sup>27</sup>The Farm Work and Farm Life in Ontario Since 1890 oral history project is Ontario Government Record Series RG 16-200.

<sup>28</sup>This is likely the result of there being fewer women than men who both grew up on a farm and continued to live on a farm into adulthood. As Crerar (1999:46) observes:

The inheritance of land was a patriarchal prerogative and women's access to the fruits of the land came through their relations to men—fathers, husbands and sons. This legally dependent relationship greatly restricted the options of women who wished to remain single and farm on their own and made all women's futures—if they wished to remain in farming—dependent upon marriage.

in the summer of 2008 as part of a SSHRC research grant (Tagliamonte 2007–2010) through partnership with the Hastings County Historical Society.<sup>29</sup> The original reel-to-reel tapes and transcripts were digitized and copies were returned to the Society. The original interviews were conducted by students as part of a federal Opportunities for Youth project in 1975. A total of 60 interviews with prominent residents of Belleville and the surrounding Hastings County area were recorded in an effort to record first-hand narratives about the history of the region, particularly in the early part of the twentieth century, through the depression, and into the Second World War. Table 2.6 shows the distribution of speakers by sex and age. Unfortunately, there was no recorded metadata for the speakers, so birth dates had to be inferred from the interview itself or sought out by other means.<sup>30</sup> A large number of interviewees' birth dates remain unknown. At least nine speakers but as many as 26 are not native to Hastings County and therefore all 26 are excluded from the corpus for present purposes.

Table 2.6: Distribution of interviews in *Belleville 1975* by speaker sex and age.

Sex	1879	1890	1900	1910	1920	Currently Unknown	Total
Male	1	5	9	3	1	13	32
Female	4	4	5	0	0	14	28
Total	5	9	14	3	1	27	60

Transcripts were completed as part of the original project and these were made available to the UofTLVC Lab. However, these documents were not authentic representations of the speech recorded. For example, compare (6), which represents the text, as transcribed in 1975 to (7) representing a more true to speech transcript (following the UofTLVC Lab protocol).

<sup>29</sup>Gerry Boyce, former president of the Hastings County Historical Society, informed Sali Tagliamonte the principal investigator of the SSHRC project about existence of the reel-to-reel tapes and arranged for access.

<sup>30</sup>In some cases, I consulted archival records, including birth certificates, marriage certificates, death certificates, obituaries, and census records to determine the age and birth place of speakers.

(6) **Original Belleville 1975 transcripts:**

- a.
- AD:**
- How are they involved in the EMO?

**Int.:** Yeah, just the general background.

**AD:** Well, the Hastings and Prince Edward regiment is a local militia regiment and it's been in this area from 1800s and Hastings County in 1804 and Prince Edward County and anything that is a national effort becomes involved in, because militia is an army of the people as against a regular army.

- b. **AD:** That's kind of the river watching that goes on all the time and we see quite a bit of it in the paper because it's the only visible thing of what we're doing at that time towards the emergency measure act. But it's sometimes quite critical, you know. I've seen the water come up, I think it was six feet in forty minutes, oh, several years ago and the chairman of the conservation authority and myself and the engineers took the footbridge and you know, said do we close the downtown or do we say go. We asked the police to have people move their cars out of the parking area and they did so and they had to get the trucks even down, the tow trucks and I looked around the one car left and I said, "what idiot left his car there?" because the water was about that high up on us, it was mine and I forgot about it.

(7) **Transcribed under UofTLVC Lab Protocol:**

- a. <033> How are they involved in the E-M-O? </033>

<2> Yeah, **just the-** just the general background. </2>

<033> **And their relation gen-** well, **uh-** the Hastings and Prince-Edward regiment is a local militia regiment **eh** and it's been **in to- in- been and**

been in this area from **eighteen-hundreds** and Hastings-County in **eighteen-oh-four** and Prince-Edward-County and anything that is a national effort becomes **uh in which** the nation is involved in- **it becomes involved in** it because the militia is **a-** an army of the people as against a regular army **eh.** </033>

- b. <033>That's kind of the river watching that goes on all the time and **you see ah** quite a bit of it in the paper because **ah um-** it's the only visible thing of what we're doing at that time towards emergency measures **I guess.** But **it-** it's **ah** sometimes quite critical, you-know. **We've seen the ah-** I've seen the water come up **ah, um,** I think it was six feet in forty minutes, **ah, oh,** several years ago **and-** and the chairman of the conservation authority and myself and the engineers **stood on** the footbridge **and-** and you-know, said **that-** "Do we say close down downtown or do we say go." **Um** we asked the police to have people move their cars out of the parking area and they did so and they had to get the trucks even down, the tow trucks and **ah** I looked around the one car left and I said, "What idiot left his car there?" because the water was about that high up on it. It was mine. </033>

< 1 > <laughter/> < /1 >

<033> I'd forgot about it.</033>

In bold in (7), are all of the changes and additions that were required to make the original transcript compliant with the UofTLVC Transcription Protocol. Changes include XML mark-up, hyphenating proper names and compounds, spelling out numbers, and representation of false starts, partial words and hedges. This excerpt also exemplifies the pruning of discourse that plagues the original transcripts. Some whole

turns of phrase are left out, and most crucially for this thesis, many pragmatic markers, such as *I guess* and *eh* above, are not represented. For these reasons, it was necessary to re-transcribe the interviews with a secondary transcription procedure in order to capture exactly what is said on the recordings.

### 2.2.2 Construction/Compilation

This thesis makes use of a subset of the 205 interviews from these two collections of oral histories. My focus is on three of the six regions: Belleville (from Belleville 1975) and Niagara Region and Eastern Ontario (from FWFL). Figure 2.1 shows these three regions as they are situated in Ontario.

Furthermore, I limit the analysis of these collections to only speakers born prior to 1920,<sup>31</sup> with a roughly equal representation of men and women born in the 1890s, 1900s, and 1910s. In the end, this leaves thirteen speakers from Belleville, fourteen speakers from Eastern Ontario and sixteen speakers from Niagara. In total, the corpus contains 202574 words, roughly comparable to a similar corpus built from oral histories, the *Récit français québécois d'autrefois* (Poplack and St-Amand 2007).

These speech communities have been specifically chosen due to their settlement histories. Belleville, Eastern Ontario, and Niagara Region were all settled and founded by the United Empire Loyalists in the late 18<sup>th</sup> century during the first wave of Canadian settlement (Chambers 1998). The Loyalists were mostly American citizens who remained loyal to the British Crown during the American Revolution. In return for their loyalty, some 45000 to 70000 were granted land in British North America, mostly in the maritime provinces or Lower Canada (now Quebec) (Fryer 1980:307). The remaining 8000 or so men, women and children migrated to new settlements along the St. Lawrence River, Lake Ontario, or in Niagara, or farther west in Upper Canada (now Ontario) (Fryer 1980:307; French 2006:54). The Loyalists are considered

---

<sup>31</sup>The oldest speaker in the Toronto English Archive was born in 1916.

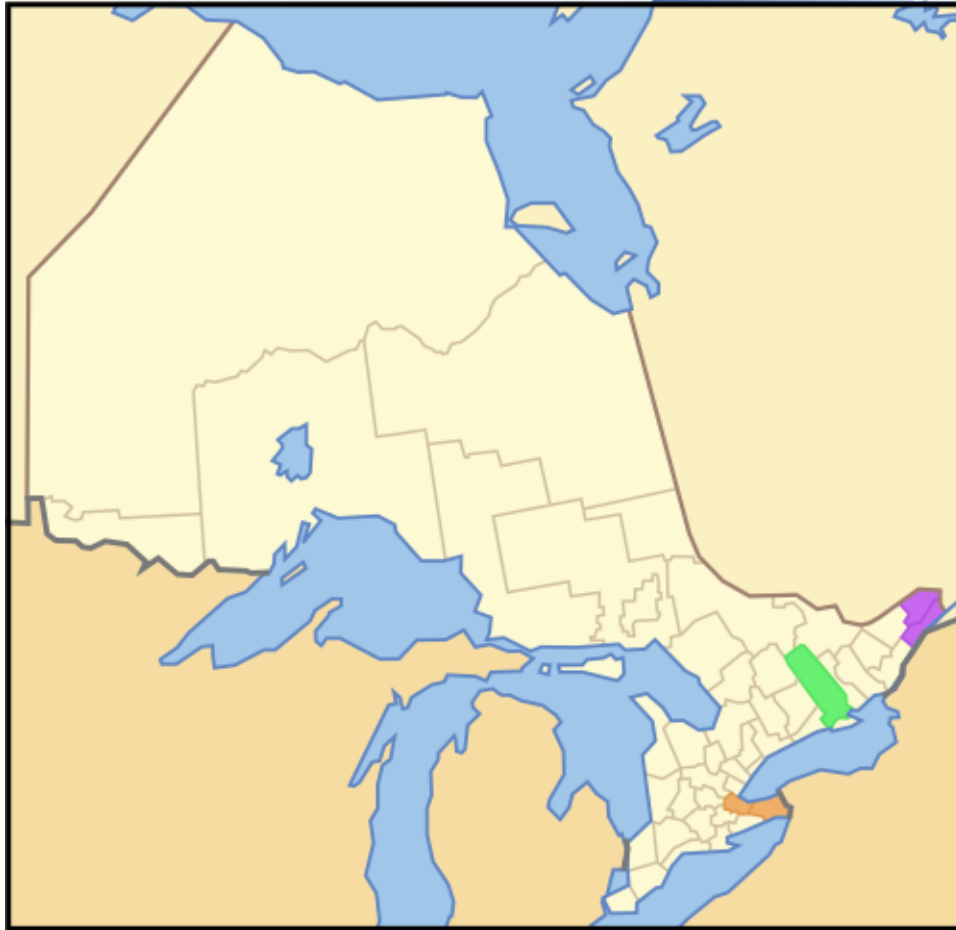


Figure 2.1: The three regions in the Earlier Ontario English data: Belleville (green), Eastern Ontario Region (purple), and Niagara Region (orange).

the founding population of English-speaking Ontario. In keeping with Zelinsky's (1992) Doctrine of First Effective Settlement, these first settlers of Ontario exerted "many subtle and largely unintentional dictates on those who succeed them" including their language (Chambers 1998:xii).<sup>32</sup> Thus, the Belleville, Eastern Ontario, and Niagara Region speech communities represent not only a lineage that can be directly traced back to Ontario English's founding speech community, but also a link in the chain of transmission of this variety, two or three generations deeper into the past than TEA represents.<sup>33</sup> In the remainder of this section I discuss the settlement histo-

<sup>32</sup>See also Mufwene's (1995) founder effect.

<sup>33</sup>The other communities included in FWFL have a more recent history, settled primarily by Irish,

ries of these three communities and describe the demographics of the speakers used in this study.

### **Belleville**

When the Loyalists first began to settle Ontario immediately after the American Revolution, the southern part of what is now Hastings County had not been surveyed and was not included in the initial land grants. However, five to ten years after the revolution, the first settlers, Loyalists and the children of Loyalists, who moved in from nearby Kingston, began to arrive (Boyce 2008:57). Prior to this settlement there was little European activity in the area (Boyce 2008:17).<sup>34</sup> The city of Belleville was first founded as the village of Meyers' Creek, named in honour of the prominent Loyalist Captain John Walden Meyers, who in 1790 established a mill and dam near the mouth of the Moira River (Boyce 2008:28; French 2006:54). At this time, around fifty Loyalist settlers took up permanent residence; nearly fifty years later the population of the area had grown to 1700 people, and by 1876, surpassed 10000 (French 2006:ii). Today, many residents of Belleville and the communities around the Bay of Quinte are descendants of the original Loyalist families (French 2006:54).

Many of the speakers interviewed in the *Belleville 1975* project, although having a prominent role in the history of Belleville, were not born and raised in the area. For this project, only speakers who were born and raised in southern Hastings County (or Prince Edward County, which neighbours Hastings to the south) were included in the analysis. One speaker was born in New Brunswick but moved to Hastings County at the age of two. In total, thirteen speakers from Belleville are examined in this thesis, six females and seven males. Their birth dates range from 1879 to 1914. Table 2.7 presents the demographic information of each of these speakers and Figure

---

Scots, and British immigrants during what Chambers (1998) refers to as the second wave of immigration to Canada.

<sup>34</sup>A short-lived Catholic mission possibly existed around where Belleville is today in the late seventeenth century, though the exact location is unknown (Boyce 2008:21).

Table 2.7: Demographics of the Belleville 1975 speakers.

Sex	Birth Year	Birth Place	Additional information
F	1879	Picton	Came to Belleville in 1918
F	1884	Belleville	
F	1897	Lonsdale	
F	1898	Belleville	
F	1901	Belleville	
F	1903	Ameliasburgh	
M	1887	Belleville	
M	1892	Huntingdon	
M	1901	Belleville	
M	1902	Belleville	
M	1908	Tyendinaga Twp.	
M	1913	Belleville	
M	1914	Chipman, NB	Moved to Trenton in 1916

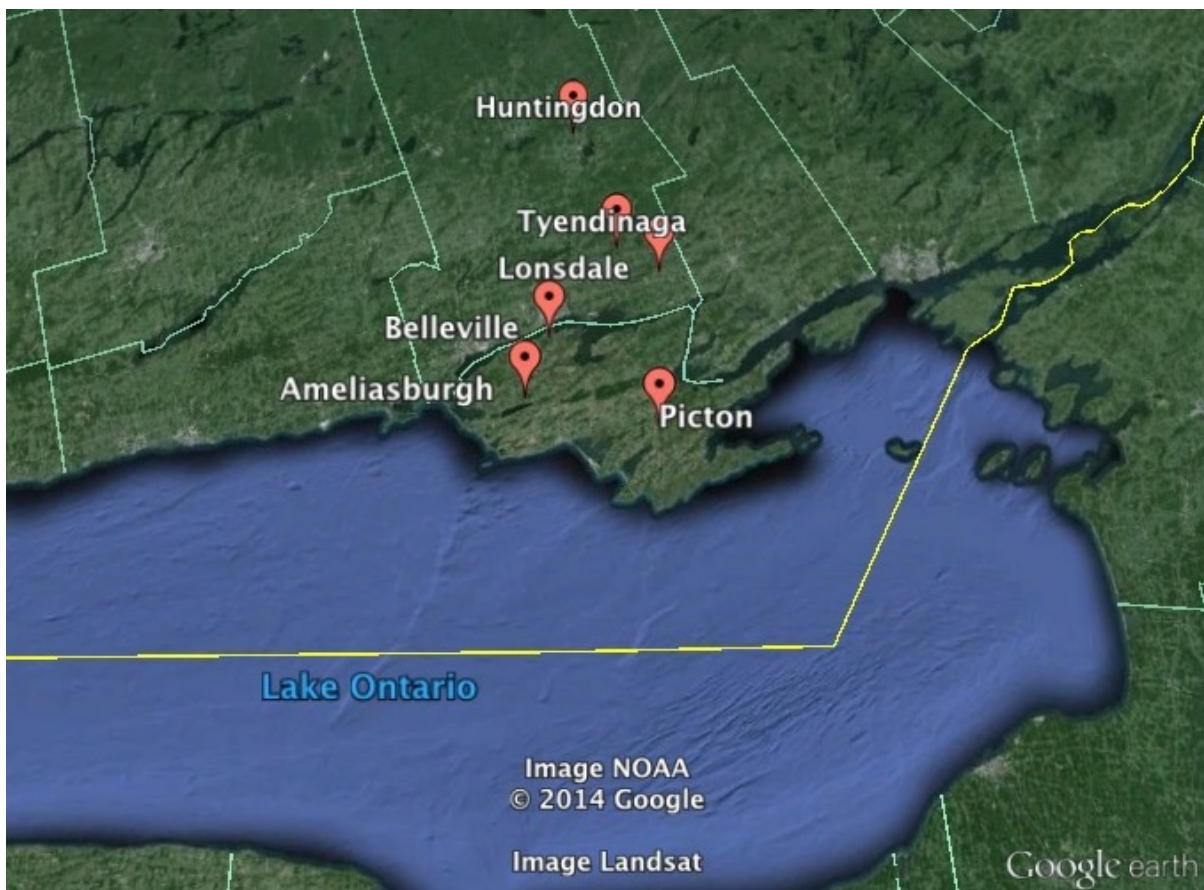


Figure 2.2: The birthplaces of Belleville 1975 speakers.



2.2 maps the location of the speakers' birthplaces.

### Eastern Ontario

The first European settlers in the most Eastern region of Ontario, in particular along the shores of the St. Lawrence River, were Loyalist soldiers and their families. In 1784, eight townships from the Ontario-Quebec border in the east to the Bay of Quinte in the west were surveyed and settlement began by those Loyalists who were granted land (Wallace 1914:98). It was determined that settlement would take place according to the corps that soldiers served. The King's Royal Regiment of New York, under Sir John Johnson, were settled in the first five of these townships. This settlement, New Johnstown, eventually became the United Counties of Stormont, Dundas and Glengarry. Many of the initial soldier settlers of what became Glengarry county, though American by birth, were of Highland Scottish descent (Wallace 1914:100). Migration directly from Scotland continued into the nineteenth century (Harkness 1946:50). Gaelic was widely spoken in the area (MacDonell 1893:150), though as of the 2011 census, only fifty people claimed fluency.<sup>35</sup> Nonetheless, this Scottish heritage is still apparent today, celebrated annually at the Glengarry Highland games. Many ethnically Rhenish-Palatine (German) Loyalists settled in the western part of New Johnstown. Like the ethnically Scottish settlers in the east, they were mostly

---

<sup>35</sup>In one interview from the *Farm Work and Farm Life Since 1890* project, a man from Kenyon recalls his aunts and uncles speaking Gaelic and singing Gaelic songs in the 1910s and 1920s. Though he did not learn the language, other children in the community learned it as a mother tongue. However, through mandatory schooling, English became the dominant language:

When I started to school there was one family not too far away, Gaelic they spoke. They hardly had any English and they also had a family- French family that came to our school more or less to learn English. So that's kind of a joke with us yes. One family Gaelic. I learned English. The other French. And here we're in between and we didn't try to learn either one of the languages. But my mother- father and mother always said there was no use of us learning because it was going to be a forgotten language. That's pretty much the end of the Gaelic around here.

(EON/M/1910)

born in America, their parents or grand-parents having arrived in New York and Pennsylvania three-quarters of a century earlier (Harkness 1946:12–13, 45).

Table 2.8: Demographics of the Eastern Ontario speakers.

Sex	Birth Year	Birth Place	Farm Location
F	1899	Cornwall	Osnabruck
F	1906	Bunker Hill	Osnabruck
F	1907	Moose Creek	Roxborough
F	1914	Dunvegan	Dunvegan
F	1915	Williamsburg	Williamsburg
F	1919	Mountain	Mountain
M	1891	Roxborough	Roxborough
M	1898	Osnabruck	Osnabruck
M	1904	Osnabruck	Osnabruck
M	1905	Kenyon	Kenyon
M	1910	Kenyon	Kenyon
M	1912	Osnabruck	Osnabruck
M	1918	Lochiel	Lochiel
M	1919	Osnabruck	Osnabruck

Of the twenty-eight interviews recorded by the *Farm Work and Farm Life Since 1890* oral history project, fourteen speakers were selected for analyses in this project. Only speakers who were born in the region were included. The goal was the most even stratification by age and sex of speakers involved in interviews of at least one hour. Six female speakers and eight male speakers, born between 1891 and 1919 were transcribed.<sup>36</sup> Table 2.8 presents the demographic information of each of these speakers and Figure 2.3 maps the location of the speakers' birthplaces.

## Niagara

Many locations throughout Upper Canada (Ontario), Lower Canada (Quebec), and the Maritimes were available to the Loyalists for settlement. Although the largest settlements in what is now Ontario were along the shores of the St. Lawrence River and Lake Ontario, the Niagara Region was of great political importance and became

<sup>36</sup>For interviews over an hour, only the first hour was transcribed.

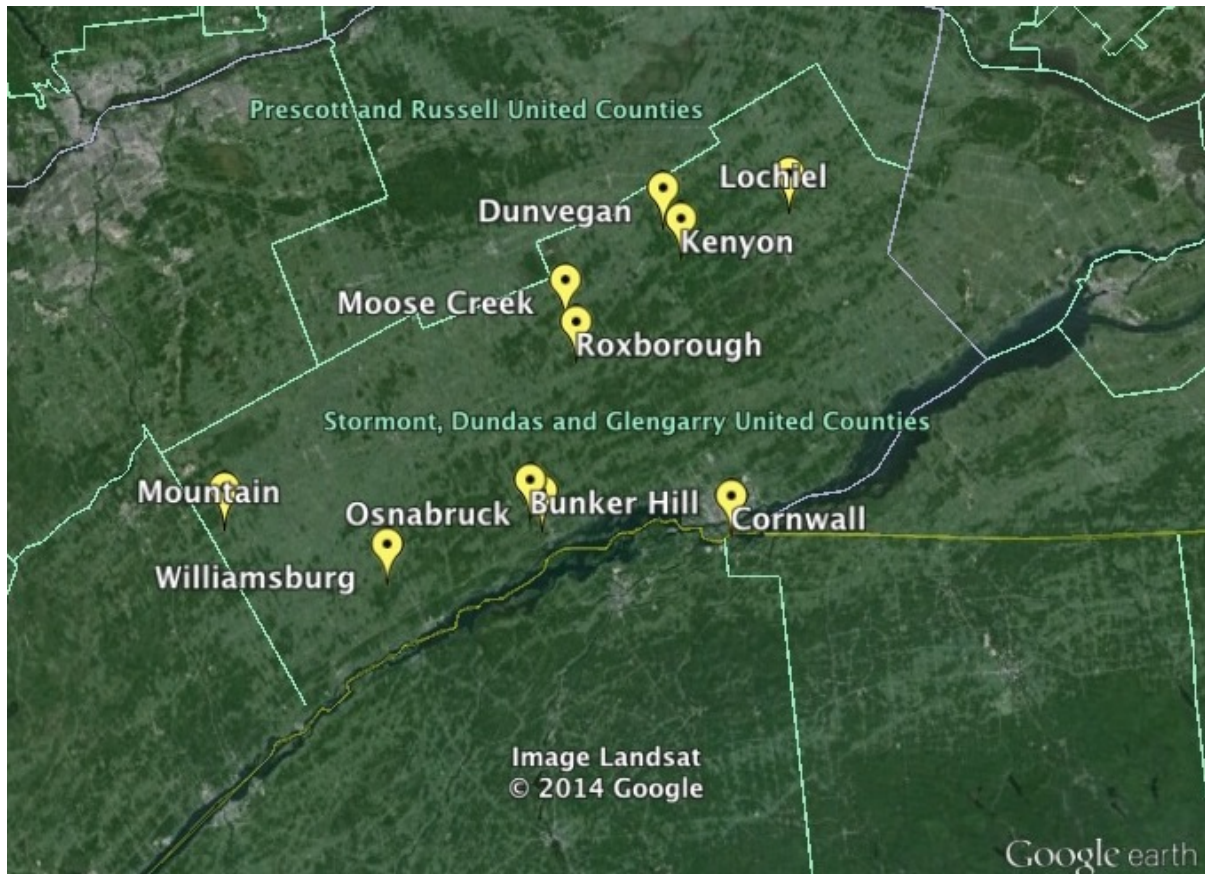


Figure 2.3: The birthplaces of Eastern Ontario speakers.

the capital of Upper Canada in 1791 and the location of the first parliament of the new crown territory (Wallace 1914:97; Carnochan 1914:13). The Niagara Peninsula had been the territory of the Mississauga First Nations and no Europeans had settled there until a small strip of Mississauga territory was purchased by the Crown in 1781 (Campbell 1966:6). Throughout the American Revolutionary War, this area had already served as a haven for Loyalist refugees, particularly from Pennsylvania, many of whom continued to squat in the area when the fighting ended (Wallace 1914:108). A prominent Loyalist regiment, Butler's Rangers, were stationed at Fort George (located in what is now Niagara-on-the-Lake). Many of these soldiers, originally from Upstate New York, were granted land in the Niagara region (Wallace 1914:109). The four-mile strip of land quickly became overcrowded and in 1784 the Crown purchased a tract of land that extended westward from Lake Ontario to Burlington, then

northwest for twelve miles, and from there to the source of the Thames River (then called the La Tranche River). Essentially, this included the whole of today's Niagara region (Campbell 1966:7). In the twenty years following the revolution, many more "Americans who much preferred the British way of government," and who faced hostilities from those who did not, migrated to the area (Nelles 1996:59).

Table 2.9: Demographics of the Niagara speakers.

Sex	Birth Year	Birth Place	Farm Location
F	1899	Waterford	Glanford
F	1902	Vineland	Beamsville
F	1903	Beamsville	Beamsville
F	1904	Grimsby	Grimsby
F	1906	Lincoln	Lincoln
F	1911	Winona	Saltfleet
F	1912	Fenwick	Pelham/Gainsborough
F	1916	Winona	Saltfleet
M	1895	Vinemount	Vinemount
M	1898	Clinton Twp.	Beamsville
M	1899	South Grimsby	Caistor
M	1902	Grimsby	Grimsby
M	1907	Wentworth	Wentworth
M	1911	Clinton Twp.	Clinton Twp.
M	1913	Vineland Station	Vineland Station
M	1917	Grimsby	North Grimsby

Of the thirty-two interviews recorded by the *Farm Work and Farm Life Since 1890* oral history project, sixteen speakers were selected for analyses in this project. Again, only speakers who were born in the region were included. Speakers were stratified by age and sex. Eight female speakers and eight male speakers, born between 1895 and 1917 were transcribed. Table 2.9 presents the demographic information of each of these speakers and Figure 2.4 maps the location of the speakers' birthplaces.

In this thesis any potential dialect differences between these three communities will be de-emphasized.<sup>37,38</sup> Rather, as three communities that were founded by

<sup>37</sup>That said, potential dialect differences will be noted.

<sup>38</sup>Likewise, the potential effects of urban-ness, social class, and gender are also backgrounded (though see immediately below). I do not wish to suggest that these social factors are play no in

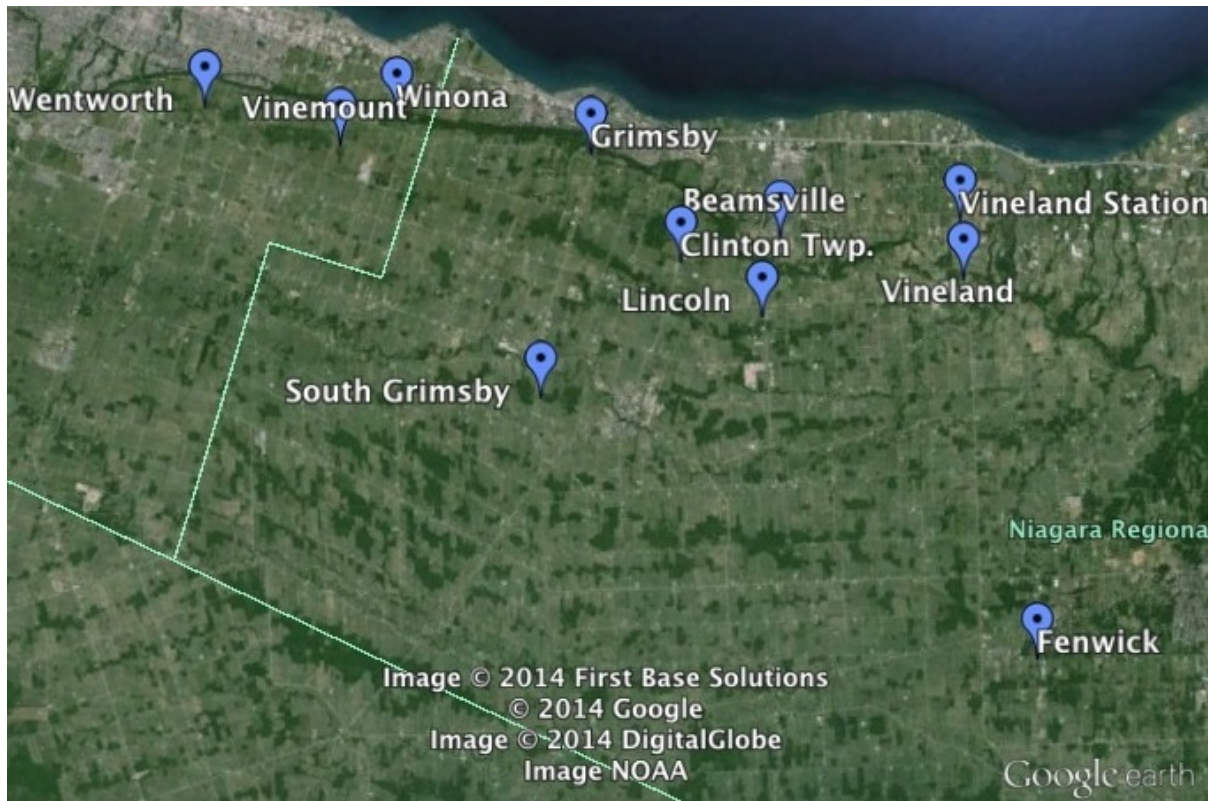


Figure 2.4: The birthplaces of Niagara speakers.

American-born, American English speakers, argued to be the source of Canadian English (Avis 1967:vii; Bloomfield 1975:5; Chambers 1998:x), these communities together will be treated as representing an earlier Ontario English, particularly relative to the Toronto English Archive. All together, the EOE communities and TEA represent an apparent time span of one hundred and thirteen years, as shown in Table 2.10 and Figure 2.5.<sup>39</sup> Although EOE by and large represents older speakers, and TEA younger speakers, there is considerable overlap with respect to the ages of speakers in TEA and EOE and some small overlap with respect to their year of birth.

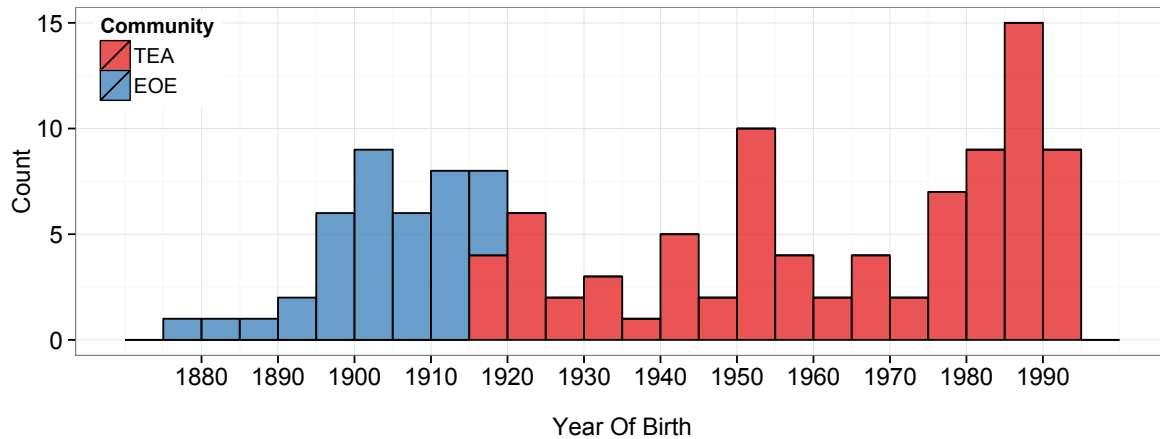
---

explaining the variable patterns of GEs and EPs. Rather, these effects are considered to be orthogonal to the questions at hand about grammaticalization.

<sup>39</sup>A subsample of eighty-seven speakers from TEA, stratified by age and sex are used.

Table 2.10: Temporal coverage of TEA and EOE.

Community	Year Of Recording	Ages	Year Of Birth
Toronto	2002–2004	10–87	1916–1992
Belleville	1974–1975	60–96	1879–1914
Eastern Ontario	1984–1985	65–93	1891–1919
Niagara	1984–1985	66–86	1898–1917

Figure 2.5: Histogram of number of speakers across 20<sup>th</sup> century

### 2.2.3 A note on possible confounds: Time, urban-ness, and social class

It is necessary to point out two possible confounds in comparing the EOE data with the TEA data and interpreting the results following the apparent time construct. While the EOE contains data from speakers born earlier than speakers in TEA, the EOE and TEA also differ with respect to the urban-ness of the communities and the social class of the speakers. While Toronto is a large urban centre and the TEA contains interviews with primarily middle class individuals, the communities in the EOE corpus, at least in Eastern Ontario and Niagara, are rural and the speakers are farmers.<sup>40</sup> In other words, while the EOE and TEA are on opposite ends of the time continuum (EOE being older and TEA being newer), they are also on opposite ends

<sup>40</sup>The Belleville data is more comparable to TEA. Although Belleville has a smaller population than Toronto, it is a city. Furthermore, the speakers in the Belleville data were by and large middle class, much like the speakers in TEA.

of the urban-ness continuum and middle-class continuum.<sup>41</sup> Therefore, the question arises as to whether the EOE can be taken as a real-time benchmark for TEA. Although differences in urban-ness and class create an imbalance in the design, this confound is consistent with respect to predictions about language change as follows: With respect to both real time and apparent time, the older EOE is expected to represent a more vestigial form of Canadian English; with respect to urban-ness, the more rural EOE is expected to represent a more vestigial form of Canadian English; and, with respect to middle-class-ness, the less middle class EOE is expected to represent a more vestigial form of Canadian English. An ideal comparative situation would be to consider real time data from Toronto; however, to my knowledge, no such data exists. This is why I have had to make “the best use of bad data” (Labov 1994:11). In other words, although I do not have access to older data that represents speakers who are as equally urban or as equally middle class as the speakers in the TEA are, having two datasets that are maximally and consistently different with respect these factors is the next best thing.<sup>42</sup> The practical result is that the EOE represents an even more vestigial form of Canadian English than a corpus of Toronto English from 1975 or 1984 would represent. Thus, differences between EOE and TEA that are discussed in this thesis, (although discussed in terms of real and apparent time rather than urban-ness and social class) are interpreted as a result of linguistic change in Canadian English.

#### 2.2.4 Oral Histories as Historical Variationist Data

Because this thesis builds many arguments around comparisons between oral history data and the sociolinguistic interviews in TEA (and other sociolinguistic corpora), a

---

<sup>41</sup>Both urban-ness (Trudgill 1974a) and middle-class-ness (Labov 2001) have been argued to be factors in linguistic change with speakers from urban centres and centrally-located social classes being on the forefront of change.

<sup>42</sup>This logic is similar to traditional dialectologists who sought out Non-mobile, older, rural males (NORMS) because they (presumably) represent the most vestigial variety of a given language.

demonstration of the comparability of oral histories and sociolinguistic interviews is essential. As just discussed, ensuring the comparability of data sources is a methodological requirement for empirically rigorous variationist sociolinguistics. Thus, researchers “need to exercise great caution when comparing corpora of socially and regionally diverse speech samples, and consider at all times the possibility that cross-corpora differences may reflect the effect of differential contextual constraints on the variation, rather than actual social or geographical variation” (Pichler 2010:586). In addition to the social and geographic, any diachronic conclusions must be subject to the same considerations.

In discussing the earlier mentioned *Récit français québécois d'autrefois*, Poplack and St-Amand (2007:717) make the case that because oral histories contain “inherent variability typical of contemporary varieties” we can consider the speech therein to be a reasonable representation of a historical vernacular, and thus comparable to today’s sociolinguistic corpora. For example, (8) exhibits variable *ne* deletion.

- (8) Parce qu’ en ce temps là l’argent n’ était pas commune, savez-vous,  
 Because in those days LÀ money-DEF NE was not common, you know,  
 elle ∅ était pas parlée beaucoup.  
 it NE∅ was not talked about much

(Poplack and St-Amand 2007:722)

Indeed, as in the *Récits du français québécois d'autrefois* (Poplack and St-Amand’s 2007), both the FWFL and Belleville data contain many examples of variability, typical of vernacular speech. This includes variation at the level of phonology, as in (9), morphosyntax, as in (10) through (16), and discourse-pragmatics, as in (18) through (21).

- (9) **Variable (ɨ):** Don’t remember. All I can remember is be<sub>[ɪ]</sub> out there hoe<sub>[ɪ]</sub> them. Keep<sub>[ɪ]</sub> them clean.



(NIA/F/1902)

- (10) **Subject-Verb Concord:** And you never knew if you **was** going to have them for a meal, at what time. ... Oh yeah, you **were** busy then.

(NIA/F/1902)

- (11) **Non-standard tense expression:** When we **come** home? It takes quite a long while to walk that far. And of course children don't always hurry home you-know! There was nothing particular we had to do when we **was** younger. But as I said, the men **done** the outside work and mother usually had the inside work done so.

(NIA/F/1899)

- (12) **Multiple Negation:** Some would have a cow and that would require hay and- and they **didn't** grow **no** hay. They grew fruit down there...

(NIA/M/1907)

- (13) **Null Relativizers:** I suppose there'd be six or eight people  $\emptyset$  come here and peel apples.

(NIA/F/1903)

- (14) **a-prefixing:** And then, soon as that was done then we were **plowing-** start **a-plowing** or maybe sow wheat.

(NIA/M/1907)

- (15) **Emphatic do:** Lost Air was- it's not like Euchre but it was a game with cards. I guess maybe older people would remember what Lost Air was. In fact, I **do** believe you can get it at the present time if you look for it.

(NIA/F/1899)

- (16) **Null Complementizers:** But I know  $\emptyset$  my mother said **that** I learned to making puddings and things you know.

(EON/F/1906)

- (17) **Null plural -s:** The teacher $\emptyset$  I told you, they always boarded at our place.

(EON/M/1891)

- (18) **Utterance Final Tags:** We had several dogs **you know eh?**

(EON/M/1905)

- (19) **General Extenders:** I did Christmas platforms **and all that kind of thing** and the one they seemed to like best was the one I did for... one Christmas I was tired of Santa-Claus pictures **and all that stuff** so I uh- I did one...

(BLV/F/1898)

- (20) **Epistemic Parentheticals:** Well we had a hired girl in too. And of course my sister was older than I was. She'd be about twelve **I suppose**, fourteen maybe. Mother **I guess**, she and Megan, that was the hired girl, **I guess** they did it all.

(NIA/F/1904)

- (21) **Approximative adverb like:** It was **like** four farms it- and we all changed and worked together.

(NIA/M/1907)

- (22) **Discourse marker like:** She uh drew by shading and crayon pictures. **Like**

she'd make the eye first and make the rest of it, the animal, afterward just by shading.

(NIA/M/1902)

In addition to the presence of such inherent variability, three more facts suggest that the speech of participants in FWFL and Belleville 1975 represents the vernaculars of these speakers and thus, is a sufficient historical benchmark for the contemporary sociolinguistic interviews in TEA.

First, the sets of questions asked during the oral history interviews in both Belleville 1975 and FWFL have a great deal of overlap with topics often included in sociolinguistic interviews. The questions asked in the interviews from the *Belleville 1975* portion of the corpus were personalized for each individual speaker but many of the topics discussed in these interviews would appear in standard sociolinguistic interviews such as school years, childhood customs, and major community events like floods. The interviews in the FWFL component of EOE all asked roughly the same questions of the interviewees and many of these questions have parallels with questions in the UofTLVC interview schedule (Tagliamonte 2006a: Appendix B). The interview schedule for the FWFL was reconstructed from Summary Analysis records included in the Archive of Ontario material. These records contain a basic skeleton of the topics discussed in each interview. Using these records, we can easily ascertain the original set of interview questions. In Table 2.11, I have outlined the interview plan of the FWFL interviews in comparison to topics included in Tagliamonte's (2006a) interview schedule.

Of the thirty topics discussed in the FWFL interviews, twenty-four overlap with topics in sociolinguistic interviews including questions about family, early life, community, and social practices. Many of these topics are those that Labov (1966b, 1972) notes are particularly useful for eliciting the vernacular of the speaker. Indeed, the

Table 2.11: Comparison of FWFL interview schedule to standard sociolinguistic interview topics.

FWFL Topic	SLX Topic
1. Age of siblings	Parents & Family
2. Age of parents when born	Parents & Family
3. Type of farm	n/a
4. Types of crops and livestock	n/a
5. Description of farm	Demographics
6. Typical yearly routine	n/a
7. Chores	Helping Out
8. Earliest recollections helping out	Helping Out
9. Doing a man's/woman's work	Work Life
10. Labour bees	Community Events
11. School	School, Peers
12. Before and after school	Games, Peers
13. Childhood friends	Peers
14. Leisure and games	Games
15. Dances	Teen Life
16. Pets	Pets
17. Special days	Holidays, Tradition
18. Family outings	Tradition
19. Going to town	Social Practice
20. Clubs	Social Practice, Hobbies
21. Hired help	Neighbourhood
22. What were you raised to consider important?	Kids these days
23. Pocket Money	Hobbies
24. Work off farm	Work Life
25. Occupational Aspirations	School, Work
26. Time with family	Family & Parents
27. City vs. farm	n/a
28. The Depression	n/a
29. Courtship/Marriage	Marriage
30. Transition to tractor age	n/a

second argument that these oral histories are comparable to sociolinguistic interviews is that there are many contexts in the interviews that elicit Casual Speech in the Labovian sense. Labov (1972:87–94) discusses five contexts in particular. The first context that can elicit the vernacular is when the recording includes speech outside the formal context. This includes interruptions in person and on the telephone, or any other time the speaker steps outside their role as interviewee, such as to offer

the interviewer a drink. Unfortunately, in the FWFL interviews, this type of speech was rarely recorded. The interviewer typically stops the tape when the interview is interrupted. The second context that effectively elicits the vernacular is speech with a third person. In many interviews in the FWFL data a speaker's spouse is also in the room and there are often interactions between the speaker and this third party. The third context for eliciting the vernacular is speech not in direct response to questions. Although the interview topics were planned in advance and more fixed than sociolinguistic interviews, detailed answers, elaboration and stories were highly valued by the researchers. Speakers in the FWFL data are not only allowed to digress, but the interviewers were encouraged to allow them to speak freely. For example, the interviewer in (23) asks how old the interviewee was when he started school and the interviewee provides a long digression about getting 'the strap'.

- (23) <1> And how old were you when you started at school? </1> 1  
 <19> To school I suppose- I lived right beside of the school. Of 2  
 course they started me when I was five years old. So, I suppose 3  
 the next year I was starting to the lake! <laughter/> Oh well that 4  
 was a great school to get the rubber straps! You know teachers 5  
 always had it nice. [...] Remember this one day, first cousin of mine 6  
 was teaching- she wasn't a very big person but she had a real mad 7  
 temper. Oh a temper like a <inc/>. [...] She lined thirteen of us up 8  
 on the front of the school and this one fellow, he was away down at 9  
 the far- or up at the far end, it don't matter. Don't ask me. When 10  
 she come to him, you know she just had such a temper. When she 11  
 start putting the strap up, the first half of dozen of us, oh boy, boy. 12  
 But she had tired herself out when she got up there, the strap would 13  
 just come down. He just kept a-laughing, the madder she got! But 14  
 she had nothing left to bring the strap down! Oh- oh boy, [...] the 15  
 way she licked us, it was terrible! But then we had it all covered 16  
 really. So I suppose that's another- it's a reason why I didn't learn 17  
 anything. I was always trying to get in trouble too much! </19> 18

(EON/M/1904)

Likewise, the speaker in (24) digresses into a narrative about his early driving expe-

riences after being asked if he remembers his first car.

- (24) <1> Do you remember your first car? </1> 1  
 <024> Yes, the first car was a Ford car with a brass radiator and 2  
 uh you had to crank it. And I used to go to Ham- uh Grimsby to 3  
 get the mail and uh I was be- below age for driving so the chief 4  
 of police in Grimsby, they only had one, he'd come out and tell 5  
 me how to start it. He'd say uh, "Son, that's no way to start that. 6  
 You pull that little wire out and turn it over once, and then put the 7  
 switch on, and she'll roll." And if I couldn't get it to go, he'd crank 8  
 it for me. </024> 9  
 <1> About how old were you? </1> 10  
 <024> I was thirteen. And that went on for a couple years and that 11  
 chief of police lost his job and they got a new constable and they- 12  
 I got stopped the first time in town. So then I couldn't go to town 13  
 anymore. So my father had a friend in Hamilton that had a son my 14  
 same age as I was and they got a special license and he uh got my 15  
 license and I had a picture and a button on my lapel to show you 16  
 I could drive a car and I came back to Grimsby and the- the town 17  
 policeman then, he'd never seen one before and he didn't know 18  
 what to think of it! </024> <1> What was the driving age? </1> 19  
 <024> Mm? </024> 20  
 <1> What was the driving age at the time? </1> 21  
 <024> Eighteen! </024> (NIA/M/1902) 22

These digressions were highly valued by the researchers and interviewers rated the quality of each interview in the original field notes based on criteria such as "elaborated on questions very well" and includes "in depth personal insights" (AO RG 16-200 Dufferin County Field Notes). Labov's fourth context is the discussion of childhood customs. As is evident in Table 2.11, school, peers, and games were all discussed at length in the FWFL interviews. For example, in (25) speaker this speaker from Niagara discusses what a typical birthday party for a child at the turn of the century might include.

- (25) <1> What would you do on a birthday? </1> 1  
 <001> Oh generally hope for a party. Not lavish like they are now, 2  
 with moving pictures or anything like that, but I think we played 3  
 games. I think I can remember some birthday parties where we 4  
 dressed up and stuff. Nursery rhyme characters or something like 5

that. </001><sup>43</sup>

(NIA/F/1906)

6

In (26), this speaker discusses her favourite toys by way of a digression.

- (26) <1> Do you remember any of the toys that you had? </1> 1  
 <009> Oh I remember a good doll, it was called Howard! Of all 2  
 names <laughter/>. Howard was named after a Baptist minister! 3  
 That was supposed to have been my name only it turned out to be 4  
 Beth! That was- the cheap thing I can remember. Most of my toys 5  
 were very much alive. Pet pigs. </009> 6  
 <1> So you were given- was it a baby to look after? </1> 7  
 <009> Oh yes. And I'd raised lots of pigs on the bottle and pigs 8  
 used to wear nightgowns and ride in my doll carriage! A friend had 9  
 an old aunt who had lost her husband and she was visiting her after 10  
 the funeral. They brought her home there and she was very much 11  
 upset and she was an old Irish lady and she was having a crying 12  
 spell one day and our neighbour didn't know how to manage her 13  
 at all. Well I walked in with a pet pig with a nightgown on and 14  
 dropped it in her lap and she stopped crying and laughed! </009> 15  
 <1> Were you responsible for raising these pigs? </1> 16  
 <009> Well not after they got passed the point where they fit in the 17  
 doll carriage. Well then I wasn't interested! </009> (NIA/F/1912) 18

The last context is the "Danger-of-Death" question. Although there are no direct questions like the Danger-of-Death question, the most popular topics of discussion (at least in Eastern Ontario) were "The Fire of 1916" and "The Fire of 1922".<sup>44</sup>

Third, as evident in (23) though (26), the data contain many narratives of personal experience. This is a characteristic feature of successful sociolinguistic interviews since narratives involve minimal attention to speech and thus a high likelihood of vernacular speech. The main goal of oral histories is to elicit first-person narratives. The FWFL field notes contain a list of "Particularly Good Stories" including stories about fires, school, games, labour bees, and among other farm related topics, making butter. (27) exemplifies one of these narratives by a female speaker born in 1899.

<sup>43</sup>Note that this short example contains two EPs and two GEs.

<sup>44</sup>In the UofTLVC Lab's interview plan, there is a question "Did anything ever happen around here, like a big fire?"

- (27) <1> Would you say that you got along better with your mother? 1  
 </1> 2  
 <22> I got along well with both of them. I only had one spanking 3  
 that I can ever remember in all my life. </22> 4  
 <1><> Do you remember what you did? </1> 5  
 <22><> Uh huh! I didn't do it! We didn't deserve it! </22> 6  
 <1><> Oh no! </1> 7  
 <22><> We had, in the springtime you know, when the grass was 8  
 green on the side of the road and the fields were damp, too damp 9  
 for anything to put- to walk on them you-know. They would let 10  
 the sheep go out. One year we had the sheep. And dad says to 11  
 watch them so they didn't run away. He put them out on the road 12  
 so that they could eat along the side. Well, mother says "you don't 13  
 need to go out for a little while because they'll be busy eating." So 14  
 we waited maybe fifteen, twenty minutes and when we went out 15  
 there wasn't sheep to be seen! Gone completely! So, we started 16  
 hunting them. We went up to my neighbours and the little girl- 17  
 neighbour friend, she come along with us and we walked for miles! 18  
 We walked practically all day hunting them! Children you-know! 19  
 We weren't old enough to know that they wasn't far away. So we 20  
 walked all day, well then come up in the afternoon, come up a 21  
 thunder-storm. And we ran to our neighbours on another road 22  
 over and uh went in and she called up dad and said we were over 23  
 there. Well in the meantime the sheep had found a little ditch and 24  
 they went under the ditch and if one sheep goes, they all go. One 25  
 sheep went in the field and all the rest of them followed and we 26  
 couldn't see them 'cause they were in the hay! Well, dad thought 27  
 we were gone playing you know and never bothered with- so I got a 28  
 spanking! My sister didn't but I did! And I always figured I didn't 29  
 deserve it because we were hunting sheep all day! It's just we're 30  
 too foolish to know where they might be close at home! </22><sup>45</sup> 31

(NIA/F/1899)

The Belleville data also contain many narratives. For example, (28), repeated from (7b), is a narrative from a male speaker born in 1914.

- (28) <033> That's kind of the river watchin' that goes on all the time 1  
 and you see uh quite a bit of it in the paper because uh um- it's 2  
 the only visible thing of what we're doin' at that time towards the 3

<sup>45</sup>This narrative contains a great deal of variation between the historical present and past morphology on verbs.



emergency measures I guess. But it- it's uh sometimes quite critical, you know. We've seen the uh- I've seen the water come up uh, um, I think it was six feet in forty minutes, uh, oh, several years ago and- and the chairman of the conservation authority and myself and the engineers stood on the footbridge and- and you know, said that- "do we say close down downtown or do we say go." Um we asked the police to have people move their cars out of the parking area and they did so and they had to get the trucks even down, the tow trucks and uh I looked around the one car left and I said, "what idiot left his car there?" because the water was about that high up on it. It was mine. I'd forgot about it. </033>

4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14

(BLV/M/1914)

## 2.3 A Note on Tools

The inferential and descriptive statistics reported in this thesis have all been conducted using *R: A language and environment for statistical computing* (R Development Team 2014). In addition to the basic base and stats packages of R, other packages have been implemented for the following specific kinds of analyses: *lme4* for mixed-effects modelling (Bates et al. 2011); *party* for conditional inference trees (Hothorn et al. 2006); and *ggplot2* for all plotting (Wickham 2009). Any code is available upon request. Concordance was performed using *AntConc 3.2.4m* (Anthony 2011).

### 2.3.1 Logistic Regression Modelling in R as Variable Rule Analysis

The standard statistical tool in variationist sociolinguistics is variable rule analysis, as instantiated most recently by the GoldvarbX program (Sankoff, Tagliamonte, and Smith 2012). GoldvarbX essentially performs fixed effects logistic regression on binary dependent variables and zero-sum/contrast coded categorical independent variables. The best model is determined by a step-up/step-down procedure that compares log-likelihoods. This thesis does not use GoldvarbX, but rather uses the

cross-platform, open source, statistical programming environment R (R Development Team 2014) for statistical analysis. Using R has many advantages; most importantly, it allows us to go beyond some of the constraints of GoldvarbX. In particular, the `lme4` package (Bates et al. 2011) allows for logistic regression that 1) is hierarchical (i.e., mixed-effects) and 2) allows for continuous independent variables. I will not go into great detail, but see Johnson (2009) and Tagliamonte and Baayen (2012) for further discussion.<sup>46</sup>

Fixed effects logistic regression assumes that every data point is independent. A problem that standard variable rule analysis faces is that if there are multiple tokens from a single speaker in your data set, this assumption of independence is violated. Hierarchical (or mixed-effects) modelling accounts for dependencies in data with random effects. Since the data presented below contain multiple tokens from individual speakers, in my statistical analyses, I use hierarchical modelling with a random intercept of speaker to account for the inherent dependence of these tokens.

Another disadvantage of GoldvarbX is that independent variables are limited to being categorical. However, variationists interested in language change often use some proxy for time (be it speaker age or year of birth) to track change. While it is possible to bin continuous factors into categories, it is not always straightforward where partitions should be made and it is almost always the case that binning results in a loss of statistical power (Johnson 2009). The `lme4` package not only allows for continuous variables but is also not limited to linear effects: different curves can also be modelled (Tagliamonte and Baayen 2012).

To summarize, while using R offers many advantages, it is important to remember that whatever platform one uses variable rule analysis is accountable statistical analysis that tests the effects of multiple independent variables on a binary dependent

---

<sup>46</sup>Other advantages of using R include: multiple ways of selecting the best model including those that penalize the complexity of the model (e.g., Akaike Information Criterion [AIC]), easy manipulation of data for calculating distributional results, changing coding structures (e.g, sum coding, treatment coding etc.) and plotting, and cross-disciplinary familiarity.

variable.

### 2.3.2 Presentation of statistical results

Variationist studies are typically built from a step-by-step discussion of distributions, proportions, total Ns, and finally logistic regression that test the effects of multiple constraints and the patterns therein are often interpreted following the three lines of evidence (statistical significance, hierarchy of constraints, and magnitude of the effect) (Poploack and Tagliamonte 2001). In this thesis, I have followed the same procedures, typically presenting a distributional analysis along with any statistical models, but for presentational purposes I have chosen to examine and statistically test individual hypotheses about the grammaticalization of GEs and EPs one at a time, by examining a single main effect and its interaction with time—as operationalized by either speaker year of birth (apparent time) or EOE vs. TEA (real time), (e.g., Tables 3.10, 4.6, and 4.11).<sup>47</sup> That said, all basic methodological steps were followed in the analyses and the statistical significance and direction of effects presented in the models below are consistent with models that include all possible conditioning factors (include the speaker sex and variable-specific linguistic factors) and their possible collinearity with other factors (although these analyses are not shown in the dissertation). In other words, in no statistical model below was a reported significant effect non-orthogonal to other possible effects discussed in the dissertation.<sup>48</sup>

---

<sup>47</sup>In some cases, only a distributional analysis is presented due to low Ns.

<sup>48</sup>This was determined through a series of distributional analyses and by including other main effects and interaction terms in models, again not shown here.

# Chapter 3

## The Development of General Extenders

*Cartoons don't have messages, Lisa. They're just a bunch of hilarious stuff you know, like people getting hurt and stuff, stuff like that.*

*-Bart Simpson*

*So... do you like... stuff?*

*-Ralph Wiggum*

### 3.1 Introduction

This chapter serves two broad purposes.<sup>1</sup> First, I address Pichler and Levey's (2011: 462) call for a real-time benchmark to assess the role of grammaticalization with respect to the development of general extenders (GEs) by replicating the diagnostic tests for the mechanisms of grammaticalization on the general extenders in Earlier Ontario English (EOE) (Cheshire 2007; Tagliamonte and Denis 2010). Second, I expand the investigation of GEs by considering their long-term trajectory and the mechanisms of grammaticalization in Ontario English in real- and apparent-time perspective by using newer statistical methods, namely mixed effects logistic regression.

---

<sup>1</sup>Parts of this chapter appear in an unpublished manuscript, Denis 2010.

I argue that there is no evidence that the mechanisms of grammaticalization act in unison, as a single, gradual process, during the development of GEs in Ontario English. Rather, there are three changes taking place. First, there is a process of lexical replacement of *stuff* type GEs but, unlike what Tagliamonte and Denis (2010) find, the rise of *stuff* type GEs is not just at the expense of *thing*. Rather, by the turn of the 21<sup>st</sup> century, *stuff* type GE variants have made massive inroads such that the variation present one-hundred-plus years earlier has become eclipsed by the new majority variant. Second, I argue that putative phonetic reduction of GEs from long forms to short forms (e.g., *and stuff like that* > *and stuff*) that has been frequently reported in the literature (*inter alia* Aijmer 2002; Erman 1995; Overstreet 1999; Cheshire 2007) is the result of a single change, which is independent of grammaticalization and which effects all GEs at a constant rate—clipping of *like that*. Lastly, to the extent to which any semantic-pragmatic changes effecting GEs can be identified, these are independent of other mechanisms of grammaticalization and thus, independent of a theorized, uniform process of grammaticalization (cf. Lehmann 1982; Heine 2003; Traugott 2003; Diewald 2010).

This chapter is organized as follows. I begin in §3.2 with a general background on GEs and a review of the relevant literature, both variationist and not. I focus on how grammaticalization theory has been implicated in observed changes to the variable GE system. In §3.3, I move on to an overview of the inventory and distribution of GEs in EOE and across the 20<sup>th</sup> century. In total, there are 643 GE tokens from EOE and 2178 tokens from TEA.<sup>2</sup> Next, I examine the long term of trajectory of grammaticalization of GEs in §3.4. I consider a real-time comparison and also abstract away from the time (and place) of the recording of the corpora used, concentrating on the *longue durée* of the mechanisms of grammaticalization in Ontario English (cf. D’Arcy 2012: on quotatives). Finally, in §3.5 I conclude that the development of

---

<sup>2</sup>All GEs in both EOE and TEA were extracted by eye and hand-tagged while reading through the transcripts. All GEs in the transcripts were extracted.

GEs is not a story about the grammaticalization of a set of pragmatic markers, but rather an instance of multiple, independent changes acting on a multifaceted variable system.

### 3.2 General Extenders and Grammaticalization

General extenders are a set of pragmatic markers which occur clause- (or phrase-) finally as in (1).

- (1) a. We used to have church socials *and things like that*.

(NIA/F/1912)

- b. I bought some eggs *and stuff* from up at Atwood.

(EON/M/1912)

The literature converges on identifying the (at least core or typical) function of GEs as having to do with set-marking.<sup>3</sup> Dines (1980:22) argues that GEs “cue the listener to interpret the preceding element as an illustrative example of a more general case.” Dubois (1992:198,182) suggests that GEs are typically used “to evoke some larger set” or “to suggest the multitude of possible elements of the set.” Cheshire (2007:157; following Aijmer 2002; Brinton 1996; Erman 2001; Overstreet 1999) observes that “what is assumed to have been their original meaning” is “indicating that the clause element to which [GEs] are attached should be seen as an exemplar of a more general set.” The set-marking function of GEs in (1) is clear. (1a) is a proposition about a set of

---

<sup>3</sup>This general set-marking (Winter and Norrby 2000) or category-implicating (Overstreet 1999) function has led to a number of names for GEs in the literature including *set marking tags* (Dines 1980), *vague category identifiers* (Channell 1994), *approximators* (Erman 2001), *discourse extenders* (Norrby and Winter 2002), and *extension particles* (Dubois 1992). However, in the last decade, Overstreet and Yule’s (1997) term *general extender* has won favour in the literature (Overstreet 1999; Cheshire 2007; Tagliamonte and Denis 2010; Terraschke 2010; Pichler and Levey 2011; Palacios Martínez 2011; Levey 2012; Parvaresh et al. 2012). I will continue to use the term general extender.

social events such as *church socials*, and perhaps *picnics, quilting bees, bazaars* and more. The speaker in (1b) is telling his interlocutor that what he purchased in Atwood was not only *eggs* but also perhaps *butter, cheese, flour* and other agricultural foodstuffs. A more formal discussion of the semantics of GEs will be discussed below in §3.4.2.

In addition to this core function, GEs are also united under a structural umbrella. Not only do GEs have a typical structural position in the syntax (phrase-finally), they also have a typical internal structure. Rather than relying solely on a functional definition of GEs, Dines (1980), Tagliamonte and Denis (2010:336–7), and Pichler and Levey (2011:449) define GEs structurally. Dines (1980:18) gives the templatic schema in (2),

(2) AND/OR [PRO FORM] (LIKE THAT)

while Tagliamonte and Denis (2010:337) give a more detailed template illustrated in Table 3.1 composed of four components: CONNECTORS, QUANTIFIERS, GENERICS and COMPARATIVES.<sup>4</sup> These four components combine in various ways and each component is optional as in the examples in (3). The connector is typically followed by a quantifier, though a quantifier is optional. The type of connector, *and* or *or*, has been used to broadly partition GEs into ADJUNCTIVE and DISJUNCTIVE forms respectively though it is possible to have GEs without a connector. See §3.3.1 below for further discussion of this partition. The third element is a generic noun. The optional comparative is found either as the last element (e.g., *like that, of that kind*) or before the generic (e.g., *sort of, kind of*).

---

<sup>4</sup>Pichler and Levey (2011) offer an even more detailed pattern that separates Tagliamonte and Denis' (2010) comparative element into two parts, a simulative (*like*) and deictic (*this/that*). Some comparatives are included with quantifiers in a modifier group, while some quantifiers are included as parts of the generic/pro-form.

- (3) a. Connector+Quantifier+Generic+Comparative:  
Pigs? Oh general ten or a dozen *or something around there*.  
(NIA/M/1907)
- b. Quantifier+Generic+Comparative:  
And we had no television of course or no radios *nothing like that*.  
(NIA/F/1899)
- c. Connector+Generic:  
They'd be starting it all in spring, seedlings and vegetables in March.  
Sowing seeds *and stuff*.  
(NIA/M/1902)
- d. Connector+Generic+Comparative:  
They loaded an ambulance with generators, lights *and things like that*.  
(BLV/M/1914)
- e. Connector+Quantifier+Comparative:  
Oh yes I was the outside. Yep. It's a farm. Milking *and all that*.  
(EON/F/1907)
- f. Connector+Comparative:  
Well you have to tell them everything you have *and like that* and then [...] if you haven't saved anything for your old age, well then they'll give you a certain amount, you see.  
(BLV/F/1879)

In addition to these templatic (or prototypical) GEs, there is also a subset of GEs that do not conform to the template but, nevertheless perform a set-marking function. Some examples are presented in (4).



Table 3.1: Typical GE templatic structure, based on Tagliamonte and Denis (2010:337)

Connector	Quantifier	Generic	Comparative
and	all	thing(s)	like that
or	every	stuff	sort of
∅	some	people	kind of
	any	one	type of
	the odd	where	of that kind
	the whole	shit	of that sort
	no	crap	of that type
	∅	baloney	around there
		∅	to that effect
			∅
etc.	etc.	etc.	etc.

- (4) a. Putting down hay for the cattle and feeding that cattle. Feeding them corn *and so on*. Yes.

(EON/F/1907)

- b. Then of course in the summertime there'd be extra help for harvest *and so on like that* but mostly during the war it was women and children.

(NIA/M/1907)

- c. And of course in the wintertime, I guess it was getting out, building forts and having a snowball fight or throwing some snowballs at somebody driving by *or whatever*.

(EON/F/1919)

- d. You know, there's kids that steal *and one thing another* but none of us kids ever- we were too good and we didn't have time to.

(NIA/F/1911)

- e. I generally played on all the teams that went out of high school. Uh the basketball, baseball, hockey, *et cetera*.

(BLV/F/1898)

Throughout this chapter, I will make reference to different levels of variation and different ways of circumscribing the variable context. Figure 3.1 presents a typology of GEs organized by the various levels of variation.

The first two sets of GEs that I will refer to in this chapter are split in terms of *PROTOTYPICALITY* as just discussed. Those GEs that fit the template in Table 3.1 are prototypical GEs and those that do not are non-prototypical. Within each of these sets is another level of variation which I will refer to as the *GENERAL TYPE* of GE. Prototypical GEs are categorized with respect to their quantifier+generic combination—*(∅+)stuff*, *(∅+)thing*, *everything*, *something*, *anything*, and *nothing*. Non-prototypical GEs are grouped into a general type if they share some specific common element—*so*, *what*, *like* etc. Within each general type is a further level of variation, the *SPECIFIC REALIZATION* of the GE. This is the most concrete level of variation and includes every different realization of GEs, regardless of common elements. For example, *and stuff*, *and stuff like that* and *and that sort of stuff* are all treated as different variants. Note that, ellipses indicate that there are a variety of specific realizations that fall under the umbrella of the previous level. For example, *thing* type GEs consist of a similar set of GEs as exemplified by the *stuff* GEs (e.g., *and things*, *and things like that*, and *and that kind of thing*). Both *stuff* and *thing* GEs also have several other specific realizations.<sup>5</sup> Within this level, there are other ways of categorizing variation. For example, I may refer to sets of GEs in terms of the connector used. Adjunctive GEs are those with *and* and disjunctive GEs are those with *or*. I will also refer to prototypical GEs in terms of length. Short GEs (in bold in Figure 3.1) contain no comparatives, while long GEs do contain comparatives.

Throughout the chapter, whatever level of variation is relevant to the present discussion will be made explicit.

---

<sup>5</sup>See Appendix A for a full list of the specific realizations of GEs considered in this chapter, along with the general type to which they belong.

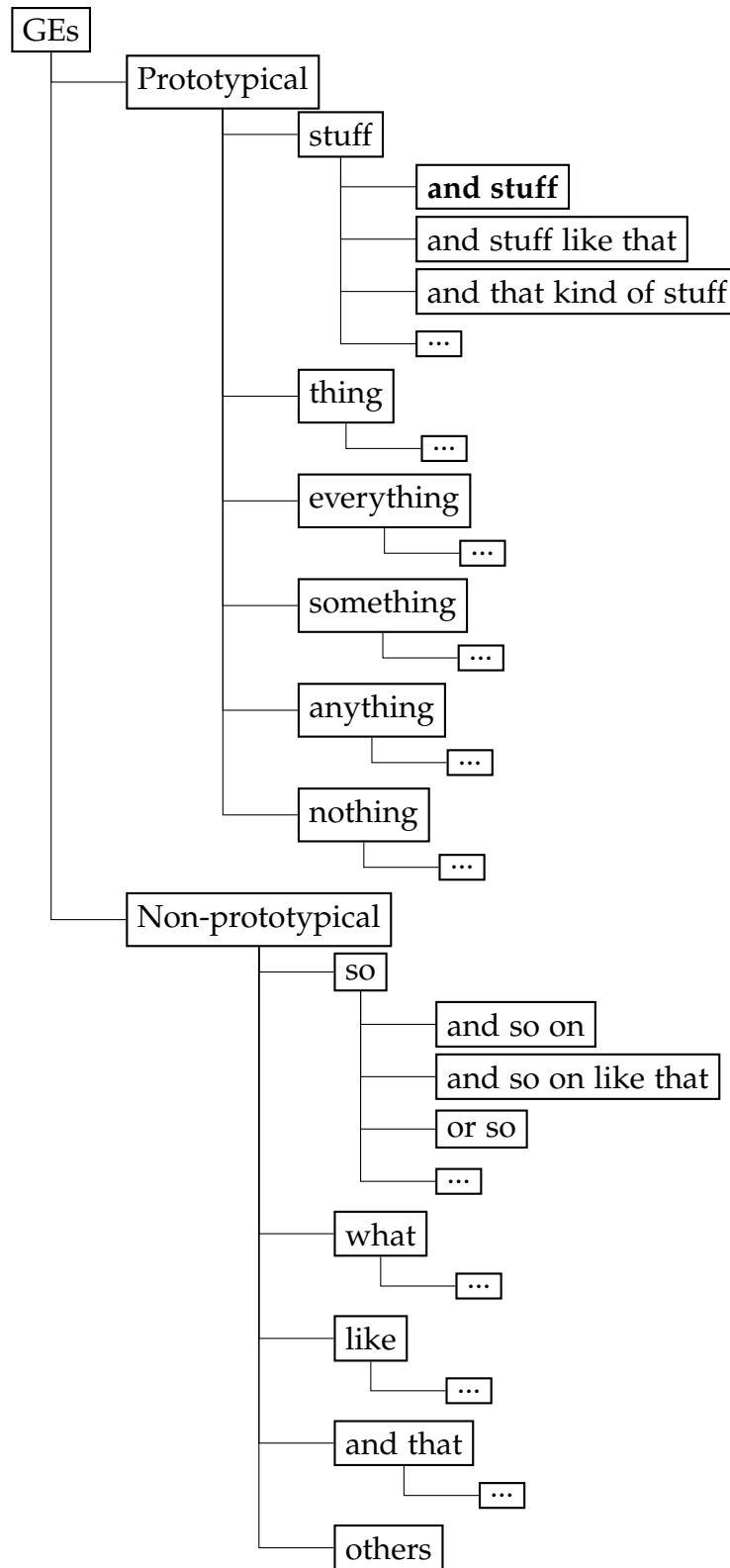


Figure 3.1: A typology of GE variation.

### 3.2.1 Sociolinguistic Approaches to GEs

Previous research on GEs has taken either a sociolinguistic or a pragmatic approach. Early sociolinguistic research has shown that both the frequency and distribution of GEs are conditioned by social factors. A high frequency of use of GEs in general has been associated with working-class speech (Dines 1980, Dubois 1992) and young people (Dubois 1992, Cheshire 2007, Stubbe and Holmes 1995, Tagliamonte and Denis 2010, Denis 2010, Denis 2011). Stubbe and Holmes (1995:72) report an interaction between sex and socio-economic class such that young middle-class females and young working-class males are the primary users. For Stubbe and Holmes, this pattern is associated with change in progress, led by young middle class females. That said, Denis (2010), following Tagliamonte and Denis (2010), argues that a change in overall GE frequency seems unlikely. By compiling the reported normalized frequencies of GEs reported in various studies in the literature (i.e., in a ‘meta-analysis’), the pattern that emerges is one suggestive of AGE-GRADING—the association of particular variants, or differential frequencies of variants with particular phases of life (Chambers 2002: 358). Figure 3.2 above, adapted from Denis (2010) and Tagliamonte and Denis (2010: Figure 2), shows a plot combining the normalized frequencies (per 10 000 words) reported in eight studies from nine different locales on three continents over the last twenty years.

With some measure of confidence we can say that there is a divide between older speakers and younger speakers with respect to the frequency of GEs. The most extensive study in terms of apparent-time depth is Tagliamonte and Denis’ (2010) study of the TEA. In Toronto, speakers over the age of forty use approximately twenty fewer GEs per 10 000 words than speakers under thirty. The data from York (Denis 2010) and Berwick-upon-Tweed (Pichler and Levey 2011:454), both in northeast England, and Wellington (Stubbe and Holmes 1995:72) in New Zealand all exhibit a parallel pattern. Given that all these results are based on data collected at different times and

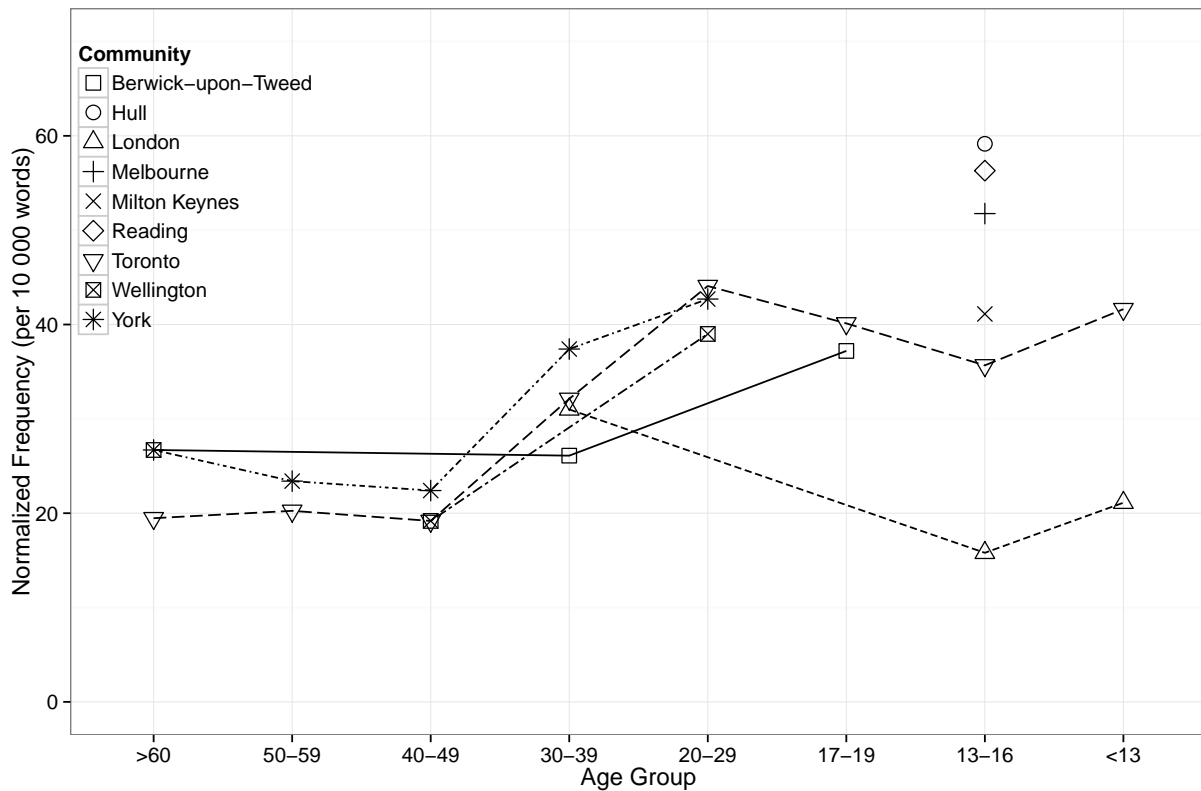


Figure 3.2: A meta-analysis of the frequency of GEs as reported in seven studies. Data originally compiled in Tagliamonte and Denis (2010:Figure 2) with the addition of data from York (Denis 2010).

in different locations, the most likely explanation is age-grading: young speakers use more GEs. The one exception to age-grading is London, where the thirty-somethings in Stenström, Andersen, and Hasund's (2002) study use nearly twice the number of GEs as the teenagers in their corpus and the pre-adolescents in Levey (2007). Taken as a whole however, the studies from Hull, Reading, Milton Keynes (Cheshire 2007) and Melbourne (Winter and Norrby 2001), each of which are equal to or higher than the frequency of older speakers elsewhere, pull up the mean frequency for teenagers.

Individual forms have also been associated with different social and stylistic contexts. Cheshire (2007:165) observes the highest use of *and that* among working class youths while middle class speakers prefer *and stuff* and *and things*. Stubbe and Holmes (1995:79) suggest that *and so on* and *et cetera* are more common in formal and written contexts while *and stuff* and *or something* are more frequent in informal

speech.

Although there is general consensus that the core function of GEs is set-marking, or more formally, that a GE entails 'there is more' (or 'there are alternatives'), the pragmatic and functionalist literature has focused on the multifunctionality of these forms (Aijmer 1985, 2002, 2013; Overstreet 1999, 2014; Overstreet and Yule 1997; Ward and Birner 1992; Winter and Norrby 2000; Youssef 1993).<sup>6</sup> The non-set-marking functions that have been ascribed to GEs include marking politeness, establishing common ground, creating social solidarity, foregrounding information, shifting topic or speaker, hedging, and approximation. This literature is diverse with respect to the methodological and theoretical frameworks with which these pragmatic markers are discussed, ranging from variational pragmatics to compositional semantics to conversation analysis. Thus, many of these functions, though nominally different, overlap in fundamental ways. Generally, we can group these functions into interactional (i.e., having to do with the organization of discourse/speech) and interpersonal (i.e., having to do with the interlocutors attitudes towards the proposition).

Overstreet and Yule (1997:250) focus on the role GEs play in the interpersonal domain as markers of social solidarity, observing that GEs "appear as markers of intersubjectivity in that type of implicit communication through which speakers indicate solidarity, an assumption of shared knowledge and social connection." The argument goes like this: because the core function of GEs is *ad hoc* set-marking and this type of category implication requires intersubjectivity<sup>7</sup> among interlocutors, GEs can function as conventionalized markers of shared knowledge and thus social closeness. Rather than marking vagueness or non-explicitness, GEs indicated invited social solidarity (but see below regarding the distinction between adjunctive and disjunctive

---

<sup>6</sup>While these approaches have their benefits, they also have several problems which I will discuss in §3.4.4.

<sup>7</sup>Intersubjectivity is tied to the idea that "individual subjective experiences of the world are necessarily distinct, yet we often indicate that we assume others share our experiences and hence our knowledge of how things are" (Overstreet and Yule 1997:254).

GEs). In a Gricean pragmatics model (Grice 1975), using a GE rather than being more explicit (and as informative as possible) can trigger an implicature (or invited inference à la Traugott and Dasher 2002): a speaker indicates to a listener that she believes that the listener has enough shared knowledge for the speaker to reconstruct a possible set. Hence, GEs mark social closeness between interlocutors. For example, in (5) from Tagliamonte and Denis (2010:336), a nineteen year old male caps off a list of classic rock bands with *all that stuff*. By using a GE, he signifies to the hearer that ‘although this is not an exhaustive list of bands I listen to, you can fill in the gaps given our shared knowledge about the genre’.

- (5) I like a lot of stuff. Classic rock. Pink Floyd, Led Zeppelin, The Who *all that stuff*.

(TOR/M/1985)

I will return to the implicata of GEs when discussing the grammaticalization mechanism of pragmatic shift in §3.4.4.

Though grammaticalization is not explicitly discussed, Overstreet and Yule (1997) make a number of observations related to grammaticalization theory. They note that those GEs that function more interpersonally, 1) appear to be shorter forms more often, 2) unambiguously attach to nominals less often, and 3) no longer tend to function to mark sets. These observations all relate to the four changes that, together, constitute grammaticalization: phonetic reduction, decategorialization, semantic change, and pragmatic shift (Heine 2003). Likewise, Aijmer (2002) observes that a number of patterns in her data from the London-Lund Corpus of Spoken (British) English can be explained by an appeal to grammaticalization theory: 1) The higher frequency of GEs without comparatives could be due to phonetic reduction (Aijmer 2002:222) and 2) the observed layering (in Hopper’s 1991 sense) of propositional, interactional, and

interpersonal functions can be thought of as a result of semantic-pragmatic expansion (Aijmer 2002:217).

The variationist sociolinguistic literature has focussed on quantitatively testing the hypothesis that GEs are grammaticalizing. In particular, Cheshire (2007) develops four diagnostics to test each of Heine's (2003:579) mechanisms of grammaticalization: PHONETIC REDUCTION or the loss of phonetic substance, DECATEGORIALIZATION or the loss of morphosyntactic properties, SEMANTIC CHANGE or the bleaching of the original meaning of an expression, and PRAGMATIC SHIFT or the use of a linguistic form in new contexts and/or with new pragmatic meanings. Each of these diagnostics, discussed in turn, is tested on a synchronic set of sociolinguistic interviews with adolescents.

Following Aijmer (2002:227), Erman (1995:145) and Overstreet and Yule (1997), Cheshire (2007:167) hypothesizes that any change in the lexical length of GEs is a sign of phonetic reduction.<sup>8</sup> Cheshire (2007:168) observes that longer forms with comparative elements (e.g., *and stuff like that*, *and that kind of thing*) are less frequent in her data of British adolescents than shorter forms without comparative elements (e.g., *and stuff*, *and things*).<sup>9</sup>

To test decategorialization, Cheshire (2007) operationalizes Dines' (1980) and Aijmer's (1985; 2002) mapping of morphosemantic features to particular quantifier-generic combinations. Dines and Aijmer assume that the generic element of each GE ought to depend on the morphosemantic features of its referent, as in (6).

---

<sup>8</sup>Tagliamonte (2012:276) argues that this may not be phonetic reduction at all, but rather a process of morphological clipping. Regardless, Givón's (1990) quantity principle predicts that more grammaticalized forms tend to be shorter than less grammaticalized forms. Likewise, Bybee, Pagliuca and Perkins (1991:38–39) argue that "SHORTNESS" is correlated with grammaticalization.

<sup>9</sup>Tagliamonte (2013:179) codes GEs for length using a three-way distinction: two words long, three words long and more than three words long.



(6)	stuff	things	thing	place	people
	$\left[ \begin{array}{c} +\text{Noun} \\ -\text{Anim} \\ +\text{Conc.} \\ -\text{Count} \end{array} \right]$	$\left[ \begin{array}{c} +\text{Noun} \\ -\text{Anim.} \\ +\text{Conc.} \\ +\text{Count} \\ +\text{Pl.} \end{array} \right]$	$\left[ \begin{array}{c} +\text{Noun} \\ -\text{Anim.} \\ +\text{Conc.} \\ +\text{Count} \\ +\text{Sg.} \end{array} \right]$	$\left[ \begin{array}{c} +\text{Noun} \\ +\text{Place} \end{array} \right]$	$\left[ \begin{array}{c} +\text{Noun} \\ +\text{Human} \\ +\text{Pl.} \end{array} \right]$

The generic form *stuff* should co-occur with a mass noun referent, such as *milk*, while *things*, being a plural (count) noun itself, should co-occur with plural (count) noun referents, such as *puzzles*. However, as Cheshire (2007) and others before her observe, this is not always the case. In fact, GEs frequently appear in coordination with non-nominal constituents. (7) and (8) show examples from the TEA of GEs attaching to *vP*'s and CPs respectively.

(7) a. his sister [<sub>vP</sub> [<sub>vP</sub> 's like a rebel ] and [<sub>vP</sub> has piercings everywhere, ] and [<sub>vP</sub> goes out ] *and all that stuff* ]

(TOR/F/1990)

b. I have to [<sub>vP</sub> [<sub>vP</sub> carry like a map around campus] *and stuff* ]

(TOR/F/1986)

In (7a), *and all that stuff* is capping the end of a list of verbal constituents (*is like a rebel*, *has piercings everywhere* and *goes out*). In (7b), it is possible that *and stuff* is modifying *campus*, but in the context of a conversation about being lost while at university, it is more likely that the speaker must *carry like a map around campus* and other such actions in order to not get lost.

(8) a. I heard that underneath the Taj-Mahal, [CP [CP they bury like the kid ] *or something* ].

(TOR/M/1988)

b. Then we got a radio and we heard everything, how-, [CP [CP what was going on ] *and stuff* ] but it was pretty cool

(TOR/F/1986)

In these two examples, the GE is attaching to a sentential constituent. In (8a), the speaker states that one fact he has tentatively heard about the Taj-Mahal is that *they bury like the kid* underneath it. In (8b), the speaker is talking about listening to the radio to find out what happened during a massive blackout. On the radio she heard the news report about *what was going on* and perhaps why and for how long the power would be out.

Cheshire (2007) operationalizes generics' expected morphosemantic mappings to test decategorialization. The hypothesis is that the more frequently a GE co-occurs with an unexpected referent, the more it is assumed to have decategorialized. Indeed, Cheshire's (2007) comparison of long and short GEs suggests that short forms (i.e., the more phonetically reduced forms) are more decategorialized than long forms (i.e., the less phonetically reduced forms). Taken together, these observations point to the possible grammaticalization of certain GEs.

Following the pragmatics literature, Cheshire (2007) takes the core meaning of GEs to be generalizing a set of entities from properties of the preceding referent or simply to mark a set (Aijmer 2002, Erman 1996, Overstreet and Yule 1997, Stubbe and Holmes 1995). To determine the extent of semantic change, Cheshire (2007) examines the extent to which this set-marking function has been bleached. Tokens in which no set-marking function is inferable are argued to be at a more advanced stage of semantic change. Cheshire (2007) observes that the set-marking function is categori-

cal with all long GEs (e.g., *and things like that*) but often not inferable with short GE forms (e.g., *and things*). That the variants at the forefront of phonetic reduction and decategorialization among Cheshire's (2007) adolescent data were also those in which the core meaning could not be inferred further supports the hypothesis that GEs are grammaticalizing. This is exemplified by short form *and things*, which attaches to a non-nominal constituent in a discussion about whether smoking should be allowed at school in (9) from Cheshire (2007:177).

- (9) I think they [teachers] say 'erm ... should you be doing that?' but maybe if like parents have given permission *and things* first they probably like might think 'why am I saying that?' you know but I think it would be because it's just like I don't know it's not really a school thing is it?

The speaker does not intend to suggest a list of possible things that parents might do in order for smoking to be allowed at school. No 'there is more' is entailed. Rather, the speaker is only suggesting that if *parents have given permission* than maybe smoking could be allowed. Thus, no set-marking function is present.

Lastly, Cheshire (2007) operationalizes collocation with discourse markers to examine the extent to which GEs have pragmatically shifted.<sup>10</sup> She argues that if a GE co-occurs with another discourse marker, then the GE is less grammaticalized than one that does not co-occur with a discourse marker. In the former case, the GE has a more propositional role, while the co-occurring discourse marker is providing the interactional and/or interpersonal function. In the latter case, the GE is assumed to be more likely functioning in the interactional/interpersonal domains itself. Cheshire (2007) finds that the GE variants most advanced in terms of phonetic

<sup>10</sup>Cheshire (2007:185) notes that her original intention was to code for and quantify the pragmatic function of each token. However, identifying a main function is complicated and difficult because these features can be simultaneously polysemous (i.e., can serve multiple functions at the same time). But see below for a discussion of Pichler and Levey (2011) who took such a quantitative approach to pragmatic function.

reduction, decategorialization and semantic change are also those most advanced in terms of pragmatic shift, as operationalized by co-occurring discourse markers.

Cheshire's (2007) evidence for grammaticalization of GEs appears strong. There is a high frequency of phonetically reduced GEs. In contrast with longer forms, these phonetically reduced forms have higher levels of decategorization, have more tokens in which a set-marking function was not apparent, and in turn have a lower co-occurrence with discourse markers. However, her focus is on a single generation and it is unclear to what extent, if any, the set of changes associated with grammaticalization are occurring through time (Tagliamonte and Denis 2010:357). Dovetailing from this question, Tagliamonte and Denis (2010) and Pichler and Levey (2011) apply these same diagnostics using apparent-time data (see below). To keep things straight, Table 3.2 summarizes the diagnostics and their predictions for change over time.

Table 3.2: Tests of grammaticalization through apparent time.

Change	Diagnostic	Prediction for change over time
Phonetic reduction	Length	Increase of short forms
Decategorialization	Type of referent	Increase of morphosemantic mismatch
Semantic change	Function to mark a set	Increase of GEs without set marking function
Pragmatic shift	Co-occurrence with DMs	Decrease of co-occurring discourse markers.

Tagliamonte and Denis (2010) take Cheshire's (2007) work as a starting point and apply these diagnostics to apparent time data. The goal was to determine if changes were happening in the GE system of Toronto English and, if so, could those changes be described as a case of grammaticalization? The data include 87 speakers interviewed in 2003, ranging in age from 9 to 85, which, according to the apparent time construct (Bailey, Wikle and Tinnery 1991), can be taken to represent a snapshot of progress through a large portion of the 20<sup>th</sup> century, from the 1920s to the 1990s.

Using variationist methodology, Tagliamonte and Denis (2010:358) report a change

in progress such that GEs composed of the generic *stuff* are increasing at the expense of GEs composed of the generic *things*. Tagliamonte and Denis' (2010:358) Figure 8 is presented here in Figure 3.3.

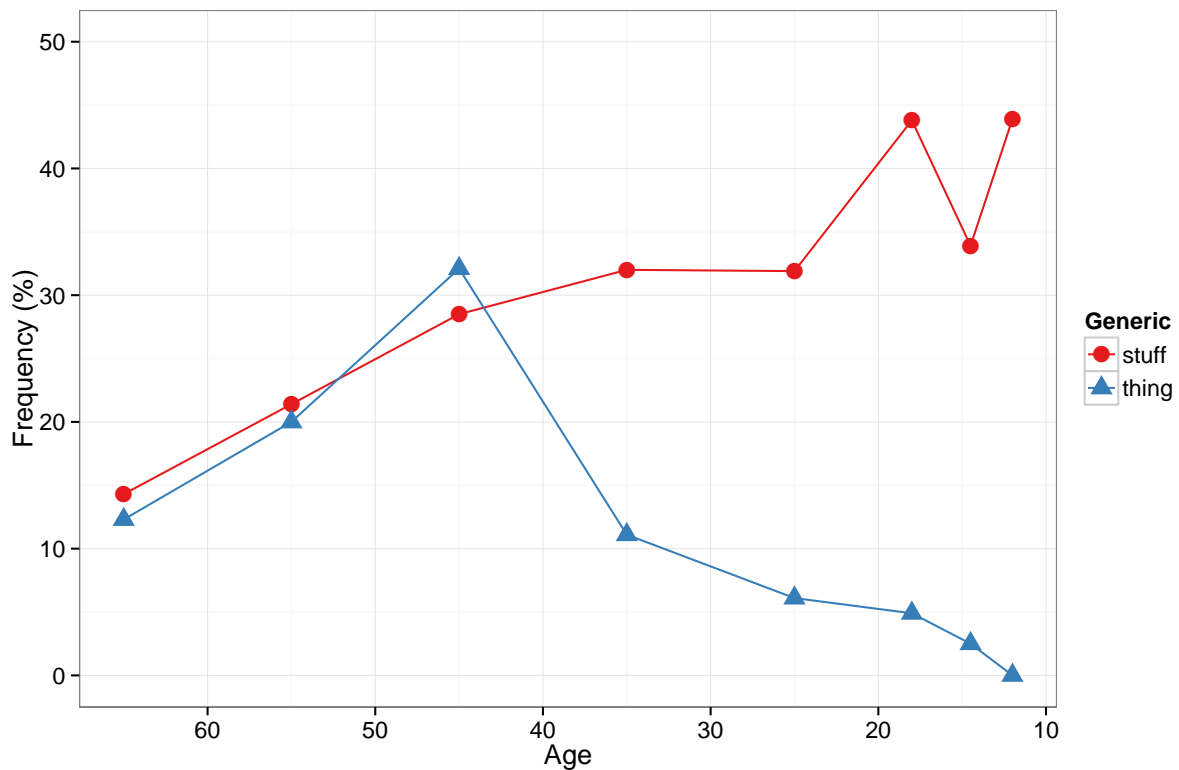


Figure 3.3: Frequency of *stuff* and *thing* type GEs as a proportion of all general extenders through apparent time in Toronto, based on Tagliamonte and Denis (2010:358). Total N =1159.

Figure 3.3 shows that GEs with both *stuff* and *thing* generics appear to increase for the first few generations. However, for those younger than forty, *stuff* forms increase while *thing* forms decrease. These results are confirmed by logistic regression (Tagliamonte and Denis 2010:360). Crucially, this figure demonstrates the trajectory of *stuff* forms from a marginal variant to the majority variant. Such surges in frequency have been argued to be signposts of grammaticalization (Mair 2004:126).<sup>11</sup> Because grammaticalization is argued to be a gradual process, if the development of *stuff* GEs is

<sup>11</sup>That said, Mair and grammaticalization theorists typically discuss surges in the raw frequency of a linguistic form in some specific context rather than the proportional increase of one variant embedded within a variable system.

best thought of as grammaticalization, than over the course of this change, evidence of increased grammaticalization over time should be apparent. Thus, Tagliamonte and Denis (2010) consider the diagnostics used by Cheshire (2007) at different points in apparent time contrasting speakers older than 50, between 30 and 50 and under 30. The results of these four tests on *stuff* type GEs are reported in Table 3.3.

Table 3.3: Diagnostics of grammaticalization of *stuff* type GEs through apparent time in Toronto English (Tagliamonte and Denis 2010).

Diagnostic	Result
Phonetic Reduction (p. 351–2)	No significant difference in the frequency of long and short forms across the three age groups.
Decategorization (pp. 352–3)	No change in the distribution of referents. Non-nominals always favoured.
Semantic Change (pp. 355–6)	Some change in the frequency of forms without a set marking function. Much less frequent than reported by Cheshire (2007). “[M]odest at best.”
Pragmatic Change (pp. 356–7)	Type and token frequencies of co-occurring discourse markers are increasing, not decreasing, through apparent time.

Tagliamonte and Denis (2010:357) conclude with respect to grammaticalization that “the system from adolescents to octogenarians is stable.” Although there are some signs of grammaticalization, it did not happen during the apparent time span of the data (1916–1993). That is, there is no evidence for *ongoing* grammaticalization in Toronto. While this is only a single case, two other cases studies, both from the northeast of England, report evidence that the grammaticalization are not ongoing. Denis (2010) examines change in the GE system of the York speech community, while Pichler and Levey (2011) consider the system in Berwick-upon-Tweed.

Denis (2010) follows the exact methodology laid out in Tagliamonte and Denis (2010), only varying the geographic context. In York, the most frequent GE is *and that*. However, embedded within the GE system is a new innovation; across the generations in the York Corpus, GEs with *stuff* have innovated and grown in frequency

(see Denis 2011 for discussion of the speakers on the vanguard of this innovation). Figure 3.4 shows the apparent-time rise of *stuff* forms from non-existence, to innovation, to community-wide adoption.

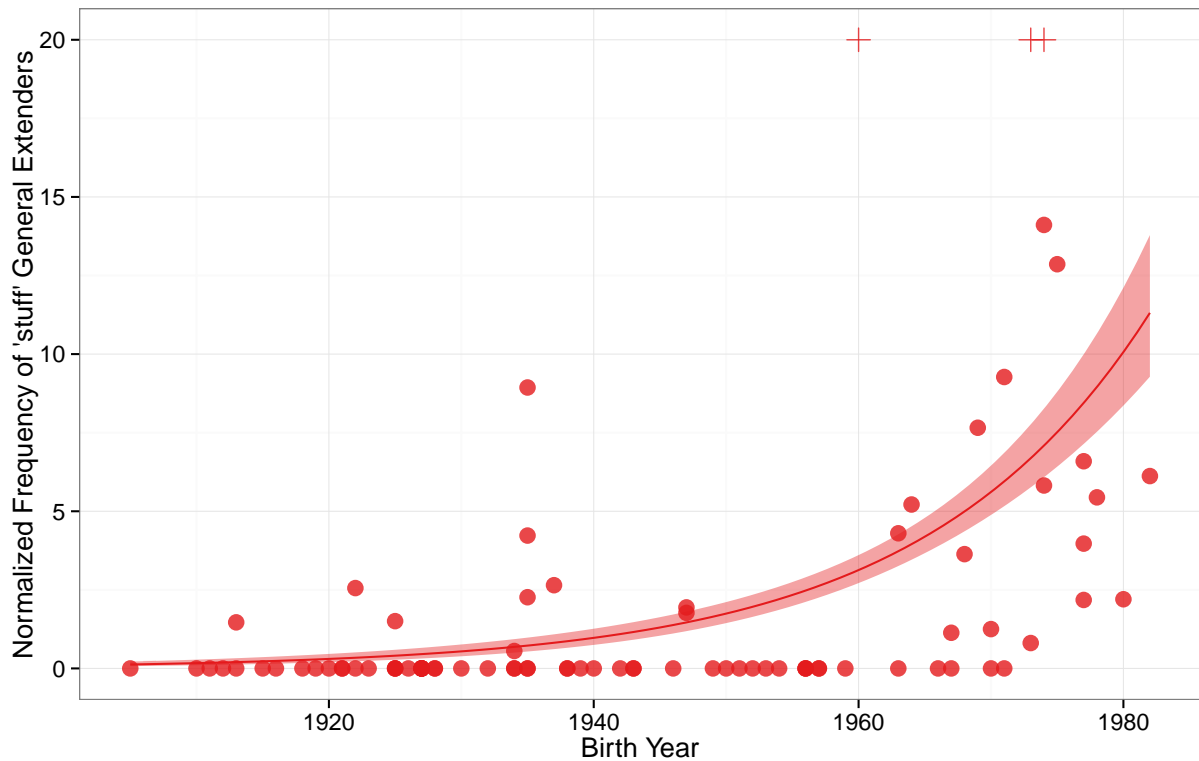


Figure 3.4: Scatterplot of the normalized frequency of *stuff* forms in York English, based on data in Denis (2010) and Denis (2011). (Crosses at the >10 line represent out of range data. Line and error ribbon plot the coefficient of Frequency by Birth Year from a poisson regression). Total *stuff* GEs = 186. Total GEs = 2156.

Thus, as in Toronto, the data in York provide an opportunity to test the diagnostics of grammaticalization over time for a GE in the process of change. Following Cheshire (2007) and Tagliamonte and Denis (2010), Denis (2010) tests the phonetic reduction, decategorization, semantic change and pragmatic shift of *and stuff* at three points in apparent time. The results are summarized in Table 3.4.

Denis (2010) concludes that although *and stuff* entered the GE system in York, constituting a change in these pragmatic markers, there is no evidence to suggest that this development involved grammaticalization. Instead, the form diffused to

Table 3.4: Diagnostics of grammaticalization of *stuff* type GEs through apparent time in York English (Denis 2010).

Diagnostic	Result
Phonetic Reduction (p. 26–7)	Significantly more short forms than long forms in youngest generation, but long forms are marginal in older generations. Contrary to grammaticalization theory, phonetically reduced variant does not appear to develop out of longer form.
Decategorization (pp. 34–6)	No change in the distribution of referents. Unexpected referent always favoured.
Semantic Change (pp. 41–44)	Qualitative analysis shows that <i>and stuff</i> is multifunctional in the same way as other GEs are among young people.
Pragmatic Shift (pp. 46)	Frequency of co-occurring discourse markers are stable through apparent time.

the community as a “pre-packaged collocate—a fixed expression, akin to *and whatnot* and *and so on*” (Denis 2010:55). Due to a lack of the longer form, phonetic reduction could not have applied. Neither is there evidence for semantic-pragmatic expansion or decategorialization. The form entered the system and increased in frequency, functioning in the same way as competing variants (e.g., *and things like that*) from the beginning.

Pichler and Levey (2011) follow the same general methodology as Tagliamonte and Denis (2010), operationalizing Cheshire’s (2007) diagnostics through apparent time.<sup>12</sup> Their analysis of Berwick-upon-Tweed “turned up no compelling evidence of grammaticalization in progress” (Pichler and Levey 2011:461–462). However, they hypothesize that the stable pattern observed in Berwick-upon-Tweed, York, and Toronto could be the result of *earlier* but now *arrested* grammaticalization. That is, grammaticalization may have occurred, or even started but stalled, prior to the time range of these studies. Pichler and Levey (2011:462) point out that grammaticalization theory allows for this. Hopper and Traugott (1993:94) argue that grammaticalization is not

<sup>12</sup>Pichler and Levey (2011) take a nuanced approach to semantic and pragmatic change. I discuss this further in §3.4.4.



required, or expected, to reach completion:

“A particular grammaticalization process may be, and often is, arrested before it is fully ‘implemented’, and the ‘outcome’ of grammaticalization is quite often a ragged and incomplete subsystem that is not evidently moving in some identifiable direction.”

Thus, Pichler and Levey (2011:462) conclude with respect to their apparent-time findings and Tagliamonte and Denis’ (2010) findings, that:

“[I]t remains to be determined whether the synchronic stable patterns of GE variability [...] are in fact the product of grammaticalization processes that may have been operative at an earlier stage of the language, predating the time-span in our corpus.”

Despite this possibility, a strong case against gradual grammaticalization remains for York. The apparent-time range of the study captures the entire ‘lifespan’ of the *stuff* type GEs in that community. The development of this variant does not follow what is expected by grammaticalization theory. Thus, what is needed to address Pichler and Levey’s (2011) concern about Tagliamonte and Denis’ (2010) results from Toronto is an extended diachronic point of view of this variety that includes a time period when contemporary innovations in the GE system were non-existent or incipient. With this type of data, we can track the development of these innovations and address the issue of whether what we see in Toronto is the result of previous, arrested grammaticalization. Pichler and Levey (2011:462, 464) explicitly call attention to the need for such work:

“While stable patterns of GE variability may in theory be compatible with a grammaticalization scenario, such a scenario awaits confirmation from diachronic analysis. Without **an appropriate real-time benchmark**, reconstruction of the diachronic transitions that have given rise to contemporary patterns of GE variability, as well as inferences that these patterns are the result of change, must remain speculative.”

“A much-needed extension to current variationist work on GEs is the incorporation of **a real-time component** to increase the time-depth of data analysed. The exploration of diachronic surrogates or oral data may enable us to ascertain the extent to which GEs have been implicated in

change, and to determine the role of grammaticalization in any trajectories of change uncovered.”

This is the point of departure for this chapter. The EOE is just such “an appropriate real-time benchmark” for examining the diachrony of GEs in Ontario English.

### 3.3 General Extenders in EOE

In this section, I present the overall distribution of GEs in EOE in comparison to Tagliamonte and Denis’ (2010) results for Toronto and the trajectory of change of the variants in the Ontario English system. The emerging picture is one of change within the system. Although the frequency with which speakers use GEs is relatively stable across time, variants—and crucially the variety of variants—have changed.

#### 3.3.1 Overall Distribution

Following the majority of the literature on GEs, the GE data in EOE is partitioned between adjunctive GEs (i.e., GEs with an *and* connector) and disjunctive GEs (i.e., GEs with an *or* connector). Overstreet (1999) observes that adjunctive and disjunctive GEs function differently with respect to face (Goffman 1967) and politeness strategies (Brown and Levinson 1987). Adjunctive GEs are said to function with respect to positive politeness; they are linguistic devices that appeal to social solidarity between interlocutors, indicating “that the speaker and hearer belong to the same group and have certain things in common” (Overstreet 1999:98). In other words, by using an adjunctive GE the speaker asserts common ground between the speaker and the hearer and indicates that “[b]ecause we share the same knowledge, experience, and conceptual schemes, I do not need to be explicit; you will be able to supply whatever unstated understandings are required to make sense of my utterance” (Overstreet 1999:99). Overstreet (1999) exemplifies this function in (10).

(10) Pam: Still having trouble with that equilibrium stuff?

Bob: Wh- yeah. It- well- just walking.

Pam: Mmhm.

Bob: Y'know uh if I pay attention I can do pretty good on flat surfaces but as soon as the surface gets uneven *an' stuff* um ... I- it- yeah it gets a little tricky.

Pam: Hmm. Bummer. (Overstreet 1999:100, ex. 4)

Bob uses an adjunctive GE in an appeal “for understanding” that he has difficulty hiking with poor equilibrium and in a response to this appeal, Pam sympathizes offering ‘Hmm. Bummer’. (Overstreet 1999:100)

Conversely, disjunctive GEs function with respect to negative politeness and are used as hedges which “minimize a threat to the hearer’s [...] face, [...] avoid imposing, and [...] mark deference” (Overstreet 1999:98). By using a disjunctive GE, the speaker indicates to the hearer a weak commitment to the proposition, saving the speaker’s face (in case of possible rejection of the proposition) and reducing the risk of threatening the hearer’s face (by reducing the pressure put on the speaker to do as requested) (Overstreet 1999:104–105). This hedging function is exemplified in (11) from Overstreet (1999).

(11) Sara: An uh uh- I’ll see ya... I- may talk to y’all sometime next week, but if not I’ll see y’all Friday a- probably Friday afternoon=Friday evening when y’all get in.

Roger: Okay.

Sara: I know y’all’ll be real tired, but hopefully we’ll like hookup *or something*.

Roger: Sure.

Sara: An’ uh I’ll talk to you later.

Roger: All right. (Overstreet 1999:106, ex. 15)

Sara avoids imposing on Roger by using the disjunctive GE which indicates that the request to 'hookup' on Friday evening, even though Roger will be tired, is only tentative.

It has been widely observed that the use and frequency of pragmatic markers in different types of discourse with different purposes, different topics, and between different types of interlocutors varies (Pichler 2010:584–6; Macaulay 2002). Because speakers are particularly sensitive to these differences, the use of adjunctive and disjunctive GEs can be operationalized as a way to ensure comparability across datasets. Data of a similar sort ought to contain roughly the same frequency of GEs and the same distribution of GE types. In other words, if two data sets contain roughly the same frequency of GEs and the same distribution of the types of GEs, we can assume that the data is of a similar sort precisely because the frequency with which speakers employ politeness strategies of different types varies depending on the range of factors just mentioned. This hypothesis is consistent with Overstreet's (1999:6–7) findings. Not only is there a marked difference in the overall frequency of GEs but also in the distribution of types in her two corpora. The corpus of informal speech among familiars contained more than five times as many GEs than the corpus of formal speech among non-familiars overall and where the informal corpus contained more disjunctive than adjunctive GEs, the formal corpus contained more adjunctives.<sup>13</sup>

Table 3.5 presents the overall normalized frequency of GEs in Belleville, Eastern Ontario, Niagara, and three age groups from Toronto. Included for comparison is also the overall normalized frequency of GEs used by adolescents in the three English towns reported by Cheshire (2007:161).<sup>14</sup>

First, note that the total normalized frequency of GEs in Belleville, Eastern Ontario and Niagara are roughly the same, around 31 GEs per 10 000 words. This is lower

---

<sup>13</sup>Although the frequencies Overstreet (1999:7) reports are not normalized, both corpora were ten hours in length.

<sup>14</sup>The Total column in the Canadian data is higher than the sum of the Adjunctive and Disjunctive columns because the Total column also includes GEs without a connector (e.g., *stuff like that*).

Table 3.5: Normalized frequency of adjunctive and disjunctive GEs in EOE, TEA, and Cheshire 2007.

	Adjunctives (N)	Disjunctives (N)	Total (N)	Word Count
BLV	17.6 (90)	7.1 (36)	27.6 (141)	51 065
EON	15.8 (107)	15.8 (107)	32.6 (220)	67 577
NIA	21.4 (179)	9.0 (75)	33.7 (282)	83 769
TOR>60	11.4 (177)	6.0 (93)	19.6 (306)	155 897
TOR30–59	13.9 (294)	9.2 (195)	25.0 (528)	211 594
TOR<30	20.9 (688)	18.1 (597)	40.8 (1344)	329 294
Reading	33.3 (234)	23.0 (162)	56.3 (396)	70 320
Milton Keynes	28.5 (238)	12.7 (106)	41.1 (344)	85 539
Hull	45.0 (343)	14.2 (108)	59.2 (451)	76 236

than the youngest people in Toronto but higher than the two age groups over 30. This may be an indication that the age-grading hypothesis discussed in §3.2.1 with respect to Figure 3.2 may not be a hard and fast rule, as the elderly speakers in EOE are using GEs at a greater frequency than the older speakers in TEA. That said, the normalized frequencies are not too far off from the eldest speakers in York (26.7 GEs per 10 000 words). Crucially, the frequencies in Belleville, Eastern Ontario, and Niagara are similar and fall within the bounds of the TEA, not reaching the higher range reported by Cheshire (2007) for adolescents in Reading, Milton Keynes, and Hull. This is the first indication that suggests that the two Canadian corpora are sufficiently comparable. Second, consider the relative distribution of adjunctive and disjunctive GEs in each community/age group. In every community there is a higher (or at least an equal frequency) of adjunctive GEs than disjunctive GEs. As Cheshire (2007:160–161) notes, this is consistent with Levey’s (2007) sociolinguistic interviews with pre-adolescents in London. Although this pattern corresponds with Overstreet’s (1999) formal corpus, rather than the informal corpus, the important point is that the frequency of GE types in corpora built of sociolinguistic interviews are consistent across communities suggesting that the data may be compared effectively.

Table 3.6: Top 6 GE types in EOE communities (%Top 6;  $H_{rel}$ ).

	Belleville (51%; 0.69)	E. Ontario (63%; 0.64)	Niagara (52%; 0.69)	All EOE (52%; 0.73)
1.	<i>and so on</i>	<i>and so on</i>	<i>and so on</i>	<i>and so on</i>
2.	<i>or something like that</i>	<i>or something</i>	<i>and things like that</i>	<i>or something like that</i>
3.	<i>and all that</i>	<i>or something like that</i>	<i>or something like that</i>	<i>or something</i>
4.	<i>or something</i>	<i>or so</i>	<i>and that</i>	<i>and things like that</i>
5.	<i>and that sort of thing</i>	<i>and things like that</i>	<i>and everything</i>	<i>and that</i>
6.	<i>and the like of that</i>	<i>and all that</i>	<i>or something</i>	<i>and all that</i>

Table 3.7: Top 6 GE types in TEA age groups (%Top 6;  $H_{rel}$ ).

	TOR $\geq$ 60 (55%; 0.65)	TOR 30-59 (67%; 0.61)	TOR $<$ 30 (80%; 0.53)	All TEA (71%; 0.60)
1.	<i>or something</i>	<i>or something</i>	<i>and stuff</i>	<i>and stuff</i>
2.	<i>and everything</i>	<i>and stuff</i>	<i>or something</i>	<i>or something</i>
3.	<i>and so on</i>	<i>or whatever</i>	<i>or whatever</i>	<i>or whatever</i>
4.	<i>and that</i>	<i>and things like that</i>	<i>and stuff like that</i>	<i>and stuff like that</i>
5.	<i>and stuff like that</i>	<i>and stuff like that</i>	<i>and everything</i>	<i>and everything</i>
6.	<i>or something like that</i>	<i>or something like that</i>	<i>or something like that</i>	<i>or something like that</i>

The next cross-community comparison to make is with respect to the specific realization of GEs across communities. Tables 3.6 and 3.7 list the six most frequent GE variants in EOE (separated by community) and TEA (separated by age group) respectively. The two numbers in parentheses in these tables are the percentage of all GEs in the respective community/age group that the top six variants represent and the relative entropy of the variation in each community/age group. RELATIVE ENTROPY ( $H_{rel}$ ) is a measure of dispersion for categorical variables that is bounded by 0 and 1 (Gries 2009:112). An  $H_{rel}$  of 0 means that all data points are represented by a single variant. If there were three variants ( $\alpha$ ,  $\beta$ ,  $\gamma$ ) and three hundred tokens, a community would have an  $H_{rel}$  of 0 if all three hundred tokens were variant  $\alpha$ . An  $H_{rel}$  of 1 means that each variant occurs with equal frequency. Again, if there were three variants and three hundred tokens, a community would have an  $H_{rel}$  of 1 if one hundred tokens were variant  $\alpha$ , one hundred were variant  $\beta$ , and one hundred were variant  $\gamma$ . In this case, the lower the  $H_{rel}$  value for a community/age group, the less *variety* of GEs that community/age group exhibits.

The youngest speakers in Toronto have the lowest  $H_{rel}$  (0.53). The group with the next lowest  $H_{rel}$  is the middle age group in Toronto with 0.61. The oldest speakers in Toronto follow next at 0.65. Lastly, when all three communities are collapsed, the EOE has the highest  $H_{rel}$  (0.73). That said, Belleville and Niagara are markedly higher than Eastern Ontario, which itself patterns similarly to the oldest speakers in Toronto.<sup>15</sup> The trend is clear: the older speakers have the most variety and the youngest speakers exhibit the least variety. Thus, over the 20<sup>th</sup> century, there has been a marked decrease in the variety of GEs used in Ontario English.

Focussing in on the specific variants in the top six, several patterns emerge. First, *and so on* is consistently the most frequent form in EOE. The variant appears in the top six with the oldest speakers in Toronto, but then falls out of favour with the middle aged and youngest speakers. Second, *or something (like that)* is the most consistently favoured form, appearing in the top six in every group. *And (all) that* follows a trend similar to *and so on*. The variant appears in the top six for all the EOE communities and the oldest speakers in Toronto, but then loses ground. Lastly, although *and stuff (like that)* is outside of the top six list for all the EOE communities (though not completely absent, see below), it steadily rises through the ranks to reach the number one spot among the youngest speakers in Toronto. Furthermore, it is the only variant to exhibit this pattern. This is consistent with Tagliamonte and Denis' (2010:358) observation that *and stuff (like that)* is rising in frequency in Toronto English.

### 3.3.2 The Long-Term Trajectory

Dovetailing from this observation, Figures 3.5 and 3.6 plot the relative frequency of the most common adjunctive and disjunctive GE types as a proportion of the total

---

<sup>15</sup>However, the proportion that the top six variants in Eastern Ontario make up is comparable to the top six variants in several enclave UK speech communities as reported by Tagliamonte (2013:176). The top six variants—*and that*, *and all*, *or something*, *and everything*, *or anything*, and *and things*—make up sixty two percent of all variants.

number of adjunctive and disjunctive GEs respectively, binning speakers by their year of birth in decade long intervals from 1890 to 1990.<sup>16</sup> The apparent-time increase of

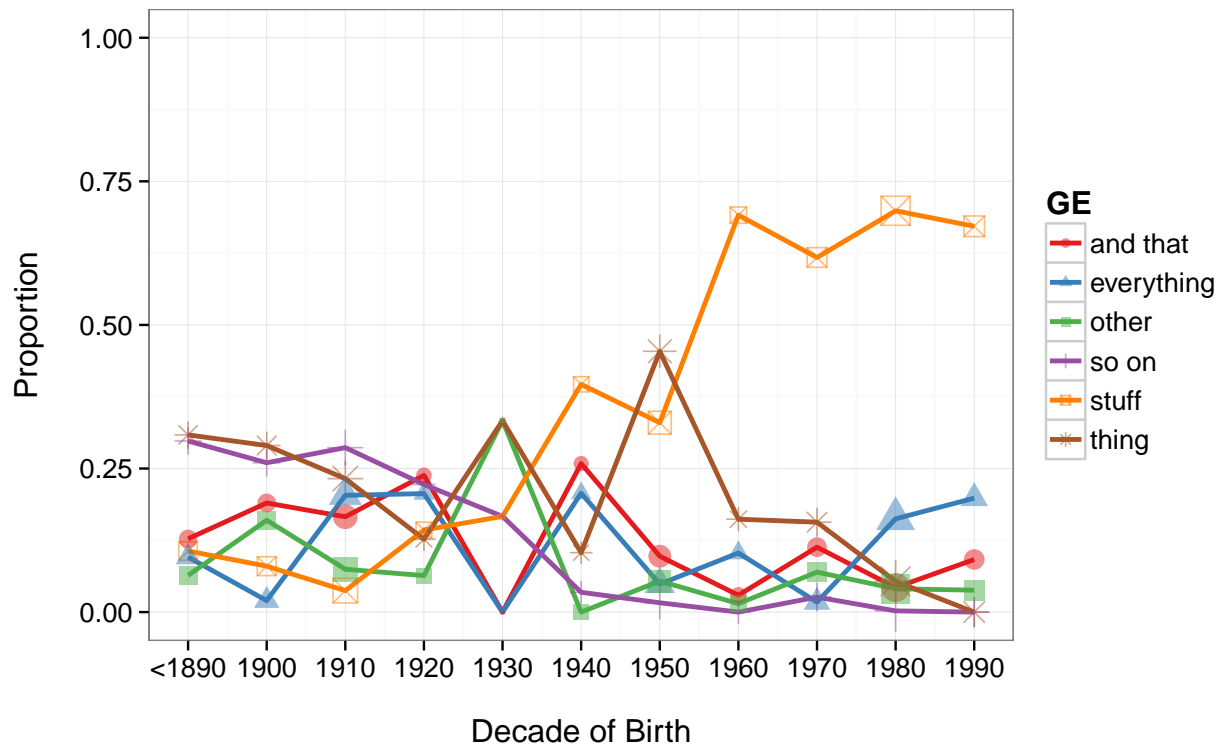


Figure 3.5: Proportion of main variants of just adjunctive general extenders over 100+ years of apparent time in Ontario English. Data point size indicates relative raw frequency.  $N = 1535$ .

$H_{rel}$  and rise of *and stuff* (*like that*) both observed above play out as a dramatic change across the twentieth century in Figure 3.5.<sup>17</sup> Concentrating on the oldest speakers, born from before 1890 to 1930, we see a a variable system with multiple variants with proportions hovering between 0.15 and 0.30. *And thing* (hex-star point) and *and so on* (cross point) type GEs lead, though we begin to see a decrease of both forms by the 1920s. Among these older speakers, *and that* (circle point) appears to be on

<sup>16</sup>The data in the 1890s bin includes speakers born between 1879 and 1889 in addition to those born in the 1890s. The data was collapsed due to the small speakers Ns in the 1870s and 1880s (1 and 2 respectively). The EOE data is also collapsed across communities. See Appendix A for a complete list of all GE forms in EOE and their frequencies.

<sup>17</sup>The frequencies of *stuff* and *thing* in Figure 3.5 differ from those in Figure 3.3 because while the latter calculates the relative frequency as a proportion of all GEs, the former does so as a proportion of just the adjunctive GEs.



the rise while other forms including *and everything* and *and stuff* rank lower.<sup>18</sup> After the perturbations among the 1930s speakers, where token Ns are generally low, a different picture is apparent: *and stuff* (open-box point) takes over the system and, despite *and things* exhibiting a peak in the 1950s, all other adjunctive variants begin to decrease in frequency. This expanded picture of GEs in Ontario English across one hundred years of apparent-time indicates that there is a more complex story than that of Tagliamonte and Denis (2010:358), who suggest that *and stuff* was replacing its next major competitor *and things*. Rather, the rise of *and stuff* resulted in a massive reduction of variation. Where the oldest speakers exhibit robust variation of variants, the youngest speakers have a variable system that is dominated by a single variant.

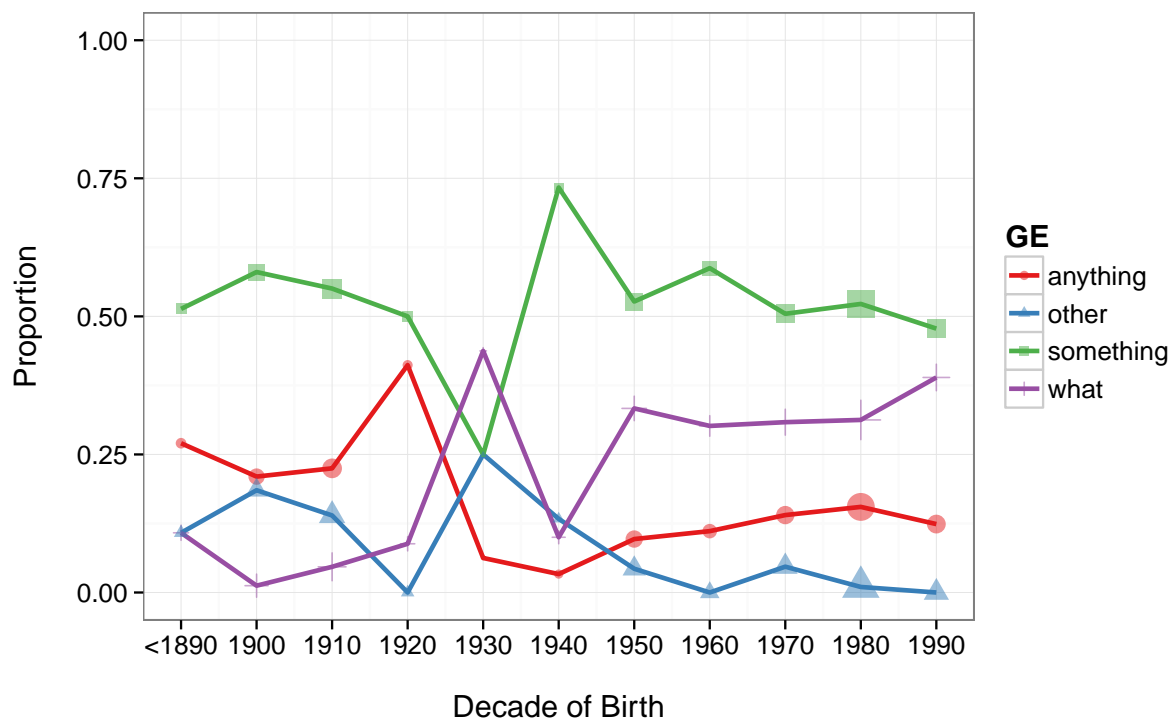


Figure 3.6: Proportion of main variants of disjunctive general extenders over 100 years of apparent time in Canadian English. Data point size indicates relative raw frequency. N = 1103.

<sup>18</sup>Tagliamonte (2013:184) argues that her data from enclave northern English dialects suggests that the contemporary frequency of *and that* in the UK is a result of “retention of a conservative northern feature” (cf. Cheshire 2007; Pichler and Levey 2011). Note that the non-marginal frequency of *and that* in earlier Ontario English supports the idea that this variant is conservative.

The trajectory of disjunctive GEs shown in Figure 3.6 is far more stable than the adjunctive variants. *Or something* (square point) is the most frequent disjunctive variant, with a proportion hovering around 0.50 across the entire hundred years. The secondary forms exhibit some change. *Or what* variants (cross point), including *or whatever*, *or whatnot* etc., are infrequent among the oldest speakers, but after the 1940s rise to the second most frequent disjunctive variant. This rise from around twelve percent to around forty percent seems to be at the expense of the other variants equally.

### 3.4 The Long-Term Trajectory of the Mechanisms of Grammaticalization

In this section, I test the hypothesis that the rise in frequency of *stuff* type GEs, from an incipient stage, is a signpost of grammaticalization of GEs in Ontario English. I replicate the diagnostics of each mechanism of grammaticalization on the EOE data following both Cheshire (2007) and Tagliamonte and Denis (2010). For each subsection, I first present the results of this replication, using EOE as a fourth and older age group to compare against the Toronto data. I then provide an updated analysis of the trajectory of each mechanism. For the case of phonetic reduction, this will involve new quantitative arguments based on the results from innovative statistical tools. For semantic change and pragmatic shift, I provide new arguments based on the semantics and pragmatics of GEs. My discussion of decategorialization will make use of both quantitative and theoretical arguments.

### 3.4.1 Phonetic Reduction

#### Background

The first mechanism of grammaticalization that will be tested is phonetic reduction. As discussed above, Cheshire (2007) follows a number of researchers who have suggested that the syntagmatic length of a GE can be used as a diagnostic of phonetic reduction. The idea is that phonetic reduction affects GEs such that whole morphemes (or components of the template in Table 3.1) are eroded away with time. Thus, short GEs are derived from longer GEs of the same type and this shortening is theorized to be part of a gradual grammaticalization process.

There are many cases of phonetic reduction in the grammaticalization literature. Although such attrition usually takes place at the syllabic or phonemic level, one other example of phonetic reduction effecting whole morphemes at a time comes from another set of pragmatic markers, epistemic parentheticals (EPs) (see chapter 4). Thompson and Mulac (1991:317) test the hypothesis that EPs, as in (12a), are “grammaticalized forms of subjects and verbs introducing complement clauses,” as in (12b).

- (12) a. *I think that* we’re definitely moving towards being more technological.  
 b. *I think* ∅ exercise is really beneficial, to anybody.

(Thompson and Mulac 1991:313)

The deletion of *that* in (12b) is crucial. For Thompson and Mulac (1991), *that*-deletion is not an example of a phonologically overt complementizer alternating with a phonologically null complementizer. Rather, they argue that the difference between (12a) and (12b) represents alternation between *I think that* introducing a complement clause and *I think* “functioning roughly as an epistemic adverb such as *maybe* with respect to the clause it is associated with” (Thompson and Mulac 1991). Implicit in this dis-

tion is that the grammaticalized form has been phonetically reduced by a whole morpheme, no longer occurring with *that*.

EOE, like TEA, exhibits variation in the length of GEs. Even among the oldest speakers in EOE there are examples of short GEs, as in (13).

- (13) a. When I was younger, you know, I worked away quite a bit, you know.  
When he could do the plowing then *and stuff* himself you know.

(EON/M/1898)

- b. We made quilts and we made everything that you could imagine, babies' clothes *and things* and sometimes we made as high as two or three-thousand dollars at our bazaars.

(BLV/F/1879)

- c. Sometimes if they had a heavier load *or something* they'd put three horses on.

(NIA/F/1899)

- d. It was a beautiful house and all that. Lovely grounds all around it *and everything*

(BLV/F/1884)

This is unsurprising in light of the fact that short GEs have a long history in English usage, as in (14).

- (14) a. With Ruffes and Cuffes, and Fardingales, *and things*.

(W. Shakespeare, *Taming of the Shrew*, iv. iii. 56, 1616)

- b. Your Noveds, and Blutraks, and Omurs *and Stuff*, By God, they don't signify this Pinch of Snuff.

(J. Swift, *Grand Question Debated*, 1729)

That said, each region of the EOE contained at least one example of a long GE form that does not appear in TEA, *and so on like that* as in (15).<sup>19</sup>

- (15) We didn't hang stockings. Each one of us would set our plate on the table at night and in the morning of course before daylight we'd come down and see what we got. But there was always one gift along with a lot of oh candies, oranges, *and so on like that*.

We will return to *and so on (like that)* below but now I turn to a replication of Tagliamonte and Denis (2010).

### Replication

Tagliamonte and Denis (2010:351–2) test the mechanism of phonetic reduction for the prototypical GEs (those that appear in the template in Table 3.1). For each general type, every individual speaker's frequency per 10 000 words of short and long versions, as exemplified in Table 3.8, was calculated. To determine the extent of phonetic reduction, the normalized frequency of long forms was subtracted from the normalized frequency of short forms. By binning speakers into three age groups, Tagliamonte and Denis (2010) argue that phonetic reduction in action can be identified if an increase in the mean of differences for each group across apparent time is

<sup>19</sup>Having become aware of this GE I have overheard it several times in Toronto, including as spoken by a 28 year old, male Toronto native (March 2014).

observed. A paired-sample *t*-test is used to determine if the mean of the differences in each group is different from zero.

Table 3.8: Prototypical GEs, short and long forms

Type	Short	Long
<i>stuff</i>	and stuff	and stuff like that, and that kind of stuff, and stuff of that sort etc.
<i>thing</i>	and things	and things like that, and that kind of thing, and things of that sort etc.
<i>everything</i>	and everything	and everything like that, and everything of that sort, and everything else etc.
<i>something</i>	or something	or something like that, or something of that sort etc.

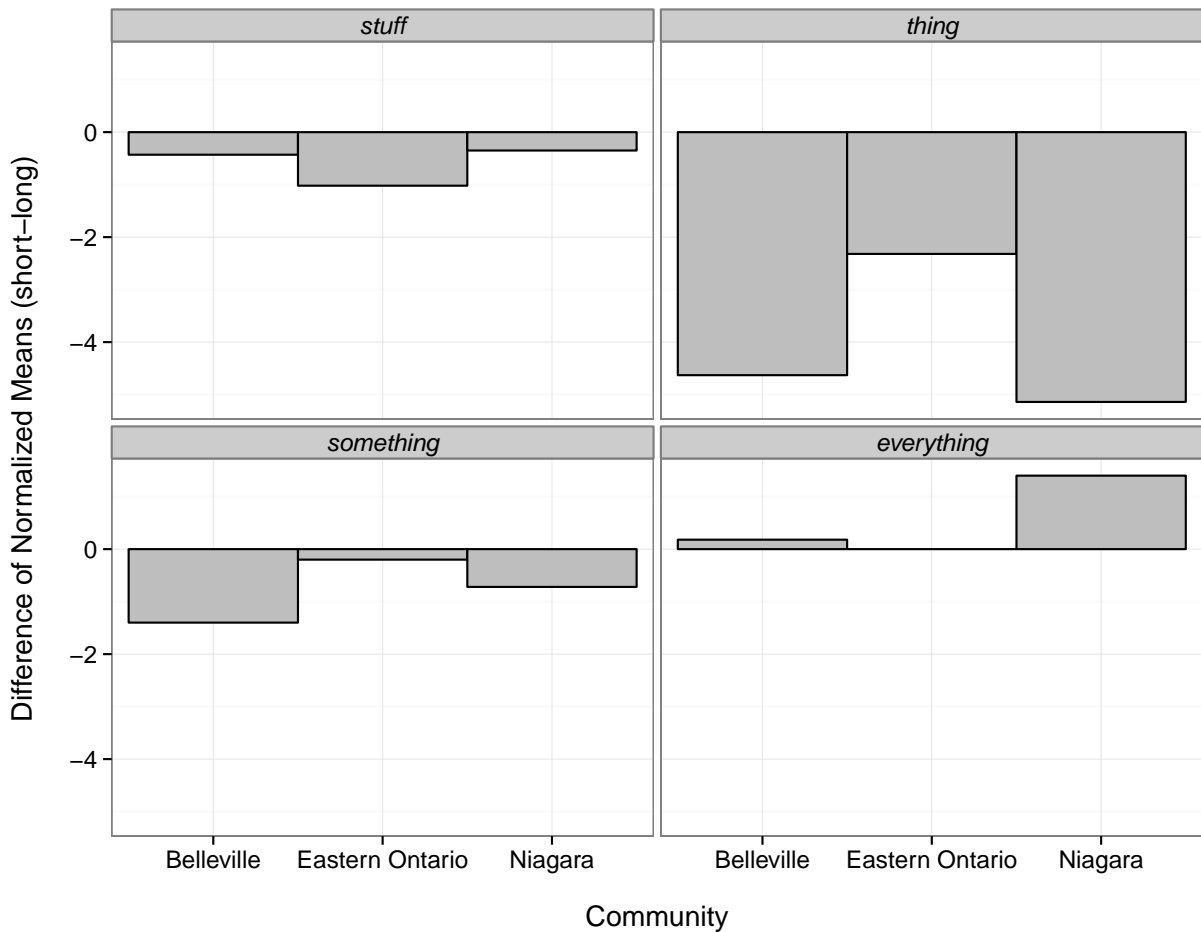


Figure 3.7: Mean individual differences of the normalized frequency of long and short *stuff*, *thing*, *something* and *everything* type GEs in Belleville, Eastern Ontario and Niagara

The first step in replicating this method is to determine if the three communities

in EOE can be treated as a single group. Figure 3.7 presents a bar chart of the mean individual differences of the normalized frequency of long and short forms of *stuff*, *thing*, *something*, and *everything* type GEs in Belleville, Eastern Ontario, and Niagara. For each individual the difference in normalized frequency of use between short and long forms was calculated and the mean in each community was plotted here.

For each community, the mean of the differences of each general GE type is in the same direction. That is, in all three communities there are more long *stuff* GEs than short *stuff* GEs. This is also the case for *thing* and *something* GEs. For *everything* GEs, there are more short forms than long forms, (though in Eastern Ontario the mean difference is essentially zero). Thus, for the remainder of this section, the three EOE communities will be treated as a single unit and act as the real-time benchmark against which to compare Tagliamonte and Denis' (2010) results from Toronto.

Table 3.9: Test of phonetic reduction in real time; Forms per 10 000 words, Paired-sample *t*-tests. Toronto data based on Tagliamonte and Denis (2010:Table 6). EOE,  $df = 35$ ; TOR >50,  $df = 31$ ; TOR 30-50,  $df = 13$ ; TOR <30,  $df = 38$ .

GE Form	EOE	TOR >50	TOR 30-50	TOR <30
<i>and stuff</i>	0.31	2.0	3.9	9.0
<i>and stuff like that</i>	0.91	2.4	2.4	5.7
	$t = -2.02, p = .05$	$t = -0.7, p = .52$	$t = 1.4, p = .17$	$t = 1.5, p = .13$
<i>and things</i>	1.27	0.4	1.9	0.2
<i>and things like that</i>	5.32	1.4	2.2	0.6
	$t = -3.08, p < .01$	$t = -2.1, p = .05$	$t = -0.1, p = .88$	$t = -1.7, p = .10$
<i>and everything</i>	1.00	1.9	0.8	3.6
<i>and everything like that</i>	0.44	0.5	0.0	0.4
	$t = 1.36, p = .18$	$t = 2.5, p = .02$	$t = 3.2, p = .01$	$t = 2.8, p = .01$
<i>or something</i>	2.49	3.0	4.4	6.7
<i>or something like that</i>	3.25	1.6	1.3	3.0
	$t = -1.76, p = .09$	$t = 1.6, p = .12$	$t = -2.5, p = .02$	$t = 2.8, p = .01$

Table 3.9 replicates the results from Tagliamonte and Denis (2010:Table 6) and additionally includes the results from EOE. The chart includes the mean normalized frequency of each GE type by length (e.g., *and stuff* vs. *and stuff like that*) and by age group. Results of a paired sample, two-tailed *t*-test, testing the significance of the

difference of the means of long and short forms in each cell are reported immediately under the means.

Tagliamonte and Denis (2010:351) find no significant difference in the means of long and short versions of *stuff* GEs. However, in EOE, the difference, such that there are more long forms than short forms, nears significance ( $p \approx .05$ ). This could be our first indication that an extended perspective on GEs in Ontario English provides empirical support for phonetic reduction as part of the grammaticalization of *stuff* type GEs in action. If *and stuff* underwent grammaticalization since its early usage in Ontario, this is exactly what we would expect. Likewise, there is evidence that *things* GEs also went through a process of (at least partial) phonetic reduction in Ontario English: the two younger age groups exhibit no significant difference in the frequency of long and short *things* forms, whereas the older two age groups exhibit a significant difference ( $p = .05$ ;  $p < .01$ ), such that there are more long forms than short forms. That said, one possible explanation for the lack of significance with the younger speakers is that *things* GEs obsolesced through the twentieth century, as evident above in Figure 3.5. In fact, of the thirty-nine speakers in the youngest age group, only nine used a *thing* type GE at all. With this caveat in mind, consider the *everything* GEs. Tagliamonte and Denis (2010:352) report that in all three age groups in Toronto, there are significantly more short *everything* GEs than long *everything* GEs. The extension of the analysis to EOE however provides some suggestion of phonetic reduction operating in the past. Although the normalized frequency of short *everything* GEs is higher than long *everything* GEs, the mean of the differences is not significant. We would expect that on a trajectory from more long to more short forms, there should be a period at which long and short forms are equally frequent. With respect to *everything*, the EOE data seem to represent that period. Lastly, as in Tagliamonte and Denis (2010:352), *something* GEs “offer the strongest evidence for phonetic reduction.” The difference of the normalized means of short and long *some-*



*thing* GEs incrementally increase across the four age groups, from  $-0.76$  to  $1.4$  to  $3.1$  to  $3.7$ . The differences are significant in the younger two age groups and not significant in the older two age groups. This is perhaps, like with *everything* GEs, because EOE and TOR $>50$  represent periods in which the long and short forms were equally likely. In sum, although Tagliamonte and Denis (2010:352) report “considerable stability” of this diagnostic of phonetic reduction, an extension of the data to an earlier point in apparent time reveals that, with the exception of *things* GEs, phonetic reduction may well have taken place in Ontario English (cf. Tagliamonte 2012:267)... Or did it?

### Independent Clipping

In the remainder of this sub-section I will argue that what has been interpreted as the phonetic reduction *cum* grammaticalization of GEs is the result of an independent change of morphological clipping. Before proceeding it is necessary to make explicit two opposing theoretical postulates. First, the grammaticalization mechanism of phonetic reduction is a gradual process that erodes individual linguistic forms as those individual forms proceed through the grammaticalization process.<sup>20</sup> If this is the case, phonetic reduction *cum* grammaticalization will affect different grammaticalizing forms at different rates because these developments are independent of one another. In the present case, the phonetic reduction of *stuff* GEs should operate independently of *things* GEs (and *something* and *everything* GEs) because it is individual linguistic forms that undergo grammaticalization, not variable systems.<sup>21</sup> The quantitative hypothesis then is that those GEs proceeding more quickly along the grammaticalization cline will exhibit a faster rate of phonetic reduction than those proceeding more slowly. The second, opposing postulate is that some diachronic changes

<sup>20</sup>This is uncontroversial within grammaticalization theory (with the possible exception of Himmelmann [2004]) but as Joseph (2004:47) argues “certain ways in which phonetic reduction is invoked in discussions of grammaticalization fly in the face of what is known about the regularity of sound change and the sorts of conditioning that can hold on sound changes.”

<sup>21</sup>This idea is implicit in Cheshire’s (2007:167) discussion of phonetic reduction—some variants began to reduce later than others.

that appear on the surface to be proceeding independently are actually the result of a single underlying rule involving variation between two competing variants in a synchronic grammar. These are the kinds of changes that traditionally involve competing grammars, differing minimally with respect to some abstract rule/parameter settings/functional head (Kroch 1989, 1994; Santorini 1992; Pintzuk 1999). The hypothesis here is that changes that proceed at a constant rate in all affected contexts are the result of a single underlying change. This constant rate pattern has been increasingly observed outside the realm of syntactic variation and competing grammars (see Fruehwald 2013 and Fruehwald, et al. 2013 for phonological change; Denis and Tagliamonte 2014a for English future temporal reference; Denis and Tagliamonte 2014b for pragmatic markers). If the process of change previously attributed to phonetic reduction in Table 3.9 is found to proceed at a constant rate, regardless of the context—in this case, regardless of the general GE type—the change is plausibly of this latter type. In other words, we can assume that the phonetic reduction of each GE observed is the result of a single underlying change, and thus, the loss of phonetic substance we observe in Table 3.9 is independent of grammaticalization. Joseph (2001:178) uses similar logic to argue against a grammaticalization theory account of the development of the weak nominative pronominal paradigm in Modern Greek (a purported “straightforward case of ordinary, garden-variety ‘grammaticalization’ via phonological reduction”). By showing that the development of these forms could have taken place via analogical change independent of grammaticalization, Joseph (2001:178) argues that “no sort of ‘grammaticalization’ as a process in and of itself is needed.”<sup>22</sup>

In an attempt to tease apart phonetic reduction *cum* grammaticalization of GEs from some more general underlying change that is independent of grammaticaliza-

---

<sup>22</sup>It would be worth-while to track the rise of these weak pronouns over time. If the whole paradigm develops at a constant rate (as opposed to individual pronouns phonetically reducing at different rates, perhaps correlated with frequency, as grammaticalization theory would predict), this would further support Joseph’s (2001) position.

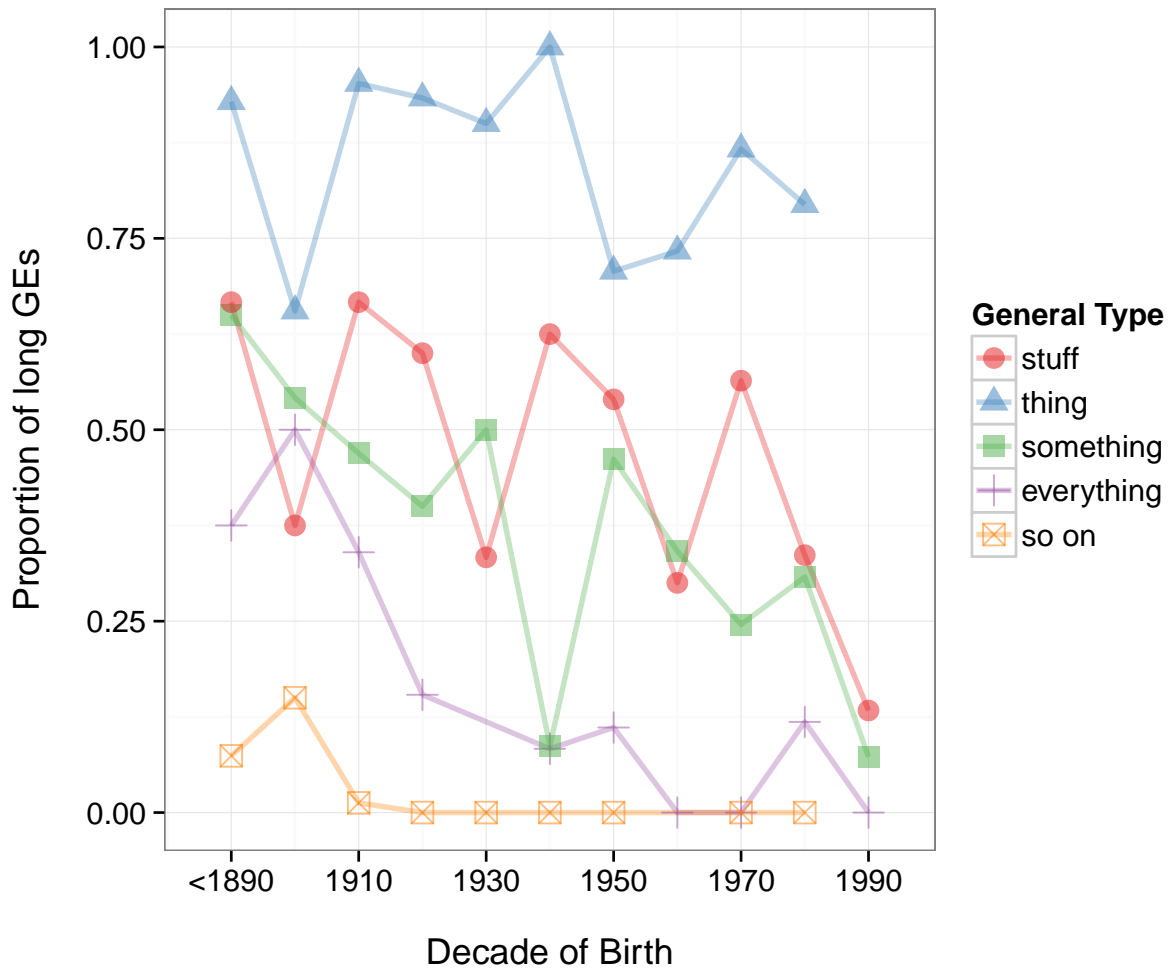


Figure 3.8: Proportion of long *stuff*, *thing*, *something*, *everything*, and *so on* GEs (vs. short forms) through apparent time. N = 2044.

tion, we begin by examining GEs from a perspective that has not previously been presented in the literature. Figure 3.8, like Figures 3.5 and 3.6, collapses across community and corpus, binning speakers into decade long age groups along the x-axis. For each age group a data point is plotted for five GE types: *stuff*, *thing*, *something*, *everything*, and *so on*.<sup>23</sup> The position of the data point along the y-axis represents the frequency of long forms as a proportion of the total number of each GE type in each age group. For example, for speakers born in the first decade of the twentieth century, approximately 50% of all *everything* GEs were long forms and 50% were short.

<sup>23</sup>These are the general types that can be conceived of as having long and short specific realizations.

Despite some jaggedness to the lines (likely due to low Ns in some cells), the trend is a decrease in the frequency of long forms. Comparing the oldest speakers to the youngest speakers, the GE types *stuff*, *something* and *everything* have all dramatically reduced the frequency of long forms. The GE *and so on like that*, which was briefly discussed above completely fell out of favour in the early half of the twentieth century. For speakers born after 1930, only the long *thing* GEs remain consistently above the 50% mark.

So far, the evidence does not disentangle the two possible causes of phonetic reduction. We simply see that for all GEs, there is a trend toward a shorter variant. In order to put grammaticalization to the test, we need a benchmark against which to compare. We might hypothesize that an alternative explanation to the trend toward shorter GEs is a different type of change, namely the deletion of the comparative element. Perhaps the reason that younger people use fewer long variants is because of a change in progress such that the comparative element *like that* is increasingly susceptible to elision as in (16).

(16) *like that* → ∅

In other words, perhaps the presence of long and short GEs represents variation between the comparative *like that* and a null comparative (cf. Tagliamonte and Smith 2005 *inter alia* on complementizer deletion). If this is the case, this process of change should effect all prototypical GEs in the same way (and thus at the same rate), since this deletion is independent of any particular GE. Some contexts (i.e., some GE types) might favour or disfavour comparative deletion to different extents, but crucially, the rate of change should be the same.<sup>24</sup>

Following Kroch (1989) (among others), this hypothesis can be tested using a logistic regression model. The dependent variable of this model will be the realization

---

<sup>24</sup>See §1.1.3 above.

of the comparative, specifically *like that* ~  $\emptyset$ , as in (17).<sup>25</sup>

(17) a. Well like, you know, help feed the cows *and stuff like that* when you're young enough.

(EON/M/1898)

b. Sowing seeds *and stuff*.

(NIA/M/1902)

The model tests the main effects of YEAR OF BIRTH of the speaker and GE TYPE. YEAR OF BIRTH is a continuous factor and has been centered around the mean in the model. The GE TYPE factor is categorical (zero-sum coded) and includes as levels, the prototypical GEs *stuff*, *thing*, *everything*, and *something*, as well as *so on*, which also exhibits variable presence of the *like that* complementizer.<sup>26</sup> If the YEAR OF BIRTH term is significant, this will be taken as evidence that there is a change in progress. If GE TYPE is significant this will indicate that different contexts favour or disfavour comparative deletion to different extents. To determine the nature of phonetic reduction *cum* grammaticalization versus an independent change, the model includes an interaction terms for YEAR OF BIRTH×GE TYPE. If this interaction is significant, this can be taken as an indication that different GE types are undergoing phonetic reduction at different rates, as predicted by grammaticalization theory. However, if the interactions are not significant while the main effect of YEAR OF BIRTH is significant, this can cautiously be interpreted as evidence that a single rule of comparative deletion is increasing in probability, affecting all GEs at a constant rate.<sup>27</sup>

<sup>25</sup>All other long GEs are excluded from the data set in order to concentrate on the hypothesis that the trend in Figure 3.8 is primarily due to the deletion of *like that*.

<sup>26</sup>GEs with the generic *anything* are excluded as the quantifier *any* is only licensed in particular environments.

<sup>27</sup>This interpretation is necessarily cautious because a constant rate is the null hypothesis. As Paolillo (2011) argues it is difficult to statistically validate "independence claims" such as this. But see Fruehwald et al. (2013) for a discussion of the difficulty usage based models would have in accounting for changes that proceed at a constant rate, independent of context.

Table 3.10: Mixed-effects logistic regression testing the fixed effects of YEAR OF BIRTH (centered, continuous), GE TYPE (*stuff*, *everything something*), and their interaction and a random intercept for Speaker on the realization of the *like that* comparative. Sum contrast coding. Coefficients reported in log-odds. Correlation of fixed effects,  $r < |0.93|$ .<sup>28</sup> N = 1777.

	Estimate	Std. Error	z-value	p-value	
Intercept	-2.446	0.722	-3.390	$7.00 \times 10^{-04}$	***
YEAR OF BIRTH (centered)	-0.038	0.014	-2.603	$9.24 \times 10^{-03}$	**
GE TYPE ( <i>stuff</i> )	2.136	0.717	2.980	$2.88 \times 10^{-03}$	**
GE TYPE ( <i>thing</i> )	3.684	0.732	5.035	$4.77 \times 10^{-07}$	***
GE TYPE ( <i>so on</i> )	-6.802	2.792	-2.436	$1.48 \times 10^{-02}$	*
GE TYPE ( <i>everything</i> )	-0.503	0.740	-0.679	$4.97 \times 10^{-01}$	
YOB:TYPE ( <i>stuff</i> )	-0.011	0.015	0.753	$4.51 \times 10^{-01}$	
YOB:TYPE ( <i>thing</i> )	0.040	0.015	2.625	$8.65 \times 10^{-03}$	**
YOB:TYPE ( <i>so on</i> )	-0.080	0.055	-1.448	$1.48 \times 10^{-01}$	
YOB:TYPE ( <i>everything</i> )	0.007	0.015	0.462	$6.44 \times 10^{-01}$	
Random intercept:					
SPEAKER	Variance = 2.21, N = 117				

The results of this model are presented in Table 3.10. There are three points to observe. First, the main effect for YEAR OF BIRTH is significant and the coefficient is negative, suggesting that younger speakers are increasingly less likely to use the *like that* comparative. This makes sense given the general downward trajectory observed in Figure 3.8. Second, the significance of a number of levels of the GE TYPE factor and their corresponding coefficients suggests that different GEs favour or disfavour the *like that* comparative at varying strengths. The significant positive coefficients for *stuff* and *thing* indicate that these GEs generally favour *like that*. Though not shown because it is the reference level of this model, *something* also has a significant positive coefficient.<sup>29</sup> The significant negative coefficient for *so on* indicates that this GE disfavors *like that*. The coefficient for *everything* is not significant. Again, these

<sup>28</sup>The high collinearity in this model is expected since the type of GE is strongly correlated with age (see Figure 3.5).

<sup>29</sup>This variant was chosen as the reference level because it seemed to fall in the middle of the other variants in Figure 3.8. Furthermore, this model uses sum contrast coding, thus the coefficients of categorical factors add to 0. Thus, the coefficient for the reference level can be calculated by taking the negative value of the sum of the coefficients of all other levels of that factor.

patterns are unsurprising considering the pattern in Figure 3.8: *so on* hardly co-occurs with *like that, everything* is in the middle, *something* and *stuff* hover around the middle, and *things* is the GE type that most frequently occurs as a long variant. Lastly, and most importantly, consider the interaction terms. Although the slope of change (i.e., YEAR OF BIRTH) for *stuff, so on, everything* are the same as the reference level *something* indicated by the non-significant interaction term, the slope for *thing* stands apart, with a significant interaction term.<sup>30</sup> To better understand this interaction, the fitted values from this model are plotted in Figure 3.9. Each data point represents an individual speaker's probability of using a long form for a particular variant. Binomial curves are fit to this data in order to help visualize the effects.

We can see that the slope for *things* is much less steep than the others, which all look essentially the same. This suggests that the change is proceeding slower for this one general GE type. Thus far, our statistical model partially support the hypothesis of phonetic reduction *cum* grammaticalization: the change is proceeding at a different rate in different contexts; phonetic reduction is gradually eroding away GEs as they gradually grammaticalize at different rates.

However, of the five general types tested, only one (*things*) deviates from the rest. If any GE type was going to behave differently, we might expect it to be *things* since this variant is rapidly moving toward obsolescence as shown in Figure 3.5. If we construct a similar model, excluding *things* GEs, the YEAR OF BIRTH:GE TYPE interaction is no longer significant.<sup>31</sup>

What remains then is a mixed story. On the one hand, we have possible evidence that supports the idea that individual GE types are phonetically reducing. Thus, we have evidence that different GEs are eroding at different rates. On the other hand, only one GE stands out from all the others as progressing at a different rate. The

---

<sup>30</sup>An analysis of deviance confirms that this interaction term significantly contributes to the model ( $\chi^2 = 13.8, df = 4, p = 0.008$ ).

<sup>31</sup>According to an analysis of deviance the interaction term does not significantly contribute to the model ( $\chi^2 = 4.8, df = 3, p = 0.186$ ).

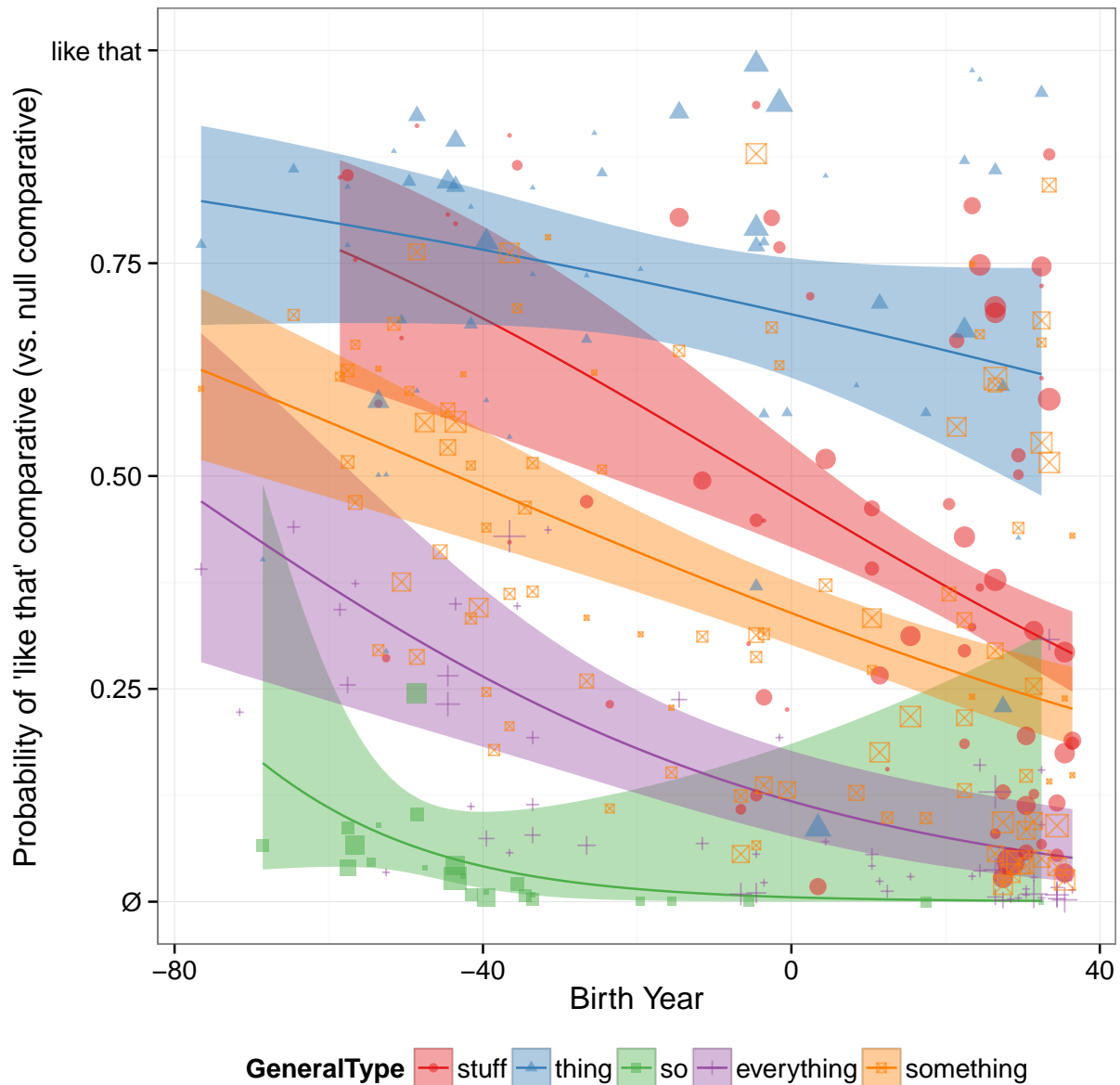


Figure 3.9: Probability of *like that* comparative (vs.  $\emptyset$ ) by GE type through apparent time. Data points represented the fitted values from the regression model in Table 3.10. Size of represents number of data points at that coordinate.  $N = 1777$ .

rates of change of all other GE types are statistically indistinguishable. Perhaps the trajectory of phonetic change for *things* is different from all other forms because of its obsolescence. For now we will move on to the other diagnostics of grammaticalization tested by Cheshire (2007) and Tagliamonte and Denis (2010) and return to a discussion of the mechanism of phonetic reduction in the context of a broader



perspective on grammaticalization below.

### 3.4.2 Decategorialization

#### Background

I now turn my attention to the second mechanism of grammaticalization: DECATEGORICALIZATION or “the loss of morphosyntactic properties characteristic of the source form[...]” (Heine and Kuteva 2005:579). As discussed above, Cheshire (2007:168) begins with the hypothesis that “[i]n a [GE] that has not grammaticalised, we might expect the head noun in the construction to always have the same syntactic and semantic properties as a preceding noun to which it relates anaphorically.” That is, *and stuff* should co-occur with non-count nouns, since the lexical noun *stuff* is a non-count noun itself. The features assumed for different kinds of generics of GEs are shown in (6) above. However, since Dines (1980), it has been observed that GEs do not strictly follow this feature matching requirement. Not only do *and stuff* and *and things* (etc.) co-occur with unexpected nominal referents, GEs frequently and unambiguously co-occur with phrases larger than nominals. GEs that co-occur with such unexpected referents are assumed to be more grammaticalized than GEs that co-occur with their expected referents.

The EOE data contains many GEs that co-occur with unexpected referents. Consider the examples in (18).

- (18) a. Stroud’s would have toys and [kitchenware *and all that stuff*] at that time.  
(BLV/F/1903)
- b. One Christmas I was tired of [Santa Claus pictures *and all that stuff*].  
(BLV/F/1898)
- c. I still have [...] the old agreements dating back to what the hired man

had to to sign. To provide so much wood and [provide so much milk  
and all this kind of stuff]. (EON/M/1912)

- d. I was one out of a group of five hundred junior farmers that went to the Royal Winter Fair from Ontario. [...] We spent a week as the guests of the Ontario Government and [we stayed at the Royal York hotel and all this kind of stuff]. (EON/M/1912)

The examples in (18) are all of *stuff* type GEs in EOE. The specific realization of the GE is nearly identical in every case: *and all that/this (kind of) stuff*. However, in each case, the GE unambiguously co-occurs with referents of different syntactic categories. In (18a), *and all that stuff* attaches to *kitchenware*, a non-count noun. In (18b), the exact same GE, *and all that stuff*, co-occurs with *Santa Claus pictures*, a plural (count) noun. In (18c), *and all this kind of stuff* caps off the end of a list of conjoined predicates (*vPs*). Lastly, in (18d), the same GE caps off a list of whole sentences (*CPs*).

We find the same distribution of referents for other GE types. In (19), *thing* type GEs are shown to attach to the same range of referents as *stuff* above.

- (19) a. We had [some sash and things] there ready for people to call for.  
(BLV/F/1898)
- b. My sister looked after that quite a lot... the picking of it, uh the currents and [berries and things], if we were busy in hay.  
(NIA/M/1907)
- c. They used to be in those big groups go down the road. Go in the houses and [have their food and things].  
(BLV/F/1903)
- d. One of the Sunday School teachers or the minister would take topic with us and we'd have a devotional period and then we'd have the social

period and we'd have lunch and so on. [We'd play games *and things like that*].

(NIA/F/1912)

In many cases the referent of a GE is ambiguous. In some instances there is even a three-way ambiguity between a nominal referent, a verbal referent, and a sentential referent, as in (20).

- (20)
- a. The old Methodist church was struck by lightning and after it took, they might have had [some political meetings *or something like that*].
  - b. The old Methodist church was struck by lightning and after it took, they might have [had some political meetings *or something like that*].
  - c. The old Methodist church was struck by lightning and after it took, [they might have had some political meetings *or something like that*].

(BLV/F/1897)

Following the previous literature, examples such as this are left aside for the remainder of this sub-section as there is no straightforward way of determining the speaker's intended structure.<sup>32</sup>

### Replication

Following Cheshire (2007) and Tagliamonte and Denis (2010), each GE in EOE was coded for the syntactic category of its referent. Then, each unambiguous token was coded for whether that type of referent was expected or not, given the GE. For example, non-count nouns were coded as 'expected' for *stuff* GEs but 'unexpected' for *things* GEs. Likewise, plural count nouns were coded as 'expected' for *things*

<sup>32</sup>There were fifteen ambiguous tokens in Belleville, twenty-six in Eastern Ontario, and twenty-six in Niagara, and four, nine, and three tokens in each respective community that were unclear for other reasons (e.g., background noise, overlapping speech, etc.).

and ‘unexpected’ for *stuff*. Furthermore, following Cheshire (2007) a distinction was made between unexpected nominal referents and unexpected non-nominal referents. By examining the distribution of expected nominal referents, unexpected nominal referents, and non-nominal referents over apparent-time, the trajectory of decategorialization can be assessed. Tagliamonte and Denis (2010:352–354) examine the distribution of referents of short *stuff* GEs, long *stuff* GEs, short *thing* GEs, and long *thing* GEs in the same three age groups as in Table 3.9.<sup>33</sup> The trend is stability. For both *stuff* GEs and long *thing* GEs there is no change in apparent-time. There is some hint of change with short *thing* GEs but even among the oldest speakers nearly half of all tokens of *and things* co-occur with a unexpected referent.

The limitation of Tagliamonte and Denis’ (2010) findings is the lack of real-time evidence. Although there was no evidence of decategorialization in progress, this does not preclude the possibility that decategorialization took place earlier in Ontario English. Given this possibility, I replicate Tagliamonte and Denis’ (2010) approach, adding an expanded temporal perspective from EOE in Figure 3.10.<sup>34</sup> Figure 3.10, a stacked bar plot displays the proportion of tokens across the same four apparent time age groups as used in Table 3.9 for each of the three referent categories (expected nominal in red, unexpected nominal in blue, unexpected other phrase in green). Each of the four facets of the chart are divided by general GE type and length.

Just as Tagliamonte and Denis (2010:354) observed, the general trend in Figure 3.10 is one of stability. Now consider each facet individually. First, the upper left facet shows that the majority of early onset tokens of short *stuff* in EOE co-occur with unexpected referents. This trend continues into the twentieth century. Although there is some indication that these oldest speakers had more of a preference for expected referents than the younger speakers (in TEA), there are only six tokens of short *stuff* in

<sup>33</sup>Low frequency disjunctive *stuff* and *thing* forms have been excluded here.

<sup>34</sup>The charts here differ from Tagliamonte and Denis (2010) in completely leaving out ambiguous tokens. In Tagliamonte and Denis (2010), ambiguous tokens were included in the ‘Other Phrase’ category. Also, in replicating this data some minor errors in the previous data file were corrected.

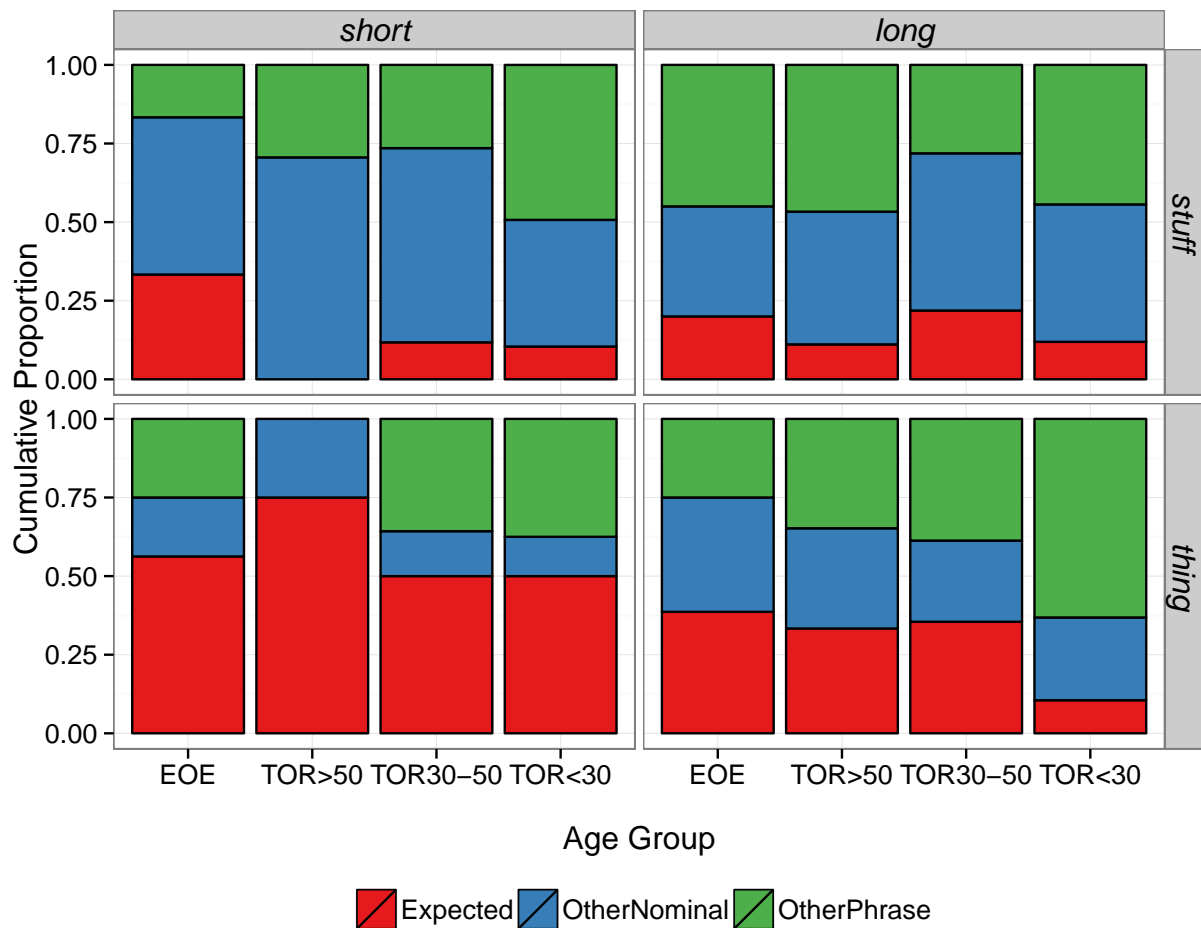


Figure 3.10: Stacked bar chart showing the changing distribution of GE referents over apparent time. *Stuff* N = 26 (EOE); 62 TOR>50; 66 TOR 30-50; 362 TOR<30. *Thing* N = 104 (EOE); 77 TOR>50; 45 TOR 30-50; 46 TOR<30.

EOE. The distribution of referents across apparent time for long *stuff* GEs in the upper right facet is stable. Tagliamonte and Denis (2010:354) suggested that short *thing* GEs were the only GEs where there was “a marked difference between the speakers older than fifty and the younger age groups” in Toronto. Indeed, for the TOR>50 group, there were no tokens of *and things* co-occurring with a non-nominal referent. However, we can now see in the bottom left facet that this difference is not the result of decategorialization in progress, but rather because the EOE group and the two younger age groups in TEA exhibit essentially the same distribution of referents, the TOR>50 group seems to be a quantitative anomaly. Lastly, in the bottom right are

the long *thing* tokens. Although the oldest three groups have near-identical distributions of referents, the youngest speakers have a lower rate of expected referents and this seems to be at the expense of a higher number of non-nominal referents. Thus, with the expanded view from EOE, the conclusion reached by Tagliamonte and Denis (2010) for TEA holds. By and large there is stability. The only exception is the possible shift of long *things* forms to use with non-nominal referents, as in (21), as the form obsolesces.

- (21) It was focused on the family, and how it works, and how it matures *and things like that*.

(TOR/F/1983)

### The Problem with the ‘Sideways’ Perspectives

Before continuing, recall from Chapter 2 that examining variation ‘sideways’ as has been done in Figure 3.10 is potentially misleading. Thus, I turn to an examination of the decategorialization of GEs from an ‘accountable’ perspective. Figure 3.11 shows a series of stacked bar plots. The four age groups are listed along the x-axis as in Figure 3.10. Along the y-axis is the proportion of all adjunctive GEs, partitioned by the general type of GE (colour coded). There is one facet for each referent type: *non-count nominal*, *plural nominal*, *singular nominal*, *CP*, *gerundal phrase*, and *VP*.

The major finding here is that *stuff* type GEs (in red) rise in frequency, regardless of context. For every referent, *stuff* GEs are least frequent in the EOE group but monotonically rise in frequency such that for the youngest age group, *stuff* represents well over half of all GEs in all contexts. But how do we assess decategorialization of *stuff*? There are two potential indicators. First, if decategorialization is active we would expect to find that amongst the oldest speakers, there are referents that never co-occur with *stuff* GEs but with later generations *stuff* GEs begin to co-occur with

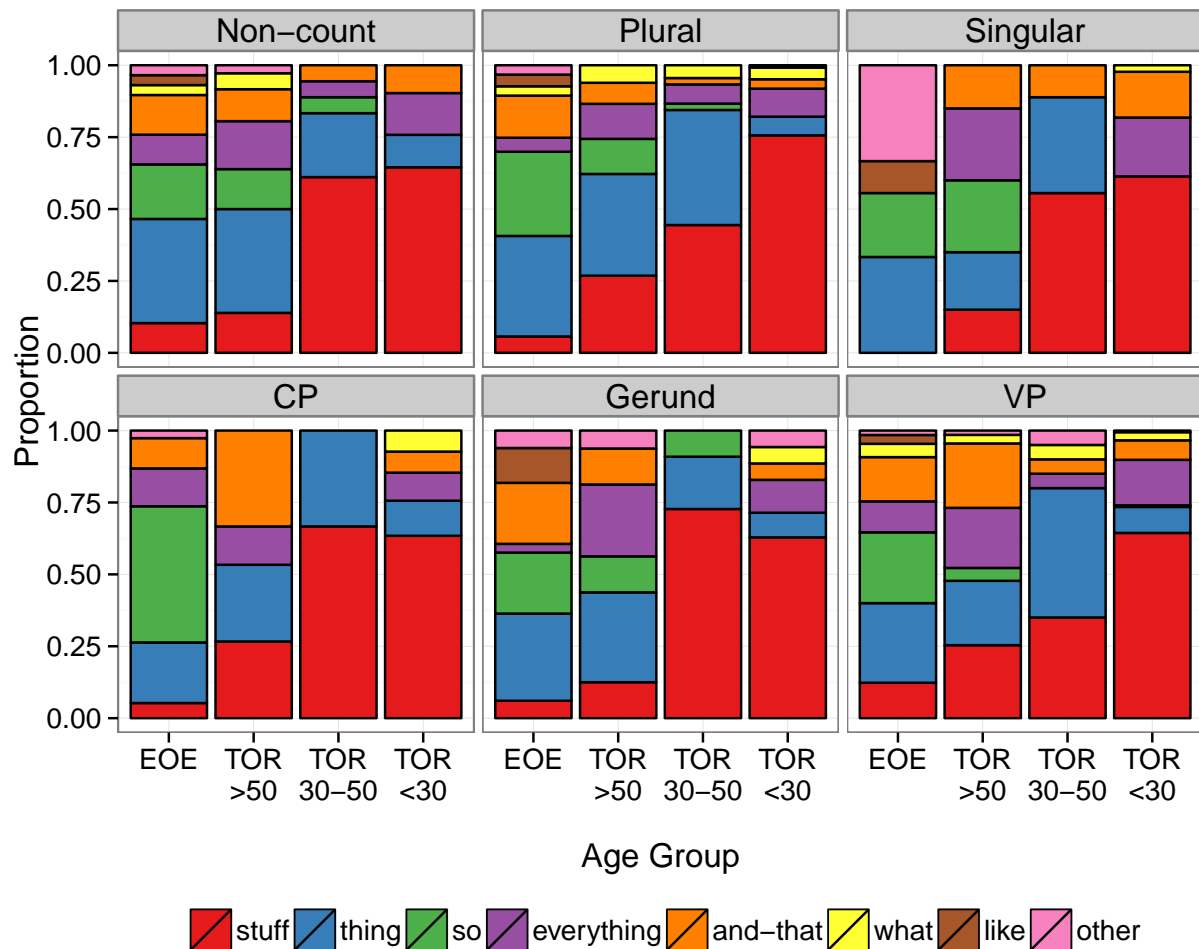


Figure 3.11: Stacked bar chart showing the changing distribution of GE variants over apparent time by referent. Non-count  $N = 174$ ; Plural  $N = 373$ ; Singular  $N = 82$ ; CP  $N = 100$ ; Gerund  $N = 95$ ; VP  $N = 329$ .

that referent. This might be the case for singular referents. There are no tokens of *stuff* that co-occur with singular nominals in the EOE data, but in the next generation there are. However, this context is also the least frequent over all. There are only nine GEs in the EOE which unambiguously co-occur with singular nominals, so we must be cautious not to overinterpret what could be an accidental gap in the data. A second potential way to identify decategorialization is to observe differing rates of advancement of *stuff* with different referents. Since *stuff* is well established as the majority variant among the youngest speakers regardless of referent, if *stuff* spread to different contexts at different times (i.e., if *stuff* gradually decategorized),

the rates of change of *stuff* with different referents would necessarily be different. However, if the rise of *stuff* GEs occurred at a constant rate across referents, it is plausible that decategorialization never occurred. That is, *stuff* occurs frequently in all contexts today not because of decategorialization but because it always occurred in all contexts and has generally and independently increased in frequency. These two different possibilities are stylized in Figure 3.12 with each line type representing the trajectory of change of four different referent types.

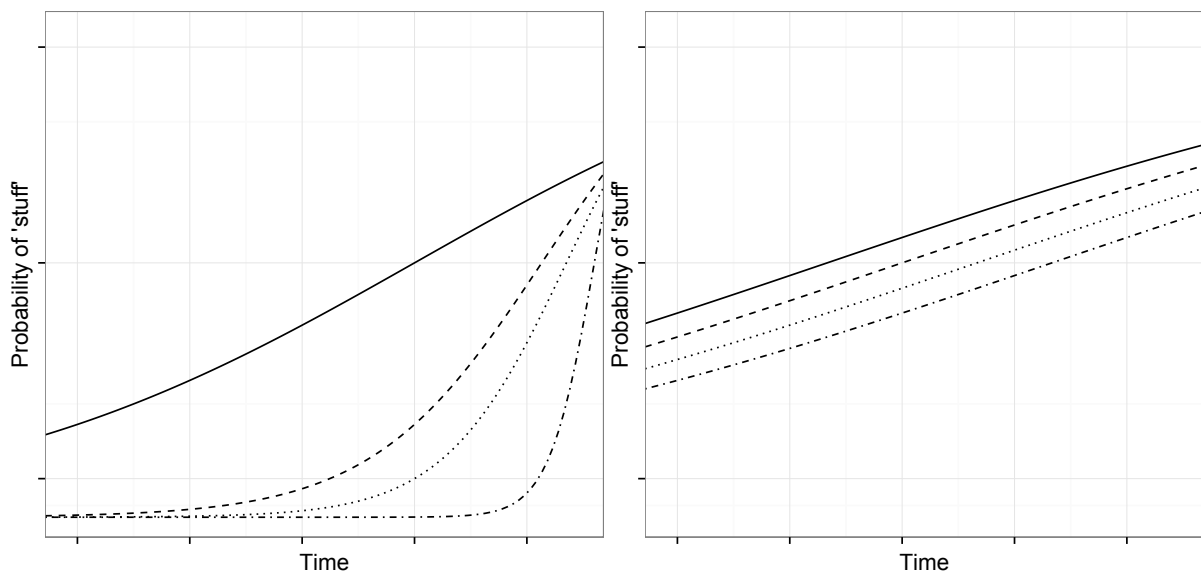


Figure 3.12: Stylized models of change. Left: Gradual decategorialization, different rates of change. Right: No decategorialization, constant rates of change.

We can test this diagnostic of decategorialization with a logistic regression model. Table 3.11 presents the results of a mixed-effects logistic regression that models the main effects of YEAR OF BIRTH, REFERENT TYPE (treatment contrast coding, non-count nominals—the expected referent for *stuff* GEs—as reference level) and their interaction, along with a random intercept for speaker, on the probability of *stuff* GEs across the entire apparent time range of the GE data. Table 3.12 presents the result of an analysis of deviance of the model in Table 3.11. If the interaction between YEAR OF BIRTH and REFERENT TYPE is significant, this could indicate decategorialization. If not, we have no evidence for decategorialization.



Table 3.11: Mixed-effects logistic regression testing the fixed effects of YEAR OF BIRTH (centered, continuous), referent type (reference level = non-count referent), and their interaction and a random intercept for Speaker on the realization of adjunctive GEs as *stuff* type GEs. Treatment contrast coding. Coefficients reported in log-odds. Correlations of fixed effects,  $r < |0.2|$ . N = 980

	Estimate	Std. Error	z-value	p-value	
Intercept	-0.355	0.192	-1.846	$6.48 \times 10^{-02}$	
YEAR OF BIRTH (centered)	0.050	0.006	8.079	$6.51 \times 10^{-16}$	***
REFERENT ( <i>plural</i> )	-0.174	0.208	-0.835	$4.04 \times 10^{-01}$	
REFERENT ( <i>sing.</i> )	0.123	0.180	0.681	$4.96 \times 10^{-01}$	
REFERENT ( <i>gerund</i> )	-0.141	0.342	-0.414	$6.79 \times 10^{-01}$	
REFERENT (VP)	0.261	0.306	0.853	$3.94 \times 10^{-01}$	
REFERENT (CP)	0.042	0.179	0.236	$8.14 \times 10^{-01}$	
YOB:REFERENT ( <i>plural</i> )	-0.016	0.007	-2.277	$2.28 \times 10^{-02}$	*
YOB:REFERENT ( <i>sing.</i> )	0.005	0.006	0.784	$4.33 \times 10^{-01}$	
YOB:REFERENT ( <i>gerund</i> )	0.003	0.013	0.248	$8.04 \times 10^{-01}$	
YOB:REFERENT (VP)	0.011	0.011	1.032	$3.02 \times 10^{-01}$	
YOB:REFERENT (CP)	-0.005	0.006	-0.767	$4.43 \times 10^{-01}$	
Random intercept:					
SPEAKER			Variance = 1.87, N = 110		

Table 3.12: Analysis of deviance,  $\chi^2$  test for model reported in Table 3.11.

	$\chi^2$	df	p-value	
Intercept	3.409	1	$6.49 \times 10^{-02}$	
YEAR OF BIRTH	65.276	1	$6.51 \times 10^{-16}$	***
REFERENT	2.042	5	$8.43 \times 10^{-01}$	
YOB:REFERENT	7.086	5	$2.14 \times 10^{-01}$	

The model in Table 3.11 shows a significant main effect of YEAR OF BIRTH, confirming the rise of *stuff* GEs over time. There is no main effect of REFERENT, suggesting the non-count nominal referents do not significantly favour (or disfavour) *stuff* more (or less) than any other referent. The significance and non-significance of these main effects is confirmed in an Analysis of Deviance in Table 3.12. Lastly, although one of the interaction terms reaches significance (plural referents), the Analysis of Deviance finds that the interaction does not add explanatory value to the model. The model is visualized in Figure 3.13 which plots the fitted values from the model and quasibinomial curves are fit to the data. Comparing this model to the two potential

outcomes in Figure 3.12, we see that the situation for *stuff* in Ontario is consistent with a constant rate of change and not consistent with decategorialization.

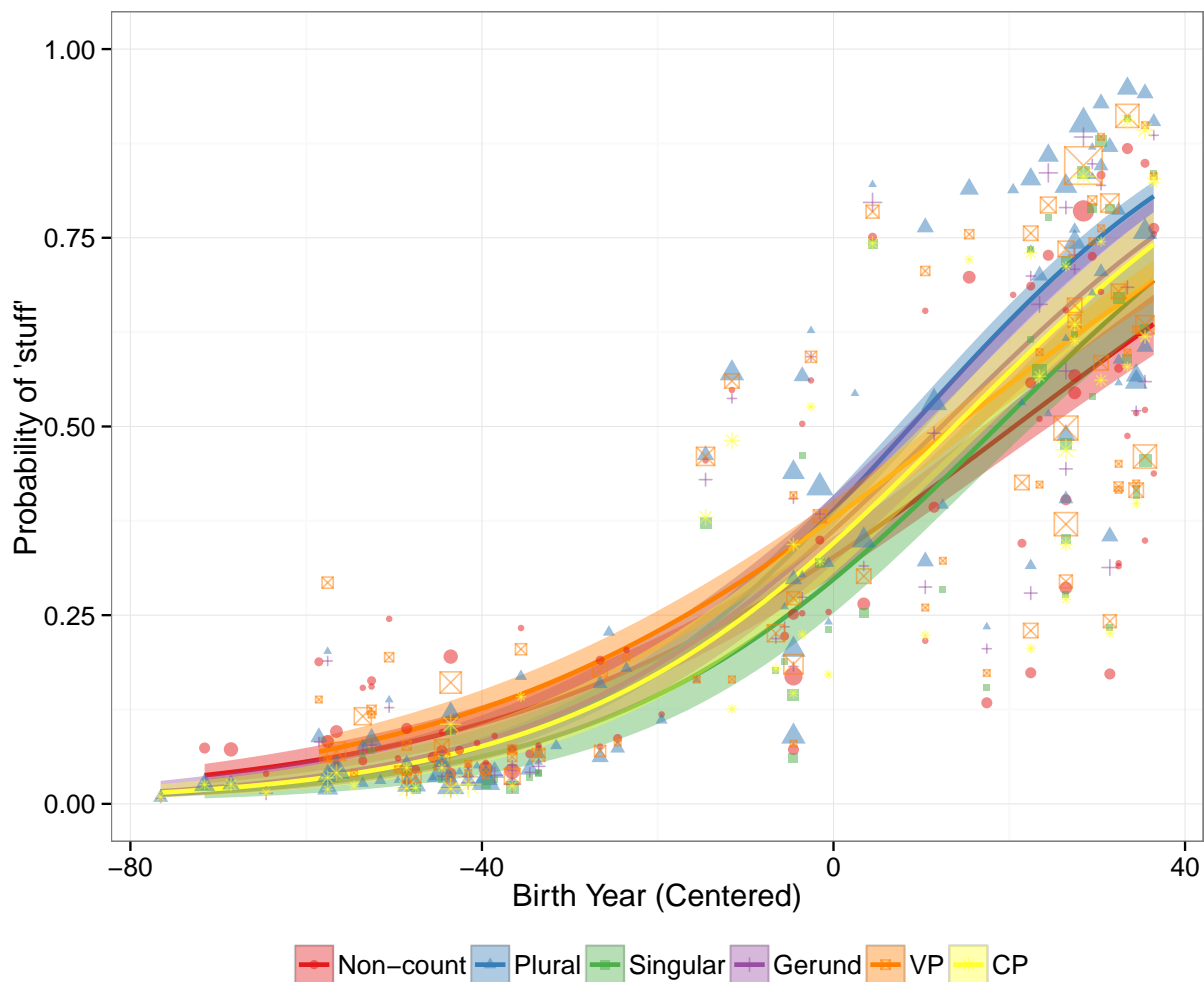


Figure 3.13: Probability of *stuff* GE (vs. all other adjunctive GEs) by referents over apparent time. Fitted values from the model in Table 3.11. Dot size represents the number of tokens at those coordinates.  $N = 980$ .

### Decategorialization? Or, Toward a Formal Semantics of GEs

In the last part of this section I will propose that the lack of decategorialization of GEs is at least in part, not theoretically unexpected. Despite the general assumption in the GE literature that GEs should coordinate with like nouns—and like nouns only—English generally allows for coordination of different types of nominals. Consider

(22), examples of coordinated nominals, one non-count, the other count.

- (22) a. bangers and mash  
 b. sash and doors  
 c. meat and potatoes  
 d. milk and cookies  
 e. soup and crackers

Though the examples above might be considered fixed expressions, there are examples in the EOE of GEs capping the end of a list that contains both count (e.g., *dishes*) and non-count nouns (e.g., *furniture*), as in (23).

- (23) My grandfather gave me all the furniture and- all the antique furniture and dishes *and so on*.

(NIA/F/1912)

It is unsurprising that 'plural, count' *and things* co-occurs with non-count nominal referents and 'non-count' *and stuff* co-occurs with count nominals and thus, inaccurate to assume that for a GE to co-occur with an unexpected nominal referent that this is an indication of decategorialization.

In a type-driven, compositional semantics, such as the model discussed in Heim and Kratzer (1998), coordination is assumed to combine elements of the same type. Furthermore, it is commonly assumed that not only non-count and count nominals, but also verb phrases, are predicates of the same type  $\langle\langle e, t \rangle\rangle$ . If we assume that GEs are of type  $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$  (that is they take an argument of type  $\langle e, t \rangle$  and return an object of type  $\langle e, t \rangle$ ), then GEs combine with their  $\langle e, t \rangle$  type referents via the simple

compositional rule of Functional Application.<sup>35</sup> Thus, no compositional acrobatics are required to coordinate different types of GEs with nominals and verbs.<sup>36</sup> From a compositional perspective, GEs with verbal referents are not decategorized versions of GEs that modify nominal referents. Rather, the same GE with the same compositional properties is at work.<sup>37</sup>

The story becomes more complicated however because, as observed above in (20), GEs also function to extend CPs (i.e., whole propositions) and DPs (i.e., nominals with a determiner). From a compositional semantic perspective then, there must also be homophonous GEs which combine with elements of type  $t$  (propositions) and elements of type  $e$  (individuals). Decategorialization may very well have taken place then. That said, there is no evidence for context expansion, from nominals and verb phrases to DPs and CPs in the Ontario data presented in Figure 3.13

However, it may also be the case that a process of decategorialization took place that was completely independent of the grammaticalization of GEs specifically. It has long been observed that the generic nominal *things* (along with *stuff*, *shit*, *junk*,

---

<sup>35</sup>Functional Application is a composition rule that takes an argument of a particular type and returns an object of a particular type, possibly the same type. More formally Functional Application is defined by Heim and Kratzer (1998:44) as:

If  $\alpha$  is a branching node and  $\{\beta, \gamma\}$  is the set of  $\alpha$ 's daughters, and  $\llbracket \beta \rrbracket$  is a function whose domain contains  $\llbracket \gamma \rrbracket$ , then  $\llbracket \alpha \rrbracket = \llbracket \beta \rrbracket(\llbracket \gamma \rrbracket)$ .

In this case (for example),  $\llbracket \text{and stuff like that} \rrbracket$  is a function (of type  $\langle et, et \rangle$ ) whose domain contains arguments of type  $\langle et \rangle$  (such as  $\llbracket \text{eggs} \rrbracket$ ,  $\llbracket \text{butter} \rrbracket$ ,  $\llbracket \text{eggs and butter} \rrbracket$ , and  $\llbracket \text{run} \rrbracket$ ).

<sup>36</sup>Note that for nominals the GE must be coordinated with the NP (or NPs) within the DP headed by *his* and not coordinated with the DP itself. Consider a sentence such as (24b), which seems to be infelicitous.

- (24) a. Every Saturday, John sells his butter and eggs and stuff like that at the market.  
 b. #For example, last Saturday John was selling his neighbours onions there.

The semantic awkwardness of (24b) arises because the set to which *and stuff like that* is generalizing in (24a) is restricted by the determiner *his*. The general set of stuff that John sells must be his own. Therefore, (24b) is infelicitous, since John is selling his neighbours onions and not his own.

<sup>37</sup>A fully fleshed out semantics for GEs is beyond the scope of this thesis but an approach using Hamblin style semantics which evokes a set of alternatives, similar to what has been used for the semantics of questions (Hamblin 1973; Ramchand 1997; Kratzer and Shimoyama 2002; Zabbal 2004), Focus (Rooth 1992), and disjunction (Alonso-Ovalle 2006) would be a reasonable approach. For now, the compositional mechanics are the main focus.

*nonsense* etc.) is highly semantically bleached and “can be readily used to apply to living beings, objects of any shape and description, liquids, abstract concepts, events, etc.” (Fronek 1982:637; see also O’Keeffe 2004:12, Quirk et al. 1985:76–77).<sup>38</sup> Thus, a situation in which the GE *and stuff* entered Ontario English already bleached of its non-count noun meaning (and thus already able to take on any of the three combinatorial possibilities) is possible. Such a situation is also consistent with the results presented in Table 3.11 and is what I have previously argued is what happened in York, UK (Denis 2010).

### 3.4.3 Semantic Bleaching

#### Background

In this section, I will examine the third mechanism of grammaticalization, semantic bleaching. For many grammaticalization theorists, semantic bleaching is at the core of grammaticalization. For example, Heine (2003:583) argues that bleaching “precedes and is immediately responsible for” the other three mechanisms of change. Thus, for grammaticalization to be in action with respect to the development of GEs, at a minimum, evidence of semantic bleaching is required. In so far as the literature has argued that speakers are sometimes able to use GEs without implicating a set, we may find just such evidence (Erman 1995, Overstreet and Yule 1997:253, Overstreet 1999:43; Aijmer 2002:227, Cheshire 2007:175).

For many utterances containing a GE, it is easy to imagine the set or category that the GE evokes.<sup>39</sup> For example, *and stuff* in (25) could evoke a set of agricultural foodstuffs.

---

<sup>38</sup>Recently I saw a magnet for sale at a gift shop declaring “*The best things in life aren’t things*”, a play on the layering of *things* older meaning of physical material or items and the bleached meaning.

<sup>39</sup>Albeit perhaps not *the* set the speaker intended (see Channell 1994). The set(s) evoked are often *ad hoc* and non-lexicalized (Overstreet 1999:43; Cheshire 2007:175).

(25) They could take their butter and eggs *and stuff* out [to market].

(BLV/F/1897)

In many, if not most, cases the set/category is *ad hoc* and non-lexicalized. For example, in (26), *things like that* might be evoking the set of ‘things kids who grew up on a farm did for fun.’

(26) I remember jumping off the beams into the straw, *things like that*. Burying each other in the wheat when I’d have kids come to play.

(F/1912/NIA)

Although, for the analyst, expressing the set evoked often requires some wordiness, it is typically still easy to express the idea, as demonstrated by O’Keefe (2004), where the GE *and all this* in (27) is marking the “category of things that an Irish teenage boy might say to tease his sister who has a facial hair problem (and even how it might be said).”

(27) I have Emm she’s fourteen and her brother slags her now he’s sixteen he would be going “look you have you have hair unde= you have a moustache” *and all this* so I do have to give out to him.

(O’Keefe 2004:17–18)

However, it has also been observed that some GE tokens lack what is believed to be their core function of set-marking/extension (Erman 1995, Overstreet and Yule 1997:253, Overstreet 1999:43; Aijmer 2002:227, Cheshire 2007:175).<sup>40</sup>

<sup>40</sup>Although Aijmer (2002:216) interprets GEs with non-nominal referents as being bleached of set marking, some times the set is simply a set of predicates. Thus, *and stuff* in Aijmer’s (2002:216) example (10), repeated in i. (with prosodic and phonetic details removed) can be interpreted as

For example, in (28), it is very difficult to imagine the possible set that the speaker might be evoking; the entailment of ‘there is more’ does not seem so strong.

- (28) [The neighbour’s trees] kind of hang over the entire yard, so all the leaves fall on our side *and stuff*, so.

(TOR/M/1984)

As Cheshire (2007:175) notes, from the analysts perspective, it is sometimes difficult to objectively determine a potential category that the speaker might be implicating. In fact for some cases, there is evidence from the surrounding discourse that the addressee did not interpret the GE as functioning to extend a set or even that the speaker did not intend for the GE to do so. This is nicely exemplified in Cheshire’s (2007:176) excerpt 22 (repeated below in (29)), a conversation about horses and horse-back riding.

- (29) yeah but because it’s in the riding school if he’s got lessons *and things* you can’t ride him but he doesn’t usually have that many lessons

(Cheshire 2007:176)

Although the speaker uses *and things*, the subsequent discourse suggests that she did not intend to extend the set which includes *lessons*. The speaker does not want to give the impression that the horse has a busy schedule and thus her riding time is cut short, so she mitigates her statement to let the hearer know that there aren’t too many lessons and therefore, there aren’t too many situations in which she can’t extending a set that might also include *eating health foods*.

- i. A: Does it inhibit exercise of various kinds?  
 B: Oh it does absolutely. Any large movement.  
 A: So you’re going to have to- you’re going to have to watch your weight *and stuff* then.

ride her horse. Crucial, the mitigated statement is only one about lessons, and not one about lessons and the whole set of things like lessons. Likewise, consider (30) from Tagliamonte and Denis (2010:355), a passage about some classmates who made a video for a teacher.

- (30) AS: We threw a baby shower, and then all of a sudden our marks got raised.  
 Yeah, she was going to cry. ‘Cause we made like a video.  
 INT:Yeah, I saw it.  
 AS: You remember, right?  
 INT:Yeah.  
 AS: So ah, we made her a video *and stuff*, and then she’s like, “Guys, I’m going to cry.” She’s like, “But I won’t.”

(TOR/M/1986)

Although the speaker uses the GE *and stuff*, the topic of conversation—that is, what was going to make her cry—is the video and not other things like videos.

To determine the extent of semantic bleaching, Tagliamonte and Denis (2010:355) (following Cheshire 2007:176) code each *stuff* and *thing* GE token for whether or not a set/category can be inferred. This included only tokens where no general set could be determined and where there was some indication from the surrounding discourse that the speaker and/or hearer did not interpret the GE as set-extending (Cheshire 2007:176). I will go on to replicate this analysis, but first consider an alternative method.

### **Another Potential Method**

Wagner, Hesson, and Bybel (2014) take issue with the accountability of coding for set-extension, pointing out that the subjective nature of determining whether or not a particular GE is being used to extend a set introduces possible inconsistencies in



cross-study comparisons. Not only is it often difficult for individual analysts to make objective decisions in individual cases, but in some cases analysts will disagree (Wagner, Hesson, and Bybel 2014:11). In most cases opposing analyses of particular cases can often be justified one way or the other. Following standard variationist argumentation (see also Poplack and Tagliamonte 1999), Wagner, Hesson, and Bybel (2014:13) argue that “[w]ithout knowing the speaker’s actual intensions, a variety of interpretations [are] possible.” Indeed, this is an unavoidable methodological problem for corpus-based research, which is further exacerbated by the fact that research on GEs comes from multiple and (often incompatible) theoretical perspectives (i.e., functionalists and generativists; corpus pragmatics and variationist sociolinguistics etc.) with different benchmarks of evidence and evaluation.

The way forward however is for researchers to be increasingly conscious of using replicable and objective coding procedures. One such procedure is introduced by Wagner, Hesson, and Bybel (2014) and further implemented by Wagner, Hesson, and Little (2014). A simple binary decision tree that takes into account the number of referents within the syntactic and discourse context and syntactic ambiguity is used to objectively code GE tokens for REFERENTIALITY.

Using this decision tree was shown to increase inter-rater reliability of coding for set-extension by nearly 20% (Wagner, Hesson, and Bybel 2014:25). Such inter-rater reliability is a highly desirable result. However, this particular approach suffers from the problem of consistent errors. While their algorithm guarantees that a particular set of GEs that are not semantically bleached are coded correctly, it also guarantees that all structurally ambiguous GEs (that are not necessarily bleached) are coded as not set extending. This is problematic since a full 20% of GE tokens in the Ontario data being examined presently are structurally ambiguous. A hypothetical token such as (31), being structurally ambiguous, would be coded as ‘not set extending’.

(31) Gord is a typical Canadian. He loves maple syrup and stuff.

However, despite the obvious ambiguity of referent, potential sets are easily inferable for the analyst whether the GE attaches to the nominal *maple syrup* (e.g., {maple syrup, Tim Horton's, Kraft Dinner, ...}) or verb phrase *loves maple syrup* (e.g., {loves maple syrup, plays hockey, participates in the Canadian Shift, ...}).

### Replication

For replicability reasons, I follow the 'intuitive' coding method of Cheshire (2007) and Tagliamonte and Denis (2010) for the EOE data.<sup>41</sup> The results are presented in 3.13, along with the results from TEA, as reported in Tagliamonte and Denis (2010).<sup>42</sup>

Table 3.13: Test of semantic bleaching in real time; Frequency of non-set extending GEs by age group, GE type, and length. Toronto data from Tagliamonte and Denis (2010:Table 7)

	EOE		TOR >50		TOR 30-50		TOR <30		TOTAL	
	%	N	%	N	%	N	%	N	%	N
<i>and things</i>	0.0	0/17	9.0	1/11	4.3	1/23	0.0	0/8	3.4	2/59
<i>and things like that</i>	0.0	0/100	5.8	4/68	0.0	0/38	2.3	1/42	2.0	5/248
<i>and stuff</i>	0.0	0/6	5.7	2/35	12.2	7/57	9.8	32/326	9.7	41/424
<i>and stuff like that</i>	0.0	0/20	5.7	3/53	0.0	0/30	8.1	13/160	6.1	16/263

Tagliamonte and Denis (2010:356) observe that "semantic bleaching is modest at best" in Toronto. Indeed, the rates of non-set-extending GEs in TEA are much lower than Cheshire (2007) reports across the board (between 18 and 32 percent). As we can see in Table 3.13, for the two oldest generations in Toronto there are only 10 and 8 tokens respectively in which no set could be determined. Although the raw frequency of non-set-extending GEs is higher in the youngest age group, the proportion does not consistently rise indicating no increasing, ongoing bleaching.

<sup>41</sup>Like Tagliamonte and Denis (2010) a three-way coding distinction was made between those tokens that clearly mark a set, those that clearly don't, and a 'maybe' category for borderline cases.

<sup>42</sup>It is important to note that the coding of set-extension for EOE and TEA is likely more comparable than any other two datasets since the same analyst (me) coded each corpus. Therefore inter-coder differences are avoided.

That said, the data from EOE add nuance to the story in Ontario. Not a single token of GEs in the three EOE communities lacked a set-marking function. Putting the data from EOE and TEA together, we might interpret the results in Table 3.13 as an indication of early semantic bleaching of GEs taking place in twentieth century Ontario English. I will revisit the implications of these results in the next two sections.

### 3.4.4 Pragmatic Shift

#### Background

The last mechanism of grammaticalization, which operates hand in hand with semantic bleaching, is pragmatic shift. The general argument from the grammaticalization literature is that grammaticalizing forms will shift from expressing propositional functions to interactional and interpersonal functions. Assuming that the set-marking function is the propositional function/meaning of a GE, then semantically bleached GEs are likely not vacuous, but rather are serving other (interactional/interpersonal) functions. From a grammaticalization theory perspective, these functions developed gradually with GEs expanding across different pragmatic contexts, triggered by semantic bleaching.

Cheshire (2007:178–183) observes that GEs operate across a range of discourse functions using evidence from the addressee's response, or if the token is within a speaker turn, how the discourse proceeds. Working with Fischer's (1998, 2000) model of discourse structure, she finds GEs operating in terms of information management (introducing discourse-new elements), textual organization (sectioning off reported speech), speech management (indicating a lack of appropriate word), turn-taking (as turn-yielding devices), and with respect to the interpersonal relationships between interlocutors (in terms of politeness and social solidarity).

Although the initial variationist approach to examining such a situation would

be to code every token for the pragmatic function exhibited, Cheshire (2007:183–184) points out a number of pitfalls with this approach. Chiefly, it is difficult for the analyst to ground an interpretation of each GE token objectively. As Labov (1994:549–50) puts it “[t]here is no reason to think that our notions of what we intend or the intentions we attribute to others are very accurate, or that we have any way of knowing whether they are accurate.” A systematic analysis of multiple, exemplary tokens of GEs, for example Aijmer’s (2002) and Cheshire’s (2007), can reveal the range of different functions that GEs *can* serve, but identifying the range of functions is not the same as identifying how any given token is functioning. That is, not all tokens are clearly categorizable on the basis of objective criteria. The coding problem is multiplied by the fact that GEs tend to be simultaneously multifunctional. Any given token can function propositionally, interactionally, and interpersonally. Thus, it is ill-advised to attempt to tease apart a single or core function for every token (Cheshire 2007:183).

As such, Cheshire (2007:185) devises an alternative diagnostic for pragmatic shift, operationalizing the presence/absence of co-occurring discourse markers. Following Aijmer’s (2002:2) observation that co-occurring discourse markers help “addressees to deal with the ambiguity of pragmatic particles”, the working assumption is that GEs that express non-propositional functions (i.e., GEs more advanced in terms of pragmatic shift) will co-occur with other (interactional/interpersonal) discourse markers (such as *like, you know, I mean*) less frequently than GEs which have not shifted (Cheshire 2007:185).<sup>43</sup> Cheshire (2007:185) finds that in two of the three speech communities she examined, the GE types that were found to be most advanced with respect to phonetic reduction, decategorialization, and semantic bleaching tended to co-occur with discourse markers less often than forms that were less advanced with

---

<sup>43</sup>Pichler and Levey (2010:19) point out that the opposite interpretation might also have some currency in the literature, citing Margerie (2007). However, Margerie’s (2007) work discusses contiguous n-gram collocations (e.g., *really kind of*), rather than the co-occurrence of different DMs within an utterance as Cheshire (2007).

Table 3.14: Percentage of co-occurring discourse markers (N) and type frequency.

	EOE		TEA>50		TEA30–50		TEA<30	
	%	Types	%	Types	%	Types	%	Types
<i>things</i>								
short	17.6 (17)	1	36.4 (11)	2	43.5 (23)	4	40.0 (10)	1
long	25.0 (100)	11	25.0 (88)	6	30.0 (40)	6	50.0 (48)	5
<i>stuff</i>								
short	33.3 (6)	1	16.6 (30)	3	36.0 (50)	6	54.4 (296)	9
long	35.0 (20)	2	30.5 (59)	6	28.9 (38)	6	58.0 (200)	10

respect to the diagnostics of those mechanisms.

However, Tagliamonte and Denis (2010:357) found that with respect to the diagnostic of co-occurring discourse markers across apparent time there was no evidence for ongoing pragmatic shift in Toronto.

### Replication

Table 3.14 presents the percentage of tokens of both long and short forms of *things* and *stuff* GEs that co-occurred with another discourse marker in the EOE, plus the results for the same diagnostic from TEA.<sup>44</sup> Following Cheshire (2007:185), turn initial discourse markers were excluded as they typically function to create “coherence between speaker turns.” The table also includes a column for the raw frequency of different types of discourse markers that co-occur with each GE type.

There are two crucial observations to make from Table 3.14. First, as observed by Tagliamonte and Denis (2010:186), the frequency of GEs that co-occur with discourse markers generally increases across time. This is true regardless of GE type and length. Though there is some fluctuation with the middle age speakers in Toronto (e.g., this group has the highest frequency of co-occurring discourse markers for *and things* but also the lowest frequency of co-occurring discourse markers for *and stuff like that*), the youngest speakers have a higher frequency of co-occurring GEs than the

<sup>44</sup>The data is based on Tagliamonte and Denis (2010:357) but the numbers are slightly adjusted due to a slightly different coding schema.

oldest two groups of speakers. The second observation is that there is no systematic correlation with the length of a GE. In some cells, shorter GEs have more co-occurring discourse markers (e.g., for TEA>50 *things*) but in other cells, longer GEs have more (e.g., for TEA>50 *stuff*).

All told, these two observations suggest that there is no evidence for pragmatic shift on the basis of co-occurring discourse markers. The hypothesis was that tokens of GEs that are shifted toward more interpersonal functions will be less likely to co-occur with other discourse markers serving a similar function. If pragmatic shift was underway throughout the twentieth century, we would expect a decrease in the rate of co-occurring discourse markers over time. Furthermore, since the shorter variants of GEs are argued to be on the forefront of grammaticalization (Aijmer 2002; Overstreet 1999; Cheshire 2007), shorter variants are hypothesized to co-occur with discourse markers less frequently than longer forms. That there is no evidence for this hypothesis is unsurprising considering the argument above in §3.4.1 that shorter GEs are not shorter due to phonetic reduction *cum* grammaticalization.

### Problems and Other Methods

A major problem with this diagnostic is that the increase over apparent time may be caused by something else. Co-occurrence of GEs with other discourse markers will necessarily be collinear with the frequency with which individuals use discourse markers generally. Thus, the increase (or non-decrease) of co-occurring discourse markers across apparent time may be independent of pragmatic shift (or non-pragmatic shift) of GEs. Figure 3.14 collapses across all GE types and examines the proportion of GEs that co-occur with other discourse markers over apparent time in each of the four age groups, separated by colour. The curves are scatterplot smoothing lines created by locally-weighted regression (see Baayen 2008 34–35; Labov et al. 2013; Tagliamonte and Denis 2014). There is a heightened use of co-occurring

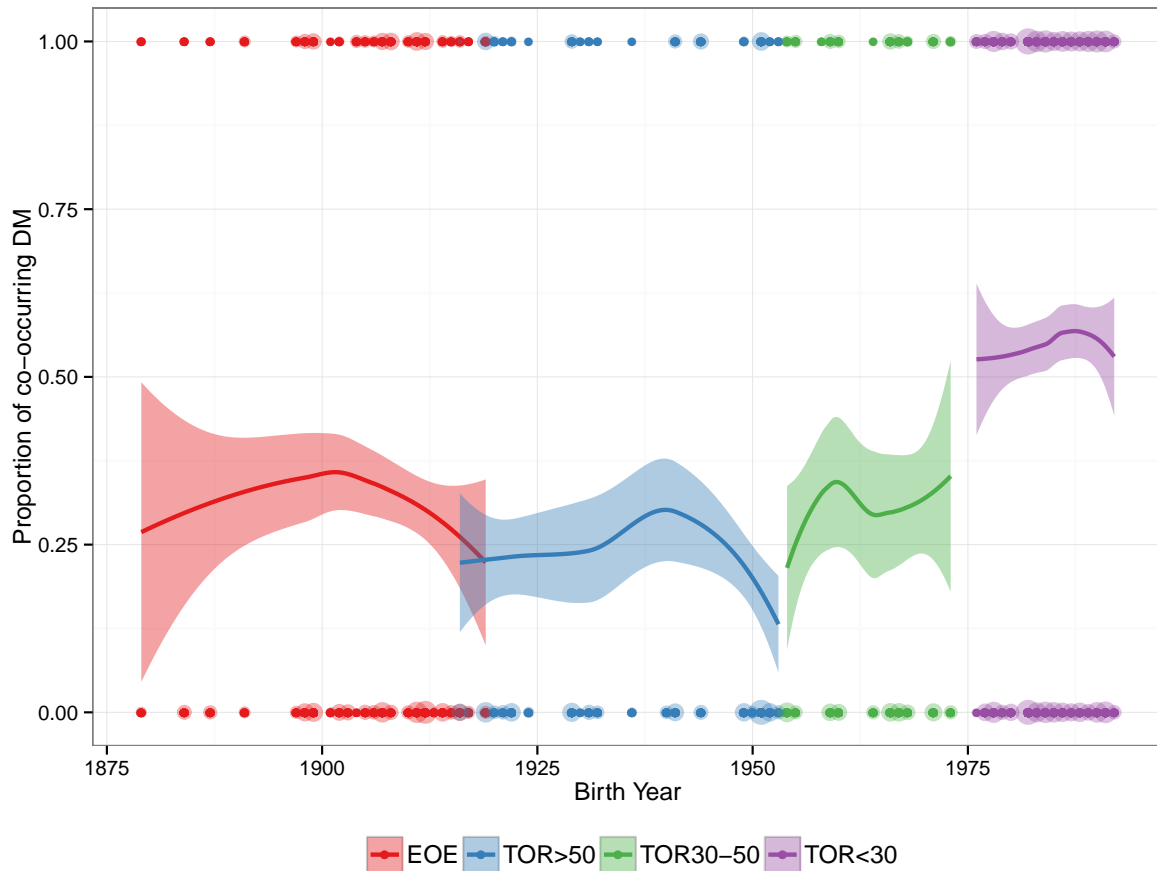


Figure 3.14: Trajectory of the proportion of GEs that co-occur with other discourse markers over apparent time. Curves represent locally-weighted regression scatterplot smoothing (with confidence interval). Span parameter = 1.  $N = 2821$ .

discourse markers among the youngest generation, from around thirty-five percent in the oldest two generations to well over fifty percent. However, this may well be due to a generally higher rate of use of discourse markers among younger people, as observed above in Figure 3.2 for GEs themselves. If this is the case, then it is difficult to tease apart this possible age-grading effect from any effect that pragmatic shift might have on co-occurring discourse markers without knowing the apparent time trajectory of the use of discourse markers generally.

Pichler and Levey (2011) forego the use of this diagnostic entirely, opting rather to use a coding taxonomy for semantic-pragmatic change together. The taxonomy

Table 3.15: Pichler and Levey's (2011:452) taxonomy of semantic-pragmatic change in GEs

Stage	Function(s)
Stage 0	set-marking (contingent on intersubjectivity)
Stage 1	set-marking and interpersonal/interactional
Stage 2	interpersonal/interactional
Stage 3	punctor (devoid of referential and pragmatic meaning)

includes four stages of change, as in Table 3.15.

GEs at Stage 0 are strictly used to extend sets and no interpersonal or interactional functions are apparent. Tokens at the next stage are multifunctional, serving both a set-marking function and interpersonal and/or interactional functions. At Stage 2, GEs are semantically bleached of their set-marking function, serving only interpersonal and/or interactional functions. Lastly, GEs at Stage 3 serve as punctors: pragmatic markers that have been classified as “nervous tics, fillers, or signs of hesitation” (Vincent and Sankoff 1992:205).

Although Pichler and Levey (2011) use multiple coders to cross check the categorization of individual tokens, this method still suffers from some subjectivity. In particular the distinctions between Stage 0 and Stage 1 and between Stage 2 and Stage 3 are fuzzy. Because GEs at Stage 0 mark sets, and this is contingent on intersubjectivity between interlocutors and linguistic features that function as markers of intersubjectivity are function interpersonally, it seems that any linguistic feature that marks a set is inherently functioning interpersonally (as described in §3.2.1) and would thus fall into the category of Stage 1. In fact, this idea is implicit in many discussions of the meaning and function of GEs. Pichler and Levey (2011:450) themselves state that “set-marking GEs are inherently intersubjective because they appeal to common frames of knowledge by implicitly evoking a more general set.” Aijmer (2002:240) argues that “by referring to common ground”, the main function of *and things* is to suggest that “we [interlocutors] have a lot [in] common and therefore you



know what I mean” and thus, this function “is associated with positive politeness.” Overstreet and Yule (2002:787) observe that speakers using GEs are “making a call upon familiarity with assumed common ground” and thus, the GE functions as “a marker of intersubjectivity.” For each of these researchers, it is acknowledged that at the core of a GE is *both* a propositional and interpersonal function and these are inherently linked.

Likewise, GEs at Stage 3, which function as punctors, seem to necessarily function interactionally. Vincent and Sankoff (1992:208–209) discuss four prosodic contexts in which punctors are used for different interactional functions including “regulation” (marking hesitations, false starts etc.), “segmentation” (including the marking of topic or parentheticals/appositions), and “discourse” (marking “transition between stretches of discourse”). Thus, as an empirical metric this taxonomy may suffer from subjective categorization on part of the analyst. However, as a theory of the development of semantic-pragmatic change of GEs, this taxonomy is important.

### A Theoretical Trajectory of Semantic-Pragmatic Change of GEs

Given the caveats just discussed regarding Pichler and Levey’s (2011:452) taxonomy of semantic-pragmatic change, a simpler model of change is more accurate. To the extent that GEs are undergoing semantic-pragmatic changes, the trajectory is something like (32).

(32) SET-EXTENSION+INTERPERSONAL(/INTERACTIONAL) > INTERPERSONAL(/INTERACTIONAL)

That is, GEs at an earlier stage (inherently) function both propositionally and interpersonally (and possibly interactionally). Over time, the propositional function of set-extension has been lost: this is apparent in the few tokens in TEA where no set-marking function could be determined. Thus, the situation of change in (32) is consis-

tent with Heine's (2003:591–2) *bleaching model* of grammaticalization, as schematized in (33).

(33)  $ab > b$

The bleaching model of change holds that semantic developments “entail a loss in semantic content of the item concerned” (Heine 2003:591). For example, Heine (2003:591) suggests that when demonstratives develop into definite articles, they lose their deictic meaning (yet retain their definiteness).<sup>45</sup> This is exactly what seems to have happened with GEs. Those GEs more advanced with respect to semantic-pragmatic change are those tokens for which the set-extension (i.e., propositional) function has been bleached, yet the interpersonal function has been retained (as in Table 3.13).

If we consider GEs to have always had an interpersonal function, as argued for here, we can now explain the pattern in Figure 3.14. Ignoring the probable age-grading with the youngest generation, there was no change to the frequency of co-occurring (interpersonal) discourse markers as semantic change took place (at least in the expected direction), precisely because the interpersonal element of GEs was always present. This is consistent with Waltireit's (2006:75) observation that pragmatic markers tend to develop from linguistic elements that “already have some properties typical for discourse markers.” I will return to this idea in Chapter 5.

Heine (2003:591) claims that cases of semantic bleaching like this are the “*sine qua non* for grammaticalization to happen.” However, all told, does the evidence presented in this chapter support an account consistent with grammaticalization theory?

---

<sup>45</sup>In a feature-geometric view of morphosyntactic features of nominals such as Cowper and Hall (2002), this bleaching might involve a delinking of the [deictic] feature, which is dependent on the [definite] feature.

### 3.5 The Development of GEs: Multiple, Independent Changes

I would like to argue that the story of the development of GEs, a set of pragmatic markers, is not a story about canonical grammaticalization. Rather, the changes documented in this chapter are attributable to other, multiple, independent changes. I will discuss the evidence for this argument in this last section of the chapter, but first I review what would be necessary to definitively identify a change as grammaticalization according to grammaticalization theory.

Although some grammaticalization theorists discuss grammaticalization as a single process and others discuss it in terms of a series of processes, it is critical to grammaticalization, as a theory of language change, that the component mechanisms are interrelated. (34) presents a sampling of grammaticalization theorists' views on the correlation of a series of changes that together make up grammaticalization.<sup>46</sup>

- (34) a. Lehmann (1982/1995:v) "Grammaticalization is a process leading from lexemes to grammatical formatives. A number of **semantic, syntactic and phonological processes interact** in the grammaticalization of morphemes and of whole constructions."
- b. Heine (2003:583): The four mechanisms of grammaticalization "and **the way they are interrelated**" account for the process of grammaticalization "irrespective of how one wishes to define a 'distinct process'".
- c. Heine (2003:583): "[G]rammaticalization, as conceived here, is above all a semantic process. Desemanticization results from the use of forms for concrete meanings that are reinterpreted in specific contexts as more abstract, grammatical meanings. Having acquired grammatical meanings,

---

<sup>46</sup>Bolded portions highlight these correlations.

these forms tend to become increasingly divergent from their old uses: they are used in new contexts (extension); lose in categorial properties characteristic of their old uses, hence undergo decategorialization; and tend to be used more frequently, become more predictable in their occurrence, and, consequently, lose in phonetic substance. Thus, **the four mechanisms are not independent of one another**; rather desemanticization precedes and is immediately responsible for decategorialization and erosion.”

- d. Traugott (2003:643–4): “[E]arly grammaticalization can therefore be seen as **a complex set of correlated changes**: i. structural decategorialization; ii. shift from membership in a relatively open set to membership in a relatively closed one (i.e., from lexical category to syntactic operator category) in the context of a specific construction; iii. bonding (erasure of morphological boundaries) within a construction; iv. semantic and pragmatic shift from more to less referential meanings via invited inferring [and] phonological attrition, which may result in the development of paradigmatic zero (Bybee et al. 1994).”
- e. Diewald (2010:19): “[I]t has been common knowledge from the very beginning of modern work on grammaticalization that **grammaticalization processes are of a composite nature**.”
- f. Diewald (2010:20): “The distinctive and unique feature of grammaticalization is generally seen in **its particular combination and serialization of several processes**.”

Thus, to identify grammaticalization, one must demonstrate that each sub-process, change, or mechanism, however conceived, is active, or more specifically in Heine’s (2003) view, that semantic bleaching has diachronically triggered the other mecha-

nisms. If grammaticalization theory is to be a theory of change, it must be falsifiable. Therefore, any case in which the activity of one or more mechanisms cannot be demonstrated must not be a case of grammaticalization.

The task of the grammaticalization theorist is made difficult by the fact that, as Diewald (2010:35) observes, “[t]here is growing agreement that none of [these] sub-processes is restricted to grammaticalization.” We may find several kinds of changes in progress which resemble sub-processes of grammaticalization but, independently of the other processes, they cannot be considered grammaticalization. However, if analysts can demonstrate that the series of processes/changes/mechanisms are active in the course of some change, and that these processes/changes/mechanisms are not otherwise independent, then I would argue that this is a true case of grammaticalization. With this in mind, Table 3.16 summarizes what I have argued for in this chapter regarding each mechanism of grammaticalization and the innovative *stuff* type GEs in Ontario English across the 20<sup>th</sup> century.

Table 3.16: Summary of the mechanisms of grammaticalization on *stuff* type GEs in Ontario English.

Mechanism	Result
Phonetic reduction	Putative phonetic reduction is the result of an independent process of morphological clipping of <i>like that</i> .
Decategorialization	No quantitative evidence of decategorialization from expected to unexpected referents for either short or long GEs. Theoretically this is unsurprising given the proposed compositional semantics of GEs.
Semantic bleaching	The beginnings of bleached GEs (i.e., GEs without a set-extension function) (but not correlated with GE length).
Pragmatic shift	No decrease in the use of co-occurring discourse markers. But, since GEs, as inherently intersubjective, have always had an interpersonal function, this is not unexpected.

So what then are we left with? I have argued that what many have argued is the PHONETIC REDUCTION of GEs is an independent change that is not related to the gradual erosion of phonological material, associated with grammaticalization. I have

shown that *and stuff* has always, from its inception, co-occurred with unexpected referents and thus, there is no evidence of DECATEGORYALIZATION. In fact, if the semantics of GEs are as I have proposed, decategorialization is moot. There is some evidence for SEMANTIC BLEACHING of GEs across the 20<sup>th</sup> century. None of the tokens in the real-time benchmark data could be conceived as lacking a set-marking function but there are tokens in the more recent data from TEA where no set-marking function is observed. Lastly, the evidence for PRAGMATIC SHIFT from co-occurring discourse markers runs counter to the expectation. However, because I have argued that GEs did not change from expressing a propositional to an interpersonal function and, therefore, the interpersonal function was always present, it is no surprise that the frequency of co-occurring discourse markers is stable.

In summary, the only mechanism of grammaticalization that I have found to be ongoing since the late nineteenth century in Ontario English is semantic bleaching. However, since grammaticalization theory defines grammaticalization as the confluence of all of these mechanisms, grammaticalization theory is not a model for the development of GEs in Ontario English. Instead, it is a complex variable system that has undergone a number of independent changes across the last one hundred and twenty years. *Stuff* type GEs have risen in frequency and become the majority form in the variable system. This has occurred as a result of LEXICAL REPLACEMENT, as suggested by Tagliamonte and Denis (2010). At the same time, there has been a growing tendency to delete the comparative element *like that*. Lastly, there is potentially a new semantic-pragmatic change underway—the bleaching of the propositional meaning of GEs, leaving just the interpersonal meaning, present in the implicata of non-bleached GEs. Thus, what we have is a case of semantic-pragmatic change that did not follow the gradual path of grammaticalization delineated by grammaticalization theory.

## Chapter 4

# The Development of Epistemic Parentheticals

*'I suppose someone could have seen them,' said Lewis slowly.*

*'I'm pretty sure someone did,' said Morse.*

*'And you think you know who it was?'*

*Again Morse nodded. 'So do you, I think.'*

*Did he? In such situations Lewis had learned to play it cleverly. 'You mean...?'*

*Last Seen Wearing, Colin Dexter, 1976[2007], Pan Books*

### 4.1 Introduction

In this chapter, I examine the development of another set of pragmatic markers, epistemic parentheticals (EPs) and test a number of diagnostics of grammaticalization that have been proposed in the literature over real and apparent time. Because EPs have spent less time in the variationist 'spotlight' than general extenders, a second goal is to demonstrate that EPs are a variable system, the realization of its variants being subject to a series of linguistic constraints.

This chapter is organized as follows. In §4.2, I discuss both the grammaticalization theory and variationist literature that has addressed EPs, paying specific attention to potential diagnostics of ongoing grammaticalization that have been proposed. I then go on to describe how I delimit the variable context of the EPs system. In §4.4, I examine the distribution of the EPs system across the twentieth century and test three diagnostics that have been implicated in the grammaticalization of EPs operationalized as linguistic constraints on variation: intervening material, syntactic position, and the epistemic/doxastic strength of the verb. I then make some conclusions.

## 4.2 Background

In what follows, I treat the set of epistemic/doxastic<sup>1</sup>, complement taking, matrix verbs that co-occur with first person pronominal subjects as a variable set of pragmatic markers that express a degree of speaker commitment to a proposition.<sup>2</sup> EPs have been variously referred to as (or been included in the set of) “comment clauses” (Dehé and Wichmann 2010), “main clause-like comment clauses” (Quirk et al. 1985:977; Kaltenböck 2013), “modal particles” (Aijmer 1997), “epistemic qualifiers” (Caton 1969, cited in Aijmer 1997), “epistemic quantifiers” (Givón 1989:134), “epistemic phrases” (Thompson and Mulac 1991), “complement-taking (mental) predicates” (Thompson 2002; Van Bogaert 2011), “parentheticals” (Fischer 2007; Kaltenböck 2007), “parenthetical verbs” (Urmson 1963), “epistemic/evidential parentheticals” (Brinton 2008), “epistemic/evidential complement-taking predicates” (Rodríguez Louro and Harris 2013), and “pragmatic particles” (Östman 1982).

---

<sup>1</sup>The literature refers to EPs as relating to epistemic modality (typically ‘what is known’). However, these features seem to relate less to speakers’ knowledge and more to their beliefs. While beliefs are sometimes considered to be part of epistemic logic, we might more accurately relate EPs to doxastic modality (i.e., concerning ‘what is believed’) (for general discussion see Kratzer [2006] and Stalnaker [2002], see also Aijmer [1997] on EPs as relating to ‘beliefs’).

<sup>2</sup>I exclude (semi-)factive verbs such as *know* and *realize*, which presuppose that their complement propositions are true (*inter alia*, Karttunen 1971; Hooper 1975; von Stechow 1999), in order to focus on verbs that, themselves may vary with respect to speaker commitment.



As exhibited in (1), EPs can occur in clause-initial position (1a-b), clause-final position (1c), or clause-medial position (1d).

- (1) a. *I think* that they used to go in the winter as well.  
(BLV/F/1897)
- b. *I suppose*  $\emptyset$  I hoed likely from the time I was big enough to hoe the hoe.  
(NIA/M/1898)
- c. I can recall staying overnight one time too in an hotel *I think*.  
(EON/F/1914)
- d. I done all those kind of things whichever a young person done *I guess* on the farm.  
(EON/M/1912)

The literature on EPs come from either a grammaticalization perspective or a variationist perspective. As with GEs, grammaticalization theory informs the variationist work.

Thompson and Mulac's (1991) analysis of EPs, briefly discussed in Chapter 2, not only lays the foundation for all subsequent work on EPs, but is a prime example of quantitative reasoning in grammaticalization theory, setting out multiple, testable hypotheses about the development of EPs. Beginning with the argument that the (apparent) alternation between overt and null complementizers in English is "better understood as an alternation between constructions like [(2a)], in which *I* and *think* are main subject and verb, with *that* introducing a complement clause, and constructions like [(2b)] and [(2c)] in which *I think* is an epistemic phrase, [...] functioning roughly as an epistemic adverb such as *maybe* with respect to the clause it is associated with," Thompson and Mulac (1991:313) propose that EPs have undergone a

process of grammaticalization. Their strong hypothesis is that tokens such as (2b) are grammaticalized forms of tokens such as (2a), having grammaticalized from main clauses introducing complement clauses to epistemic adverbial phrases (Thompson and Mulac 1991:317–8).

- (2) a. I think *that* we're definitely moving towards being more technological.  
 b. I think  $\emptyset$  exercise is really beneficial, to anybody.  
 c. It's just your point of view you know what you like to do in your spare time *I think*

(Thompson and Mulac 1991:313, ex. 1–3)

Their evidence is twofold: 1) the subject and verb combinations that most frequently appear without *that* in contexts like (2b) are those subject and verb combinations that also most frequently appear in a position other than before the complement clause as in (2c); and, 2) the verbs that appear without *that* are epistemic. Thus, the cline of grammaticalization of EPs according to Thompson and Mulac (1991) is as in (3) (where P is a proposition).

- (3) Stage i: [<sub>CP</sub> I think [<sub>CP</sub> that P ] ]  
 Stage ii: [<sub>CP</sub> [<sub>ADV P</sub> I think ] P ]  
 Stage iii: [<sub>CP</sub> P [<sub>ADV P</sub> I think ] ]

Thompson and Mulac (1991) propose that the grammaticalization of EPs is due to reanalysis. However, as discussed in Chapter 1, if such a change is reducible to some other well-established method of linguistic change, there is no need to appeal to grammaticalization (see e.g., Joseph [2001]). That said, Kearns (2007:478) points out that the proposed reanalysis in (3) is simply “not consistent with the general assumption that the *surface* form is not altered by the underlying structural change.”

That is, the difference between Stage i and Stage ii is not just structural—the result of rebracketing—but the contents of Stage i and Stage ii differ, with Stage i exhibiting an overt *that* where Stage ii has nothing. Kearns (2007:478) argues that if EPs did develop via reanalysis, there must be an intermediate structure between Stage i and Stage ii, a stage which exhibits a null complementizer, as in (4).<sup>3</sup>

(4) [CP I think [CP Ø P ] ]

In fact, Jaeger (2010:35) points out that even Thompson and Mulac's (1991) results are not consistent with the strongest version of their hypothesis since when EPs are excluded from their data, "there is still considerable variation in *that*-mentioning [i.e. the tokens that contain an overt complementizer -DD] that needs to be accounted for." In other words, since there are clear examples of non-EP main verbs that introduce embedded clauses without an overt complementizer, it is unlikely that the absence of an overt complementizer with epistemic verbs is a clear indication that they are functioning as EPs. For example, Aijmer (1997:8) observes with respect to one such non-EP main verb (*say*) that "[i]t would be [...] strange to say that *John said* in *John said Bill was fat* with zero has been reanalysed as an adverb."

Other grammaticalization theorists have proposed different accounts of the development of EPs. Most notably, Brinton (1996; 2008:48) rejects Thompson and Mulac's (1991) 'matrix-clause hypothesis' as "the chronology of events proves difficult to establish." In particular, Brinton (1996:246–8) finds that Old English lacked a sufficient number of Stage i type tokens—that is, matrix clauses with overt *that* complementizers<sup>4</sup>—and that the most common EP constructions in Middle English were crucially not those that occurred with the most Ø complementizers, as predicted by Thomp-

<sup>3</sup>Kearns (2007:503) argues that null complementizers correlate with an increase of "the informational prominence" of embedded clauses.

<sup>4</sup>Rissanen (1991:283) suggests that the Ø variant "may have been the unmarked link in speech throughout the Old and Early Middle English period." See also Warner (1982).

son and Mulac (1991). Rather, Brinton (1996:241) argues that EPs developed from a Middle English relative clause structure, which itself developed from a series of Old English “syntactically complete clauses with an anaphoric demonstrative referring back to the preceding clause, that is, relative clauses” as in (5). Here, the pronoun *þæs* functions to introduce a “sentential relative clause.”

- (5) Habbað we to þæm mæran micel ærende, / Deniga frean,  
 Have we to him celebrated<sub>DAT.SG</sub> big errand, / Danes<sub>GEN.PL</sub> lord<sub>DAT.SG</sub>,  
 ne sceal þær dyrne sum / wesan, *þæs ic wene*  
 nor shall there secret any / be, *this<sub>GEN.SG</sub> I think*  
 ‘We have for the famous lord of the Danes, a great errand; nor shall anything  
 there be secret, *of this I think.*’

(Beowulf, cited in Brinton 1996:241 [morpheme gloss -DD])

By Middle English, another set of EPs, in addition to those containing anaphoric pronouns (as in (6)), develop: *as/so* parentheticals. Brinton (1996:250-1) argues that *as/so* in these EPs, as in (7), also function as relative pronouns.<sup>5</sup>

- (6) He lese shal; *thereof have I no doubt*  
 He lose shall; *thereof have I no doubt*  
 ‘He shall lose; *thereof I have no doubt*’

(Chaucer, *Canterbury Tales*, cited in Brinton 1996:249 [gloss -DD])

<sup>5</sup>Although Brinton (1996:241) suggests that Stage i EPs (i.e., *þæs ic wene* ‘of this I think’ in (5)) are essentially relative clauses, Fischer (2007) argues that they are “more loosely connected paratactic clause[s],” essentially independent, adverbial clauses. Given a parallel development in Dutch, Fischer (2007) concludes that a relative clause stage (Stage ii in (8) below) is not necessary and that EPs developed out of these independent clauses directly. In later work, Brinton (2008:45) seems to agree with Fischer’s (2007) assessment.

- (7) “She hath ynough to doone, hardyly,/ To wynnen from hire fader, *so trowe*  
 She has enough to do, hardily,/ to part from her father, *so believe*  
*I.*”  
*I*  
 ‘She has enough to do, assuredly, to get away from her father, *so I believe* [= ‘which I believe’]’

(Chaucer, *Troilus and Criseyde*, cited in Brinton 1996:250 [gloss -DD])

Brinton (1996:251) suggests that at the next stage, the anaphoric pronominal, whether *as*, *so* or a *demonstrative*, is deleted, leaving just the subject and verb.<sup>6</sup> At this stage, the subject and verb were reanalyzable as (syntactically mobile) adverbial phrases. Thus, the diachronic development of EPs for Brinton (1996:252, simplified) is summarized in (8).

- (8) Stage i: [They are poisonous.]<sub>i</sub> *That<sub>i</sub> I think.*  
 Stage ii: They are poisonous, *as I think.*<sup>7</sup>  
 Stage iii: They are poisonous, *I think.*  
 Stage iv: *I think, they are poisonous./They are, I think, poisonous.*

Of the variationist work that has discussed EPs, the majority of researchers have considered them tangentially, in the context of variation between the overt/*that* and null/ $\emptyset$  complementizer (Tagliamonte and Smith 2005; Blondeau and Nagy 2008; Torres Cacoullos and Walker 2009a).<sup>8</sup> This variation is exhibited in (1a) and (1b) above and, within a single utterance in (9) from Tagliamonte and Smith (2005:289) below.

<sup>6</sup>Alternatively, “*as* is grammaticalized as a pure subordinator introducing an adverbial clause” (Brinton 1996:251).

<sup>7</sup>This stage is likely unnecessary. See note 5.

<sup>8</sup>This is the reverse situation with the grammaticalization theory literature, which focuses on EPs and tangentially discusses complementizer variation.

- (9) I wish *that* forty or fifty years ago I'd as much confidence. I wish  $\emptyset$  I'd had it then.

Tagliamonte and Smith (2005) examine vernacular dialect data representing four geographically isolated communities in Scotland, Northern Ireland, and northern England. They argue that the variation between the overt and null complementizer is part of a long term change, underway since at least Early Middle English (Warner 1982; Rissanen 1991:279–82). Furthermore, Tagliamonte and Smith (2005:300) observe a near-categorical rate of the  $\emptyset$  variant with *I think* (along with *you know* and *I mean*). Their results are particularly relevant to the present discussion of the development of EPs as they argue that what they observe “provides striking confirmation of Thompson and Mulac’s (1991) hypothesis in highly vernacular English usage” (Tagliamonte and Smith 2005:300); *I think* is not a matrix clause with a subordinate complement clause in these cases, but rather is functioning as an EP. Tagliamonte and Smith (2005) offer further evidence, which they suggest supports Thompson and Mulac (1991): first person, singular pronominal matrix subjects (i.e., those that look like EPs) favour the  $\emptyset$  variant (Tagliamonte and Smith 2005:303); the presence of a modal or negation between the subject and verb favours  $\emptyset$  (Tagliamonte and Smith 2005:304); when the subject of the complement is a personal pronoun,  $\emptyset$  is favoured (Tagliamonte and Smith 2005:304); and, present tense verbs favour  $\emptyset$  (Tagliamonte and Smith 2005:304).

However, as with Thompson and Mulac’s (1991) own results (see Jaeger 2010:35), these facts all add up to support only a *weak* version of the matrix-verb hypothesis. It cannot be the case that all occurrences of the  $\emptyset$  variant represent grammaticalized EPs. As Tagliamonte and Smith (2005:306) observe “even when the most frequent epistemic parenthetical in the data [DD: *I think*] is removed from the analysis, we still have an overwhelming preference for *zero*.” According to the strongest version

of Thompson and Mulac's (1991) hypothesis, the  $\emptyset$  variant should never occur with non-epistemic verbs, in non-present tense, or with intervening material. While it is true that the  $\emptyset$  variant is *favoured* in these contexts, not all tokens with a  $\emptyset$  variant in Tagliamonte and Smith's (2005) data are epistemic, are present tense, or contain no intervening material, as in (10).<sup>9</sup> Since at least some tokens with the  $\emptyset$  variant must not be EPs, there must be a null complementizer in at least some of these cases. This is exactly what Jaeger (2010:35) observes in Thompson and Mulac's (1991) own data and it supports Kearns' (2007:478) argument.

- (10) a. To prove  $\emptyset$  I could do it. Yes. I had to prove *that* I could do it.  
(Ayrshire)
- b. The teachers thought  $\emptyset$  this was a great idea.  
(Ayrshire)
- c. For you see, I always thought  $\emptyset$  mi mother was coming back.  
(Cullybackey)

Although the early change of *I think* to an epistemic adverbial<sup>10</sup> may have acted like "an arrowhead in the initial motivation of this change," the complementizer of all matrix verbs which take sentential complements has been changing toward a null variant via reanalysis and analogy (Tagliamonte and Smith 2005:307).

Torres Cacoullos and Walker (2009a) also use variationist methods in an attempt to tease apart EPs from sentential complement-taking predicates. Their initial hypothesis is that those frequent subject and verb collocations that have developed ("conventionalized") into EPs should behave differently from complement-taking predi-

<sup>9</sup>I am very grateful to Sali Tagliamonte for allowing me search through the original token file used in Tagliamonte and Smith (2005) to find examples b. and c. Example a. appears in print.

<sup>10</sup>Note that Tagliamonte and Smith (2005:305–6), following Palander-Collin (1997) link the initial grammaticalization of *I think* to an earlier grammaticalization of *me thinks*.

cates (Torres Cacoullos and Walker 2009a:9). For one, the high frequency collocations should have a higher rate of the  $\emptyset$  variant. Indeed, the EPs *I think* and *I guess* co-occur with *that* at a substantially lower rate than the general average (five and three percent respectively, versus thirty-one percent for infrequent collocations). This leads Torres Cacoullos and Walker (2009a:21) to ask “if the highly frequent collocations are discourse formulas that express speaker stance, are they better analyzed as belonging to the lexicon as fixed or frozen (discourse-pragmatic) units or to a productive grammar as instantiations of a construction with open-class positions?” To find out, they examine the linguistic constraints on *that*/ $\emptyset$  variation in both high frequency collocations and all other complement-taking predicates, hypothesizing that if the high-frequency collocations are fixed discourse formulas, their constraints will differ from the constraints on infrequent collocations (Torres Cacoullos and Walker 2009a:9). They find a strong favouring effect for *that* with frequent collocations when an adverbial is present between *I* and the verb and argue that this indicates that these collocations are discourse formulas. In other words, once a high frequency collocation (as in (11a)) is split apart by intervening material (as in (11b)) (and is thus, no longer a discourse formula), it behaves more like any other complement-taking predicate.

(11) a. I think  $\emptyset$  Brian went to McGill.

(Torres Cacoullos and Walker 2009a:ex. 21a)

b. I personally think that it is well worthwhile.

(Torres Cacoullos and Walker 2009a:ex. 24b)

Additionally, Torres Cacoullos and Walker (2009a:34) find a parallel constraint grammar across frequent and infrequent collocations. Taken together, these two observations indicate that high frequency collocations exhibit ‘grammatical persistence’ (Torres Cacoullos and Walker 2009a:34). Torres Cacoullos and Walker (2009a:34) con-



clude that the high frequency collocations are indeed discourse formulas, but that the grammatical conditioning of grammaticalized/-ing constructions “persist in the development of discourse formulas.”<sup>11</sup>

However, although high frequency collocations with intervening material do behave *more* like other complement-taking predicates, the rate of *that* is still markedly lower (fourteen percent versus approximately thirty-four percent). There is no clear quantitative division between high frequency discourse formula (like *I think...* and *I guess...*) on the one hand, and infrequent complement-taking predicates (like *understand...* and *be happy*) and high frequency collocations with intervening material (like *I personally think...*), on the other hand. Furthermore, since the same constraint is operative with the infrequent collocations, in the same direction and with nearly the same strength of effect, it is not clear if this evidence should be interpreted as support that high frequency collocations are discourse formulas stored in the lexicon, or simply as evidence of a consistent effect, independent of frequency and the development of discourse formula.

Thus far, I have exclusively discussed how EPs, in general, may have developed into pragmatic markers. But what about their further development? What changes have taken place in the variable system of EPs? And, are these changes triggered by (or do they even trigger) further grammaticalization?

Kaltenböck (2013) tracks what he argues to be just such further grammaticalization of EPs in the *Diachronic Corpus of Present-Day Spoken English* (which includes the *London-Lund Corpus* recorded between 1958 and 1977 and the *British Component of the International Corpus of English* compiled in the early 1990s) and the *Corpus of London Teenage Language* (a corpus of spoken, teenage English recorded in 1993). In particular, Kaltenböck (2013:287) examines what he argues is the continuing grammaticalization of *I think* as it shifts from a marker of epistemic uncertainty to a general pragmatic

---

<sup>11</sup>That said, Blondeau and Nagy (2008) were not able to replicate this result for Montreal English.

marker “with important textual and discourse-organizational functions” including as a filler, boundary marker, and introducing different perspectives (all discussed in Kärkkäinen 2003).<sup>12</sup> As evidence for this continued grammaticalization, Kaltenböck (2013:287) observes that the form *I think* has decreased in frequency in clause-final position over time. Since this position, Kaltenböck (2013:295) suggests, is the least likely position to function textually/discourse-organizationally, the form *I think*, in all positions, must have shifted toward these more advanced pragmatic functions. That is, a decrease in clause-final *I think* is taken to be indicative of an increase in textual/discourse-organizational functions. He further observes an increase in co-occurrence with other textual/discourse-organizational markers/fillers (such as *well*, *actually*, and *I mean*), which is interpreted as evidence that *I think* is also increasingly functioning as a filler (Kaltenböck 2013:299).<sup>13</sup> Kaltenböck (2013) also notes a decrease in the rate of the complementizer and an increase in “phrasal” tokens, in which the EP has scope over a non-clausal constituent such as noun phrase in (12), and independent tokens of *I think*.

- (12) Uh in the uhm, *I think* October issue of Computational uh Linguistics there’s an attempt to do something of this type

(ICE-GM:S1A-024#105)

Kaltenböck (2013:301) then goes on to examine what is being used in *I think*’s stead as a marker of epistemic uncertainty as the form itself is becoming increasingly textual in function. He argues that three EP variants wax exactly when *I think* wanes: *I’m thinking*, *I just think* and *I’m guessing*.<sup>14</sup> Each of these three forms exhibit an increase

<sup>12</sup>For Kaltenböck (2013:301), this is a change from more conceptual to more procedural meaning (cf. Traugott and Dasher 2002).

<sup>13</sup>Note that this is the opposite intuition of Cheshire (2007) with respect to co-occurrence of discourse markers and GEs.

<sup>14</sup>This analysis uses the *Corpus of Historical American English*.

in frequency (normalized to tokens per million words) in the 1990s and 2000s when *I think* decreases in frequency. Kaltenböck (2013:303-4) argues that this result is “consistent with a view that sees *I’m thinking* [and *I just think* and *I’m guessing*] as taking over some of the epistemic function from *I think*.”

There are a number of problems with these arguments with which to take issue. First, as Kaltenböck (2013:302) acknowledges, the increase of *I’m thinking* in the 1990s is “in line with the attested rise of the [stative] progressive in general (e.g., Mair 2006:89; Aarts, Close, and Wallis 2010; Levin [2013]; Smith and Leech [2013]).” However, the general increase of stative progressives in English may itself be the explanation for the increase in *I’m thinking*; the progressive variant is not replacing *I think* in epistemic contexts, but is rather a specific instantiation of a more general change in progress in the language. Indeed, the EP usage of *I’m thinking* in Kaltenböck’s (2013:304, Fig. 12.5) data increases in lock-step with non-EP usage of *I’m thinking*. The hypothesis about *I just think* may suffer from a different problem: the discourse marker use of *just* also experienced an increase in frequency in North American English in the 1990s (Tagliamonte 2005:1905). Furthermore, *I just think* never appears in clause-final position, the position which Kaltenböck (2013:295) argues privileges the epistemic function of EPs. If *I just think* was being used for this function, we would expect to find clause-final tokens.

The rise of *I’m guessing* as a replacement for *I think* would be the most convincing argument, and the examination of variation between the realizations of the verb of the EP is the main empirical thrust of this chapter; however, there are also a number of problems with Kaltenböck’s (2013) general hypothesis that *I think* has further grammaticalized. Two of these problems arise from not treating EPs as a variable system, and as such not examining frequency accountably (in a Labovian sense; Labov 1972). First, although the (normalized) frequency of *I think* has decreased in clause-final position in the *Diachronic Corpus of Present-Day Spoken English*, if we do not

know how frequently this context occurs in each time period, the result is potentially misleading, as demonstrated with quotative *be like* and internal dialogue in §3.4.2.<sup>15</sup> Likewise, although there is an increase in the number of tokens of *I think* that co-occur with other textually functioning discourse markers/fillers, we do not know the extent to which this is true of all other EP variants. Both of these observations are used, ostensibly, as evidence for the pragmatic shift of *I think*. However, the significance of these results is unknown without an accountable examination of the whole variable system. A second issue is that the apparent increase in scope of *I think* (i.e., from attaching to clausal to phrasal constituents) is taken to be evidence for further grammaticalization. However, as Kaltenböck (2013:297) acknowledges, both the narrowing (e.g., Lehmann 1995) and widening (e.g., Traugott 1995:3) of scope has been found to co-occur with grammaticalization. If both results can hold in cases of grammaticalization, than neither can be used as a diagnostic for the process.<sup>16</sup> Lastly, although Kaltenböck (2013:298) cites a number of researchers (most prominently Kärkkäinen 2003) who have argued that *I think* is shifting away from expressing “epistemic qualification” to expressing more textual/discourse-organizational functions, the example he provides of the pragmatic marker being used as a “turn-taking signal (rather than an epistemic qualifier)” does not clearly show either a lack of epistemic uncertainty nor definitive use as a turn-taking signal. The example is reproduced in (13).

---

<sup>15</sup>Kaltenböck (2013:294) does present a figure labeled “Relative change of *I think* within the set of [EPs] overall”, which sounds like it could treat the EPs as a variable, however in the prose this chart is described as showing that “[i]n relative terms, i.e., compared to the other positions within the set, the proportion of final uses has also decreased significantly.” This analysis is still restricted to *I think* alone and does not calculate the frequency of *I think* relative to other variants.

<sup>16</sup>Kaltenböck (2013:297) does suggest that perhaps we are not dealing with changes to the syntactic scope of *I think* but rather with the semantic-pragmatic scope, and as such an increase in use with phrasal constituents might be more like Himmelmann’s (2004) context generalization. It is difficult to see how this is not a syntactic phenomenon, even if the semantic-pragmatic scope is also affected.

(13) A: Yes Well *I think*

B: <unclear words> was there that's the wonderful thing about it but it's an awfully long time ago

A: *I think* he *I think* the reason is that it's neither very alternative and free range like Bryanston, and nor is it too stiff like Rugby

(ICE-GB:S1A-054#99)

Kaltenböck (2013:298) argues that the turn-taking function is clear, evidenced by "A's repeated attempt to gain the floor." An alternative interpretation is that B interrupts A's first turn; A tries again with the same utterance, beginning with '*I think*', hesitates at 'he'; A then repairs his utterance, beginning again with *I think*, seemingly expressing a degree of commitment to the proposition beginning with 'the reason...'. Since Kaltenböck (2013:298) provides no metric with which to evaluate the pragmatic functioning of *I think*, this interpretation seems as likely as one in which *I think* is functioning as a turn-taking device. Thus, while EPs may be changing and continuing to grammaticalize, Kaltenböck's (2013) evidence is insufficient.

The one variationist study that has highlighted the variation in the realization of the verb of EPs and examined their continuing grammaticalization is Rodríguez Louro and Harris' (2013) study of Australian English. Rodríguez Louro and Harris' (2013) focus on the verb *reckon*, a variant that has received little attention in the literature. Following the terminology of Thompson (2002), they consider both epistemic(/doxastic) verbs and evidential verbs, which include *find* and *be said*, to determine where in this functional space *reckon* lies. They perform two separate analyses: one with all epistemic/evidential verbs, regardless of the subject, and one with only first person singular pronominal subjects. Together, they find that *think* is the overwhelming majority verb accounting for sixty-five percent of all epistemic/evidential verbs (whatever the subject) and seventy-two percent of tokens with first person,

Table 4.1: Rodríguez Louro and Harris' (2013:432) predictions for the grammaticalization of EPs. Reproduction of their Table 6

Factor group	Prediction
Syntactic position	The more grammaticalized the epistemic/evidential parenthetical, the greater its syntactic mobility and the more likely it will occur clause-medially or clause-finally.
<i>That</i> or <i>zero</i>	The more formulaic the epistemic/evidential parenthetical, the more likely it will occur with zero complementizer.
Main clause adverbials	The more grammaticalized the epistemic/evidential parenthetical, the less likely adverbials will intervene between the main clause subject and verb.

singular pronominal subjects.<sup>17</sup> Treating the set of epistemic/evidential verb constructions as a sociolinguistic variable, Rodríguez Louro and Harris (2013:431) set out, in this later analysis, to determine “the extent of *I think*, *I guess*, *I believe*, and *I reckon* grammaticalization [...] that is, the extent to which these first-person [epistemic/evidential complement taking predicates] have become epistemic/evidential parentheticals.” They make a number of quantitative predictions, based on much of the literature discussed above, to test for grammaticalization. These predictions are given in Table 4.1.

Despite their stated goal of determining the extent to which the top five EPs have grammaticalized, the low Ns for all variants except for *I think* mean that it is difficult to tease apart the constraints significantly favouring any one of these variants and the constraints disfavouring *think* in a variable rule analysis. Thus, I concentrate my discussion of their results on their statistical model of *I think* (versus all other variants) (Rodríguez Louro and Harris 2013:434). This model yields three important results. First, *I think* is favoured in tokens without an overt *that* complementizer. Given Rodríguez Louro and Harris' (2013:432) second prediction, this result is consistent with *I*

<sup>17</sup>Of just the epistemic(/doxastic), non-factive verbs which this chapter is concerned with, *think* represents seventy-seven and eighty-two percent in each analysis respectively. Focussing just on the first person cases, *I guess* is the next most frequent variant at 5.3 percent. *I believe* is third with 3.5 percent, *I suppose* is at 2.8 percent and *I reckon* represents a mere 1.8 percent of tokens.

*think* being more formulaic (cf. Torres Cacoullos and Walker 2009a). Second, *I think* is favoured in clause-initial position, suggesting that this form is less grammaticalized. This seems unexpected given the fact that the vast majority of the literature identifies *I think* as the most advanced EP along a cline of grammaticalization. However, this result might be misleading, since a full 77.1 percent of non-clause-initial EPs are realized as *I think* (Rodríguez Louro and Harris 2013:435). Furthermore, the favouring of clause-initial position might be confirmation of Kaltenböck's (2013) suggestion that *I think* has grammaticalized further from an epistemic qualifier to being used with textual/discourse functions, since these later functions of *I think* never occur clause-finally.<sup>18</sup> Lastly, the presence of an adverbial between the subject and the verb of an EP does not significantly constrain the realization of *I think*. Rodríguez Louro and Harris (2013:435–6) do note that despite the fact that a high frequency of tokens with main clause adverbials are *I think*, thirty percent of the time the adverbial is *just*. They suggest that *I just think* may be developing into a “formulaic construction in [Australian English]” (cf. Kaltenböck 2013:287).<sup>19</sup> As such, Rodríguez Louro and Harris (2013:438) conclude that *I think* exhibits the hallmarks of a grammaticalization of EPs (as in Table 4.1): syntactic mobility, a lack of overt complementizer, and little intervening material (given the stipulation about *I just think*).

However, much like in the previous chapter, it is difficult (or perhaps impossible) to determine the extent of grammaticalization without a diachronic component, which Rodríguez Louro and Harris (2013) lack. If *I think* is continuing to grammaticalize, we should see evidence for its increasing grammaticalization through time. That is, we should be able to find evidence for the predictions that Rodríguez Louro and Harris propose, as well as other diagnostics of grammaticalization, across time.

<sup>18</sup>Although Rodríguez Louro and Harris (2013:438–9) do not connect this result with Kaltenböck's (2013) hypothesis, they note that Mullan (2010) has observed that *I think* “is predominantly used to express what she labels an ‘organizational function’ [...] [which] occurs in turn initial position.”

<sup>19</sup>This conclusion is subject to the same caveat discussed above: it is not distinguished from any potential independent increase in the frequency of *just* (see Tagliamonte 2005).

This is the point of departure for this chapter. Using the combination of the *Earlier Ontario English* data and the *Toronto English Archive*, I track the EP variable system in Ontario English across the twentieth century in both real and apparent time in order to assess the diachronic evidence for these developments.

### 4.3 Variable Context

As I have done above in Chapter 3 with GEs, I treat the set of EPs as a variable system. As such, a variable context must be delineated (Tagliamonte 2006a:70). As with GEs, I circumscribe the variable context in terms of an amalgam of structural and functional properties (see Sankoff et al. 1978; Dubois 1992; Tagliamonte and Denis 2010; Pichler 2010; Denis and Tagliamonte 2014a). For the purposes of this chapter, EPs must conform to both a functional definition and structural definition. Functionally, EPs are pragmatic markers which express a degree of speaker commitment to a proposition. Following a strict functional definition, the variable context might include modal adverbials such as *maybe*, *possibly*, and *probably*, which speakers also use to decrease their commitment to the truth of propositions.<sup>20</sup> However, these adverbials fall outside of the present structural definition of EPs. This structural schema used to delimit EPs is presented in (14).

All EPs consist, minimally, of a first person, singular pronoun and a non-factive, epistemic/doxastic verb.<sup>21</sup> A few EPs can optionally occur with a modal between

<sup>20</sup>Östman (1982:153) excludes EPs and modal adverbials such as *maybe* from the core set of pragmatic markers (particles for him). However, he lumps them together as “modal particles”. Interestingly, *I guess* is included in his set of core pragmatic particles, while *I suppose* and *I believe* are listed as peripheral, where peripherality is defined in terms of a lack of clear separation between pragmatic function and propositional meaning. As Brinton (1996:31) points out, there is “little agreement among scholars about the inventory of forms to be included in the category of pragmatic markers” and other authors consider all EPs to be (core) pragmatic markers, while others still do not consider EPs to be (at least core) pragmatic markers at all.

<sup>21</sup>Unlike Rodríguez Louro and Harris (2013), who first examine EPs with the full range of subjects (both pronominal and nominal) and then focus in on just the first person singular EPs, I only consider first person singular EPs. That said, the variable context for this study and for Rodríguez Louro and Harris’s (2013:422) more focused analysis are essentially the same, though I do not consider



the subject and the main verb, as in (15). In EOE, these include *think*, *imagine*, and *say*, and, along with these three, *guess* in TEA. This modal is typically *would* (but sometimes *could*, *can*, and *should*). There are three EP types that are exceptions to the optionality of modals. When used in an EP, the main verb *say* requires a modal, whereas *recall* disallows modals. EPs that contain either bare present tense verbs or modal auxiliaries are both included in the variable context.

(14)

<i>I</i>	<i>(would)</i>	<i>(neg)</i>	}	<b>think</b>
				<b>suppose</b>
				<b>guess</b>
				<b>believe</b>
				<b>imagine</b>
				<b>figure</b>
				<b>feel</b>
				<b>gather</b>
				<b>expect</b>
				<b>suspect</b>
	<i>I</i>	<i>would</i>	<i>(neg)</i>	<b>say</b>
<i>as</i>	<i>I</i>			<b>recall</b>
	<i>it</i>	<i>(would)</i>	<i>(neg)</i>	<b>seems to me</b>

(15) a. Oh *I would think* we went once a week mh-mm.

(EON/F/1914)

b. Well I imagine I was six, *I would imagine* I was six.

(EON/F/1914)

---

Thompson's (2002) 'evidential' verbs such as *find*.

- c. There were a lot of people coming from, you know, houses where *I'd guess* they were making probably double in, you know, family income from what my folks were making.

(TOR/F/1976)

All EPs can also be negated as in (16), with the exception of EPs with the main verb *recall* (\**as I don't recall*). Although some dialects of English allow for *guess* EPs to be negated, as shown in (17) from an oral history of a North Carolinian speaker<sup>22</sup> as we will see in §4.4.2, there are no examples of *I don't guess* in EOE or TEA. Both affirmative and negated EPs are included in the variable context.

- (16) a. We had guinea hens. Now that's something nobody raises anymore, *I don't think*.

(NIA/F/1904)

- b. Oh *I don't suppose* there were more than a dozen apple trees.

(NIA/F/1904)

- c. *I don't expect* he had too much money to bank as far as that went.

(EON/F/1906)

- (17) *I don't guess* you ever saw a knotted countrypin.

(Blue Ridge Parkway Folklife Project Collection (American Folklore Center  
1982/009))

Negated EPs, particularly in initial position where they may be functioning as matrix verbs, are considered to involve neg-raising, in which the negative marker of the subordinate clause is raised to the matrix verb but the negation is still interpreted as

<sup>22</sup>The full recording and transcript is available from the Library of Congress here <<http://www.loc.gov/item/qltooooo9>>.

applying to the subordinate clause (Israel 2004:709). For example, (16c), implies (18).

(18) I expect he didn't have too much money...

Though the neg-raised (16c) and non-neg-raised (18) are semantically equivalent, the neg-raised paraphrase is sometimes considered to express less commitment to a proposition than non-neg-raised variants (see Kearns 2007:486; Rodríguez Louro and Harris 2013).<sup>23</sup>

Consider the contrast between (19a) and its neg-raised paraphrase in (19b). This is part and parcel of a general phenomenon discussed by Horn (1989) of “contrary-negation-in-contradictory-clothing.” Israel (2004:708) describes the phenomenon as taking “the form of an inference from a formally contradictory negation **not-p** to a strong contrary assertion **q**, effectively ignoring the logical possibility of something being neither **p** nor **q**.” In other words, negating something that is positively evaluated (e.g., *He's not nice*) can mitigate something that is negatively evaluated (e.g., *He's mean*).<sup>24</sup> I will return to the strength of commitment of negated EPs at the end of §4.4.4 below.

(19) a. *I think they wouldn't have made a very good jam even. They were too sweet.*

(NIA/F/1916)

---

<sup>23</sup>Neg-raising seems to be a property of EPs only. True main clauses do not exhibit semantically vacuous raising of the negation in an embedded clause. Consider the contrast between i. and ii. below.

- i. a. I think he didn't have a lot of money. ≈  
 b. I don't think he had a lot of money.
- ii. a. I realized he didn't have a lot of money. ≠  
 b. I didn't realize he had a lot of money.

In i, the two sentences are semantically equivalent, while in ii. the meanings are distinct.

<sup>24</sup>See Israel (2004:708–9) for further discussion and examples of the pragmatics of contrary negation.

- b. *I don't think* they would have made a very good jam even. They were too sweet.

Lastly, as observed in (1), an EP can occur prior to its complement proposition (initial or root position), after its complement proposition (final position), or within its complement proposition (medial position). In this chapter, I consider all EPs in each of these potential positions.<sup>25</sup>

## 4.4 Tracking Changes Over the 20<sup>th</sup> Century

I now turn to a quantitative analysis of the constraints on the verb in EPs, concentrating on factors that have been implicated in the grammaticalization of these pragmatic markers. I examine the distribution of variants across several sociolinguistic factors and then assess the significance of these factors using mixed-effects logistic regression modelling.

### 4.4.1 Overall Distribution in Ontario Across Time

In total, 4323 EP tokens were extracted and coded from the three corpora.<sup>26</sup> I begin by presenting the overall distribution of variants across the four communities: Belleville, Eastern Ontario, Niagara, and Toronto in Figure 4.1 and Table 4.2.

The majority of the literature has identified *I think* as the most frequent EP in a wide range of twentieth century, national varieties including American English (Thompson and Mulac 1991; Kaltenböck 2013), British English (Thompson and Mulac

<sup>25</sup>Only initial tokens are included from the variable context of *that*/ $\emptyset$  variation analyses (Tagliamonte and Smith 2005:298; Torres Cacoullos and Walker 2009a:11).

<sup>26</sup>My procedure for extraction of EPs differed slightly for the two corpora. In EOE, EPs were hand tagged by eye in every interview. In total, 1063 tokens were extracted. For TEA, every variant that was found in EOE, all possible gaps given the range of possibilities discussed above, and any other variant that has been reported in the literature were searched for systematically using *AntConc* (Anthony 2011) and hand filtered to remove non-EPs. In total, 3260 tokens were extracted. The final set of data was coded by hand for the various factors discussed.

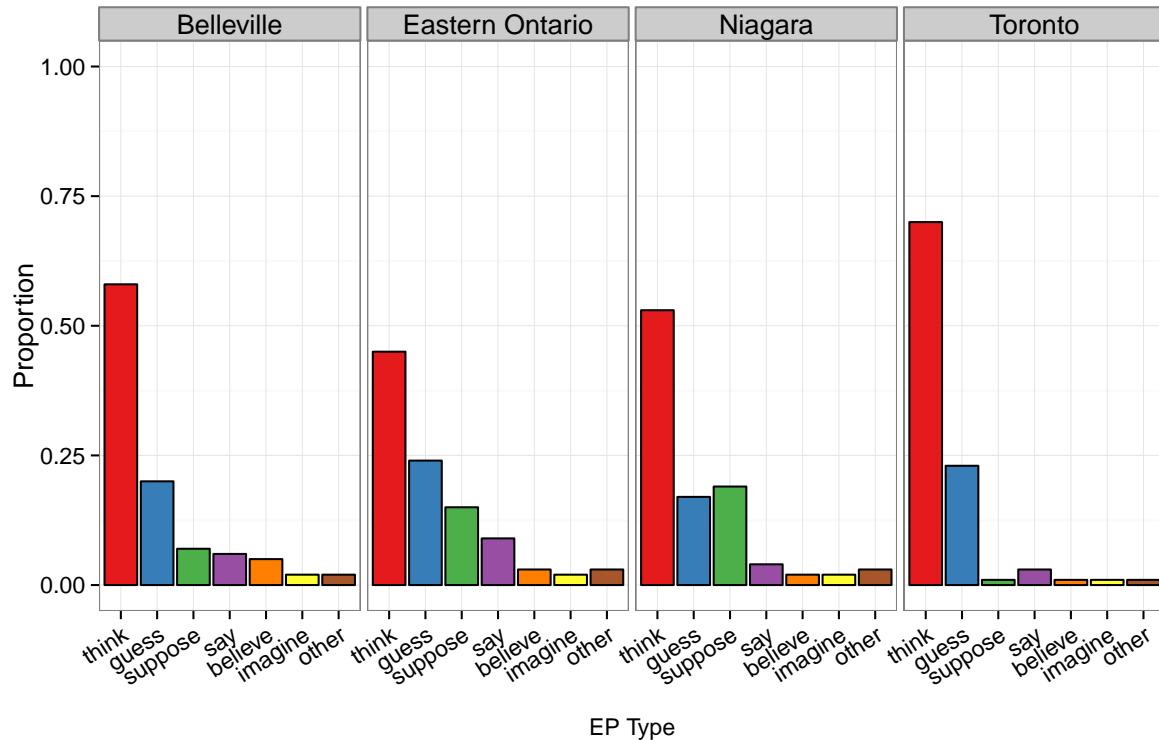


Figure 4.1: The overall distribution of EP variants in EOE and TEA, by community. Belleville N = 244; Eastern Ontario N = 349; Niagara N = 470; Toronto N = 3260.

1991; Tagliamonte and Smith 2005; Kaltenböck 2013), Australian English (Rodríguez Louro and Harris 2013), and Canadian English (Torres Cacoullos and Walker 2009a; Blondeau and Nagy 2008). The predominance of *I think* is clear in each of the four EOE/TEA communities. In all but one community, *I think* represents over half of the EP tokens, and even in Eastern Ontario, *I think* is the most common form representing 44.7 percent of tokens. Although *I think* is the dominant variant, even in Toronto where it is most frequent, the form does not reach the extreme frequencies reported elsewhere: 82 percent in Australia (Rodríguez Louro and Harris 2013:433)<sup>27</sup>; 80 percent in the *London-Lund Corpus*, 84 percent in ICE-GB, and 75 percent in the *Corpus of London Teenage Language* (Kaltenböck 2013:290, fn. 2).

Of the other variants, only *I guess* and *I suppose* ever occur more than 10 percent

<sup>27</sup>This frequency is adjusted from their table to included only epistemic verbs and leave out evidential verbs (e.g., *hear*, *sense* (see Rodríguez Louro and Harris 2013:418).

Table 4.2: The overall distribution of EP variants in EOE and TEA, by community

	BLV		EON		NIA		TOR	
	%	N	%	N	%	N	%	N
<i>think</i>	58.2	142	44.7	156	53.2	250	70.2	2287
<i>guess</i>	20.5	50	24.4	85	17.2	81	23.3	761
<i>suppose</i>	7.0	17	14.9	52	18.5	87	0.6	20
<i>say</i>	5.7	14	8.6	30	4.3	20	3.1	101
<i>believe</i>	4.9	12	2.6	9	1.7	8	0.9	30
<i>imagine</i>	2.0	5	2.3	8	1.7	8	0.5	17
other	1.6	4	2.6	9	3.4	16	1.3	44
TOTAL		244		349		470		3260

of the time. *I guess* is the second most frequent form in all communities except Niagara, where *I suppose* marginally surpasses it (17.2 vs. 18.5 percent). In the three EOE communities *I suppose* makes up a sizeable portion of the variation, but in Toronto, the form is peripheral along with *I would say*, *I believe*, *I imagine*, and the other less frequent variants. Given that Toronto represents a thirty (to forty) year real-time advance in relation to the EOE communities (and a wider apparent-time range, including much younger speakers than EOE), the contrast between Toronto on the one hand and Belleville, Eastern Ontario, and Niagara on the other, suggests that a change took place across the twentieth century. The variant *I suppose* has fallen in frequency while *I think* (and perhaps *I guess*) have risen.

Consider now Figure 4.2 which plots the frequency of *I think*, *I guess*, *I suppose* and all other EPs across apparent time, collapsing the four communities. Speakers are grouped into the decade of their birth, except for those born before 1890, who are grouped with those born in 1890 as the Ns before 1890 are small.

As hypothesized earlier, *I think*, in red, has risen in frequency across the twentieth century, though after the 1930s stabilizes at around 75 percent of the variation. In the same period, *I suppose* in green obsolesces from a maximum of 20 percent (and the second most frequent form) for those born in the nineteenth century to zero occurrences for speakers born in 1990. This obsolescence is consistent other than

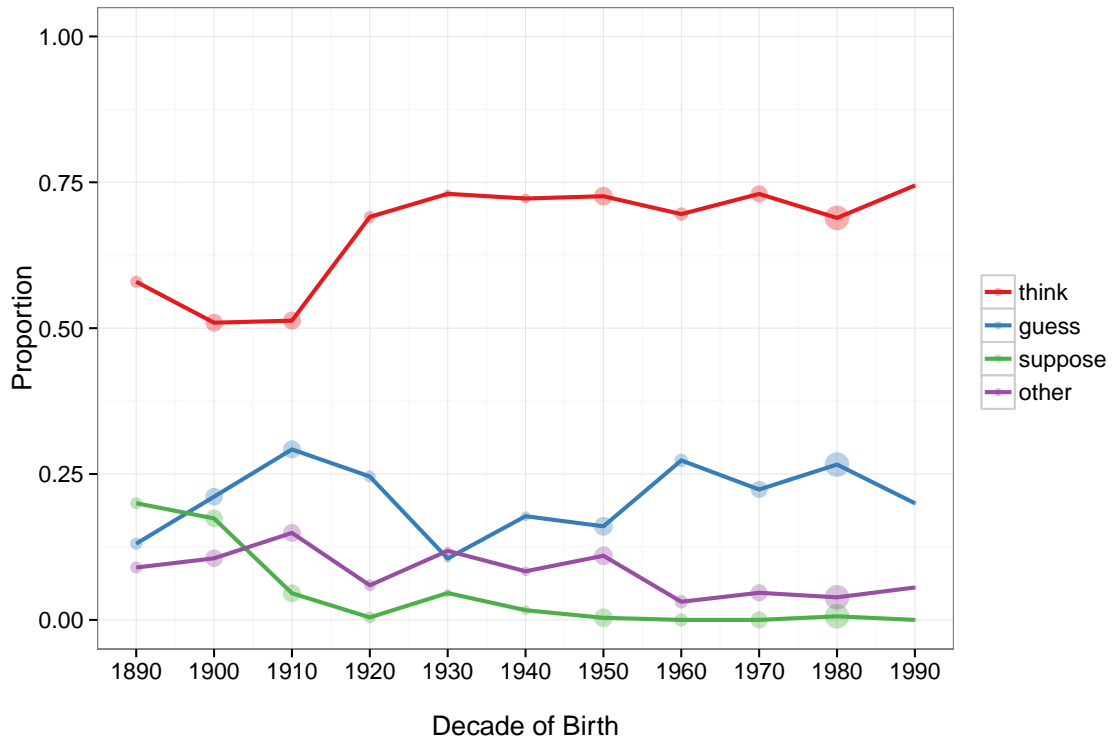


Figure 4.2: Distribution of EP variants in EOE in apparent time. Size of data point represents  $N$ s in that decade. *Think*  $N = 2835$ ; *guess*  $N = 977$ ; *suppose*  $N = 470$ ; other  $N = 335$ .

a small bump in frequency in the 1930s. The trajectory of *I guess* is more jagged. Although the initial obsolescence of *I suppose* seems to be at the expense of an increase of *I guess* rather than *I think*, once *I think* increases in earnest, *I guess* briefly recedes. But, after 1930, once *I think* stabilizes and while *I suppose* and the other variants decrease, *I guess* rises once again. For speakers born in and after 1960, the use of EPs is essentially a dichotomous variable.

Although many have argued that changes in the frequency of forms can be indicative of grammaticalization (Thompson and Mulac 1991; Bybee 2003; Mair 2004; Torres Cacoullos and Walker 2009a), by considering the frequency of these variants alone, it is unclear what role grammaticalization might have in these changes. Rather, in order to examine the potential role of grammaticalization in the development of EPs across twentieth century Ontario, we must operationalize factors that have been

implicated in the development of these pragmatic markers as a series of linguistic constraints (Poplack and Tagliamonte 2001:225). I now turn to analyses of three such constraints: intervening material, syntactic position, and epistemic/doxastic strength.

#### 4.4.2 Intervening Material: Negation, Modals, and Adverbials

Several researchers have implicated the presence or, more accurately, the absence of intervening material between the subject and verb of EPs as indicating a more advanced level of grammaticalization. The less frequently EPs occur with intervening material is taken as evidence for a range of related principles and parameters of grammaticalization, all of which involve the fusion or bonding of morphemes. Lehmann (1982[1995]:147) identifies the loss of word boundaries as the processes of COALESCENCE and FUSION, while Hopper (1991:22) and Heine (2003:579) discuss DECATEGORIALIZATION, which involves (among other changes to the morphosyntactic properties of linguistic elements) “the loss of independent word status” (Heine 2003:579).<sup>28</sup> Such decategorialization is exactly what Thompson and Mulac (1991:325) argue is happening with EPs: “[t]he lexical category of the erstwhile combination of canonical Noun+Verb can best be characterized as the secondary category ‘Adverb’.”<sup>29</sup>

Thus, if EPs are grammaticalizing, we should observe a decrease in the frequency of intervening material between subjects and verbs. There are three kinds of intervening material that occur in the Ontario data: negation, modals, and adverbials, as in (20). Each of these contexts will be examined in turn.

- (20) a. It’s not fair for the little boy *I don’t think*.

(TOR/F/1980)

- b. Well *I’d think* around about, around ten maybe, something like that.

<sup>28</sup>Brinton and Traugott (2005:110) observe that the bonding of a syntactic phrase into a single morpheme is characteristic of both grammaticalization and lexicalization.

<sup>29</sup>Brinton (1996:253) makes a similar observation.



(NIA/F/1911)

c. *I just think* they didn't plan ahead.

(TOR/F/1988)

In Table 4.3, I present the frequency of negated tokens in each of the three most frequent EP types (and all other EPs). Frequency is shown here 'sideways', that is, the percentage of negated EPs is relative to the total number of tokens of that general type. For example, in EOE, one-quarter of the type *I think* occurred with a negative marker (i.e., *I don't think*).

Table 4.3: Frequency of negated EPs in four general types over four age groups.

	EOE		TOR>50		TOR <sub>30-50</sub>		TOR< 30	
	%	N	%	N	%	N	%	N
<i>I think</i>	24.8	136/548	11.6	83/717	13.9	65/468	13.7	151/1102
<i>I guess</i>	0.0	0/216	0.0	0/237	0.0	0/105	0.0	0/419
<i>I suppose</i>	9.6	15/156	7.7	1/13	–	0/0	0.0	0/7
Other	14.0	20/143	4.8	4/83	6.5	3/46	3.2	2/63
Total	16.1	171/1063	8.4	88/1050	10.9	68/619	9.6	153/1591

Consider the *I think* EPs first. In the oldest age group, 24.8 percent of tokens are negative. In the next three age groups, this frequency drops by more than ten percent. As noted earlier, there are no negated tokens of *I guess*. For *I suppose* and the other EPs, the frequency of negated tokens declines across time.<sup>30</sup> Underlying this distribution is a decrease in negated EPs generally, at least between EOE and TEA.<sup>31</sup> Taken together, the distribution of types of negated EPs across the twentieth century has reduced in real time. This decrease in intervening material between the subject and verb of an EP might be evidence for the increasing fusion/decategorialization of *all* EPs. However, it could also be indicative of a change toward less negative-raising generally. Alternatively, since there is no significant difference across the

<sup>30</sup>This difference is significant according to a  $\chi^2$  test:  $\chi^2 = 33.9, df = 3, p = 2.1 \times 10^{-7}$ .

<sup>31</sup>The three Toronto age groups are not significantly different from each other. The overall frequency of negated EPs in Toronto is 9.5 percent.

three Toronto age groups,<sup>32</sup> and thus no evidence for such decategorialization in apparent time within that community, perhaps the difference between Toronto and the three other communities does not represent change in real time, but rather a dialect difference.

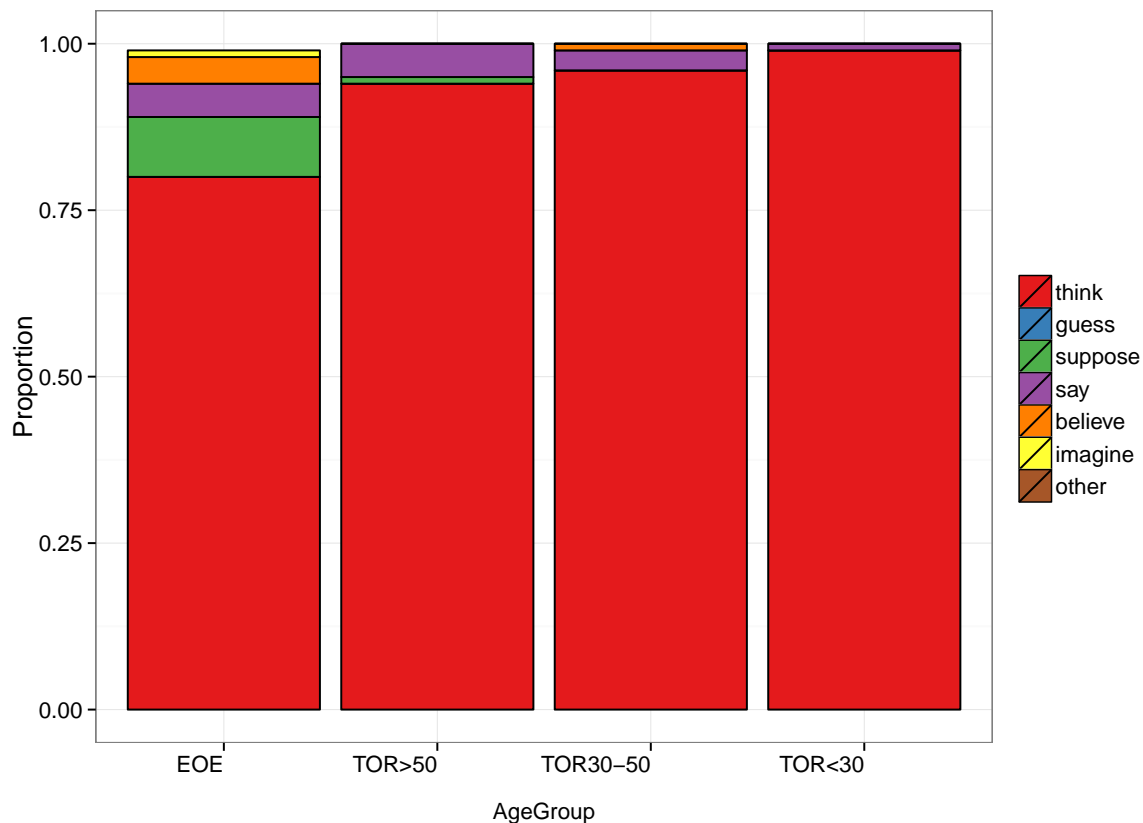


Figure 4.3: Distribution of negated EP variants in four age groups. EOE N = 171; TOR>50 N = 88; TOR30-50 N = 68; TOR<30 = 153.

Figure 4.3 shows evidence for a different kind of change. In EOE, six different EP types were negated: *I think*, *I suppose*, *I believe*, *I would say*, *I imagine*, and *I expect*. In the two middle age groups, three variants occur in the negative (*I think*, *I suppose*, and *I would say* for TOR>50 and *I think*, *I would say*, and *I believe* for TOR30-50). In the youngest age group, other than two tokens of *I wouldn't say*, the only negated EP is *I don't think*. Thus, across the twentieth century, it appears that *I don't think* has been

<sup>32</sup> $\chi^2 = 1.6, df = 2, p = 0.44$

recruited by speakers to serve as the one negative EP. One could interpret this as an instance of Hopper's (1991:22) principle of SPECIALIZATION, the decrease of options as linguistic items grammaticalize. However, the apparent specialization of *I don't think* takes place at the same time that the only other EPs that are able to be negated (e.g., *I suppose*, *I believe*, *I would say*) generally decline (see Figure 4.2). If the only available variants for the youngest generation are *I think* and *I guess* and *I guess* is not able to be negated, it is unsurprising that *I don't think* has become the only negative EP.

Table 4.4 shows the frequency of EPs with respect to the presence of modal auxiliaries. Unlike Table 4.3, here the general type *I (would) say* is excluded, as this form always requires a modal. Otherwise, the frequencies are tabulated in the same manner as above.

Table 4.4: Frequency of modal EPs in four general types over four age groups.

	EOE		TOR>50		TOR <sub>30-50</sub>		TOR< 30	
	%	N	%	N	%	N	%	N
<i>I think</i>	1.8	10/548	0.0	0/717	0.0	0/468	0.2	2/1102
<i>I guess</i>	0.0	0/216	0.0	0/237	0.0	0/105	0.5	2/419
<i>I suppose</i>	0.0	0/156	0.0	0/13	–	0/0	0.0	0/7
Other	4.0	3/75	6.5	3/46	27.3	3/11	3.0	1/33
Total	1.3	13/1001	0.3	3/1012	0.5	3/584	0.3	5/1557

There are far fewer tokens of EPs with modal auxiliaries (N = 24) than negated EPs (N = 480), so these patterns must be interpreted with caution. The presence of a modal within an *I think* EP has significantly declined over real time, from 1.8 percent of all *I think* tokens in EOE to 0.2 percent in the youngest age group in Toronto. This is suggestive of increased fusion/decategorialization of *I think*. In the oldest age groups, *I guess* and *I suppose* never occur with an intervening modal. However, in the youngest generation, there are two tokens of *I would guess*.<sup>33</sup> *I suppose* never occurs with a modal. The only EP type represented in the other category is *I imagine*.

<sup>33</sup>These two tokens come from two different speakers.

Although there is a larger peak in the TOR<sub>30–50</sub> age group, there are only eleven tokens of EPs other than *I think*, *I guess*, and *I would say* and thus, this age group is not significantly different from the other age groups.

Lastly, consider Table 4.5, which presents the frequency of EPs with respect to the presence of an intervening adverbial.

Table 4.5: Frequency of EPs with intervening adverbials in four general types over four age groups.

	EOE		TOR <sub>&gt;50</sub>		TOR <sub>30–50</sub>		TOR <sub>&lt;30</sub>	
	%	N	%	N	%	N	%	N
<i>I think</i>	0.7	4/548	0.1	1/717	0.2	1/468	0.7	8/1102
<i>I guess</i>	0.0	0/216	0.0	0/237	0.0	0/105	0.0	0/419
<i>I suppose</i>	0.0	0/156	0.0	0/13	–	0/0	0.0	0/7
Other	0.7	1/143	0.0	0/83	0.0	0/46	0.0	0/63
Total	0.5	5/1063	0.1	1/1050	0.2	1/619	0.5	8/1591

Very little can be gleaned from this table about fusion/decategorialization of EPs. Regardless, there are three important observations: 1) intervening adverbials are generally rare in Ontario English, 2) with one exception, only *I think* occurs with intervening material,<sup>34</sup> and 3) according to a  $\chi^2$  test, there is no significant change in the frequency of intervening material within *I think* across the twentieth century.<sup>35</sup> Although Kaltenböck (2013) and Rodríguez Louro and Harris (2013) suggest that *I just think* might be developing into a formulaic expression itself in British and Australian English, there are only four tokens in the youngest age group in Toronto.

<sup>34</sup>There is one token of *I don't really believe* in EOE, shown in (i).

- i. So what I'm getting at then, is that *I don't really believe* that any level of government is going to ah drop this program.

(BLV/M/1914)

This token co-occurs with an overt complementizer a likely indication that it is a main clause, rather than an EP. Indeed, that negative polarity items, such as *any* here, are licensed in the embedded clause may indicate that negated examples like these are not parenthetical.

<sup>35</sup> $\chi^2 = 4.37, df = 3, p = 0.22$

**Summary: Intervening Material**

In sum, the evidence from intervening material is scant and mixed. Although there is a decrease in the frequency of a negative marker intervening within *I think* over time, there is also a general decrease in negated EPs. There is potential specialization of *I don't think*, but this might be an artifact of the changing variable system. The frequency of intervening modals within *I think* has decreased, but even in the earliest age group, less than two percent of *I think* tokens included a modal. Perhaps surprisingly, what seems to be a strictly fused form, *I guess*, is used with a modal by two different speakers in the youngest age group. Lastly, *I think* occurs with the same frequency of intervening adverbials comparing across the oldest and youngest speakers. Although there is some evidence for increased fusion/decategorialization over time, some of this evidence could be interpreted in other ways and there is also some counter-evidence.

Next, I turn to a factor that has been consistently implicated in the grammaticalization of EPs: the syntactic position of the EP.

**4.4.3 Syntactic Position**

The syntactic position of EPs has been used as a diagnostic of grammaticalization by several researchers, as discussed above (*inter alia* Thompson and Mulac 1991; Aijmer 1997; Kaltenböck 2013; Rodríguez Louro and Harris 2013). As Rodríguez Louro and Harris (2013:424) put it, “[t]he looser the syntactic behaviour of [EPs], the more grammaticalized they are.” The general idea is that the more frequently an EP is used in non-clause-initial position (i.e., non-root or in any position that cannot be mistaken for a main clause), the farther along the cline of grammaticalization the EP is. Taking this hypothesis as a starting point, I test to see whether the variants that have been increasing over the twentieth century are becoming syntactically ‘looser,’

and thus, more grammaticalized.

### Coding Procedure

Following Rodríguez Louro and Harris (2013:424) and Kaltenböck (2013:293), I coded every EP token for its syntactic position: clause-initial, clause-medial, or clause-finally, as in (21).

- (21) a. *I suppose* we'd probably have half the farm in peaches.  
(NIA/M/1898)
- b. Took more [men] filling silo *I think* than it did thrashing.  
(NIA/M/1907)
- c. Strawberries was about middle of June *I guess*.  
(NIA/M/1902)

Many tokens were coded as being in clause-initial position, despite the EP not being the first element of the sentence. Several linguistic elements can occur in a syntactically higher position than a clause-initial EP including discourse/pragmatic markers (*so, well, anyway however, like, you know, you see*), hesitation markers (*uh* and *um*), conjunctions (*and, or, because*), adverbials (*now, then, actually*), and attention/acknowledgment/agreement markers (*hey, oh, yes, no*) (Tagliamonte 2014). Extraposed topics can also occur to the left of an EP. The presence of a resumptive pronoun in subject position was the main diagnostic of whether or not an EP was in initial position but occurred after an extraposed topic, as in (22a), or if the EP was in medial position, as in (22b).<sup>36</sup>

- (22) a. The guy *I guess* he thought she had insulted him.

<sup>36</sup>In practice, if an EP could function as a main clause, it was coded as being in clause-initial position. This was determined by whether or not a *that* complementizer could occur after the EP or not.

(TOR/F/1979)

- b. My mother *I think* went to grade five or something.

(TOR/M/1952)

Although syntactically in initial position, EPs with an overt *that* complementizer, as in (23) were coded separately. These tokens have been excluded from the analysis, on the assumption that they are main clauses and not pragmatic markers.

- (23) *I think that* we were short players on my competitive team.

(TOR/F/1983)

In addition to appearing in clause-initial, clause-medial, and clause-final positions, several EPs occur as independent or semi-independent clauses with *so* as in (24). These EPs are typically used as a quick affirmative response to an interlocutor.<sup>37</sup> In cases with *so*, *so* functions as a propositional anaphor, referring back to the previous sentence.

- (24) a. A: Must have been pretty frightening.

B: Yeah. *I guess*.

(TOR/M/1952)

- b. A: Well, it'll take you an hour and twenty minutes.

B: Yeah, yeah. *I suppose so, yeah*.

(TOR/M/1984)

Independent EPs occurred 52 times across EOE and TEA and semi-independent EPs occurred 150 times. Figure 4.4 shows a conditional inference tree of these indepen-

<sup>37</sup>Note that in both these examples, the EP co-occurs with the affirmative marker *yeah*.

dent and semi-independent EPs.<sup>38</sup> Of the two main EP variants, *I think* is more likely to co-occur with *so*, while for speakers born after 1980, *I guess* is more likely to occur independently. These tokens will be left aside for the remainder of this chapter.

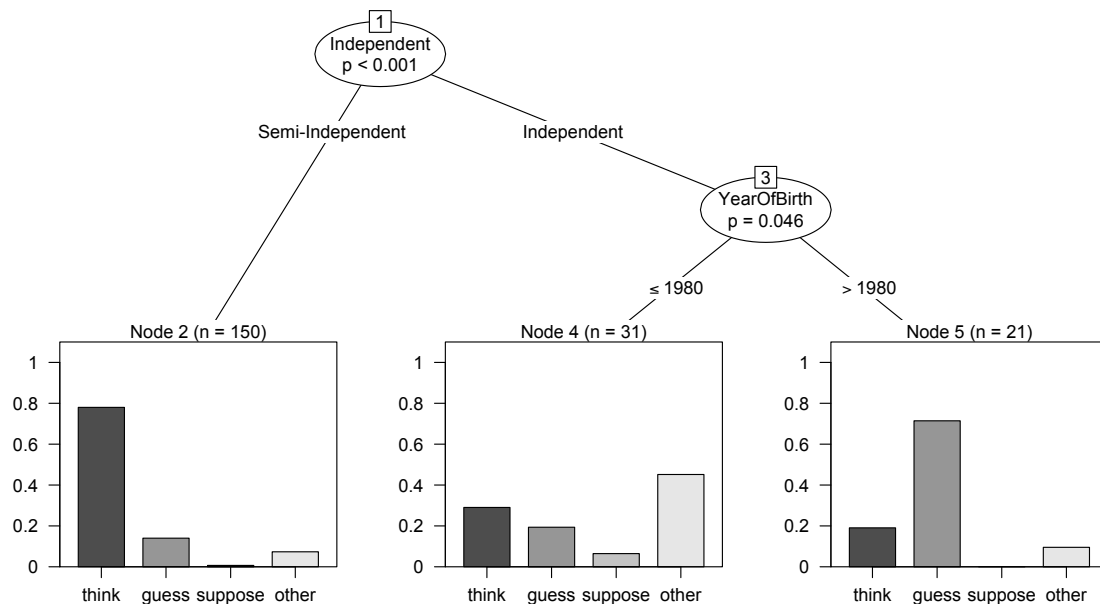


Figure 4.4: Conditional inference tree of (semi-)independent EPs and year of birth.

Any tokens that were surrounded by incomprehensible speech, were part of a false start, or were unclear for any other reason were coded as such and excluded from the analysis ( $N=99$ ).

## Distributions

Figure 4.5 presents the distribution of *I think*, *I guess*, *I suppose* and all other EP variants by the three main syntactic positions in EOE, collapsed across community, and TEA, separated into the three age groups used in Chapter 3.

Given that *I think* makes up the great majority of tokens, it is not surprising

<sup>38</sup>Conditional inference trees are a type of decision tree algorithm that uses a non-parametric test statistic to estimate “the likelihood of the value of the response variable [...] based on a series of binary questions about the values of predictor variables” (Tagliamonte and Baayen 2012:159). Conditional inferences trees are particularly useful for data exploration, particularly for spotting interactions between independent variables and for determining how to categorize (i.e., bin) continuous variables.



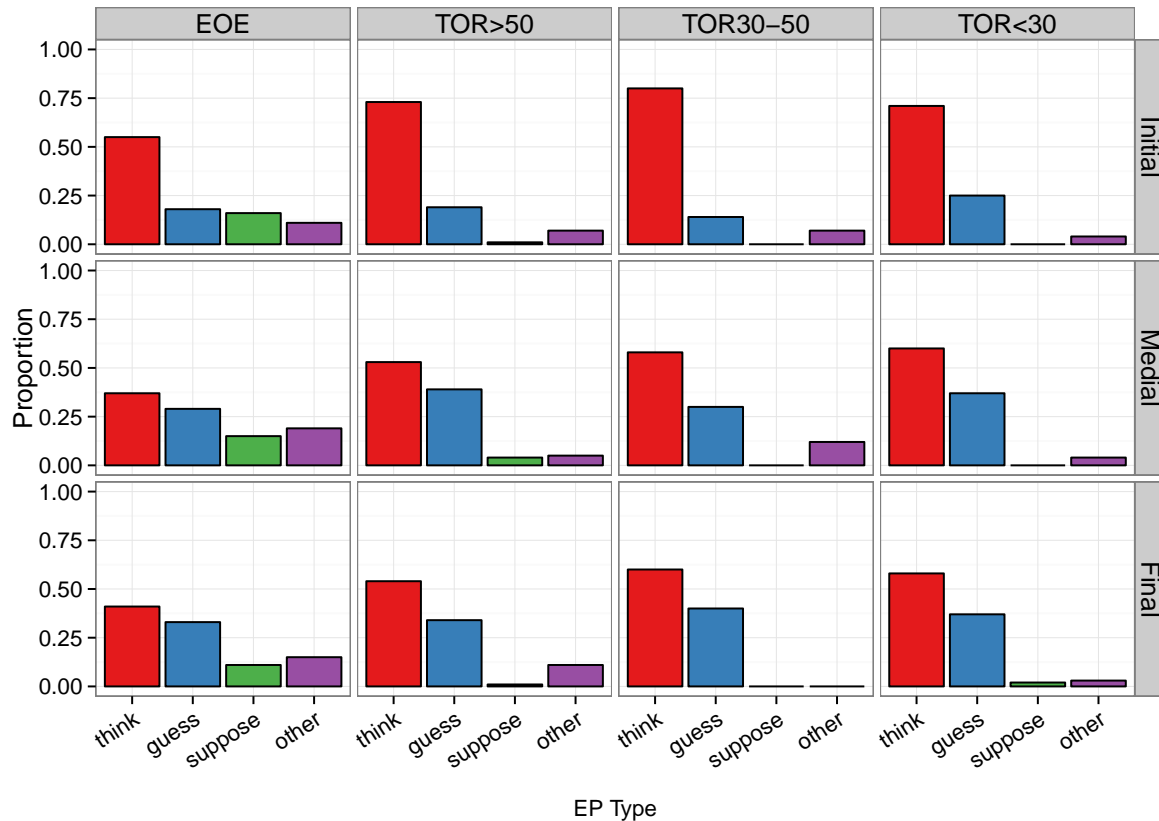


Figure 4.5: Distribution of EP variants by syntactic position in four age groups. Cell Ns: EOE, Initial = 723, Medial = 105, Final = 138; TOR>50, Initial = 716, Medial = 135, Final = 102; TOR30-50, Initial = 447, Medial = 76, Final = 40; TOR<30, Initial = 1145, Medial = 136, Final = 135.

that this variant is the most frequent variant in all positions and in all age groups. However, the distribution of variants does change across age groups and positions. In line with Figure 4.2, the frequency of *I think* (in red) increases in each subsequent age group. This increase is consistent across each position, with the exception of the youngest speakers, where *I think* is slightly less frequent in clause-initial and clause-final position than the middle aged speakers in Toronto. Again, in line with Figure 4.2, *I guess* (in blue) oscillates across age groups, but, in all positions, is consistently more frequent with the youngest speakers in Toronto than with the speakers in EOE. The decline of *I suppose* (in green) is also apparent across positions.

Comparing the distribution of different forms across positions, consistency across

age groups is the norm. For each age group, *I think* is more frequent in clause-initial position than clause-medially or clause-finally, while *I guess* is more frequent in the non-clause-initial positions than in clause-initial position.

Kaltenböck (2013:293–295) argues that there is a decrease of *I think* in clause-final position. What he observes is that the frequency of *I think* in clause-final position decreases relative to the frequency of *I think* in other positions. As I noted above, this is problematic because without knowing how frequently *I think* could have been used but was not (i.e., some other EP variant was used), it is not possible to tell if the constraint on position has changed. Figure 4.5 demonstrates that by considering *I think* relative to all other EP variants, the variant does not decrease in clause-final position. On the contrary, the form increases.

### Regression Analysis

How are these results relevant to the predictions about the grammaticalization of EPs? One of Rodríguez Louro and Harris's (2013) predictions is that more grammaticalized forms will appear in non-clause-initial positions. From a diachronic perspective, if grammaticalization is ongoing, we should see an increasing amount of syntactic mobility of grammaticalizing forms. In other words, EPs that are undergoing grammaticalization should exhibit an increased frequency in non-clause-initial positions and a decrease (or at least significantly slower increase) in frequency in clause-initial position. This is the same idea as the one discussed in §§3.4.1 and 3.4.2.

This hypothesis can be tested statistically. If there is a significant interaction between syntactic position and time on the realization of an EP as *I think* (for example), we can interpret this as an indication that *I think* has undergone grammaticalization over the twentieth century. We might expect this, given the rise in frequency of *I think*. If there is no interaction, then the rise of *I think* is not related to increasing grammaticalization of the form from a main clause to a (syntactically mobile) adver-

Table 4.6: Mixed-effects logistic regression testing the fixed effects of YEAR OF BIRTH (centered, continuous), POSITION (reference level = initial position), and their interaction and a random intercept for Speaker on the realization of EPs as *I think*. Treatment contrast coding. Coefficients reported in log-odds. Correlations of fixed effects,  $r < |0.25|$ . N = 3354

	Estimate	Std. Error	z-value	p-value	
Intercept	0.992	0.104	9.500	$<2.00 \times 10^{-16}$	***
YEAR OF BIRTH (centered)	0.013	0.003	4.110	$3.96 \times 10^{-5}$	***
POSITION(Non-initial)	-0.698	0.093	-7.482	$7.31 \times 10^{-14}$	***
YOB:POSITION	-0.002	0.003	-0.7482	$4.25 \times 10^{-01}$	
Random intercept:					
SPEAKER	Variance = 0.91, N = 127				

bial.

Table 4.6 presents the results of a mixed-effects logistic regression that tests the effects of syntactic position over time on the realization of EPs as *I think*. Given that the distributions of EPs in clause-medial and clause-final positions in Figure 4.5 are roughly identical, and there is no a priori reason to assume that either clause-medial position or clause-final position is more (or less) grammaticalized than the other, this model treats syntactic position as binary: initial vs. non-initial position.<sup>39</sup>

Table 4.7: Analysis of deviance,  $\chi^2$  test for model reported in Table 4.6.

	$\chi^2$	df	p-value	
Intercept	90.258	1	$<2.20 \times 10^{-16}$	***
YEAR OF BIRTH	16.889	1	$3.96 \times 10^{-5}$	***
POSITION	55.982	1	$7.31 \times 10^{-14}$	***
YOB:POSITION	0.636	1	$4.25 \times 10^{-01}$	

The model in Table 4.6 includes a significant main effect of YEAR OF BIRTH, confirming the increase of *I think* over the twentieth century. There is also a main effect of POSITION, such that *I think* is disfavoured in non-initial position (compared to initial position) as indicated by the negative coefficient. The interaction of these two main effects is not significant. An analysis of deviance, presented in Table 4.7, confirms

<sup>39</sup>This model excludes all independent EPs, EPs with an overt *that* complementizer, and all EPs with intervening material.

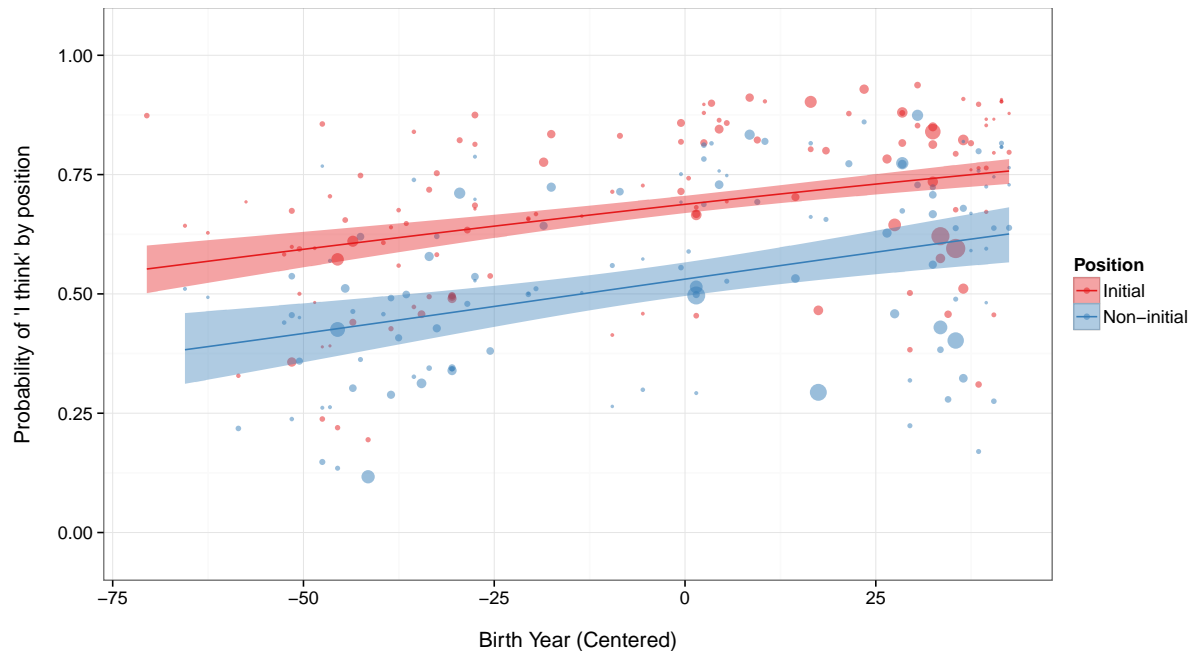


Figure 4.6: Probability of *I think* (vs. all other EPs) by syntactic position over apparent time. Fitted values from the model in Table 4.6. Dot size represents the number of tokens at those coordinates.

the significance of the main effects and indicates that the interaction does not add explanatory value to the model. The model is visualized in Figure 4.6 which plots the fitted values from the model and binomial curves for each level of the POSITION factor.

Figure 4.6 provides a visualization of the rise of *I think* across the twentieth century. Although *I think* is shown to favour initial position over non-initial position, the probability of *I think* increases at a constant rate in each position. Thus, despite changes in the frequency of *I think*, there is no evidence of concomitant increasing grammaticalization as diagnosed by syntactic position.

In Figure 4.2, *I guess* also shows some indication of increasing frequency. It is possible that this variant has increased its syntactic mobility. Table 4.8 shows the results of a mixed effects model, testing the same factors as Table 4.6 but on the realization of EPs as *I guess*, rather than *I think*.

Table 4.8 indicates that the main effect of YEAR OF BIRTH is not a significant pre-

Table 4.8: Mixed-effects logistic regression testing the fixed effects of YEAR OF BIRTH (centered, continuous), POSITION (reference level = initial position), and their interaction and a random intercept for Speaker on the realization of EPs as *I guess*. Treatment contrast coding. Coefficients reported in log-odds. Correlations of fixed effects,  $r < |0.27|$ . N = 3354

	Estimate	Std. Error	z-value	p-value	
Intercept	-1.490	0.110	-13.565	$<2.00 \times 10^{-16}$	***
YEAR OF BIRTH (centered)	-0.002	0.003	-0.709	$4.78 \times 10^{-01}$	
POSITION(Non-initial)	0.739	0.097	7.615	$2.64 \times 10^{-14}$	***
YOB:POSITION	-0.002	0.003	0.625	$5.32 \times 10^{-01}$	
Random intercept:					
SPEAKER	Variance = 0.97, N = 127				

Table 4.9: Analysis of deviance,  $\chi^2$  test for model reported in Table 4.8.

	$\chi^2$	df	p-value	
Intercept	184.009	1	$<2.20 \times 10^{-16}$	***
YEAR OF BIRTH	0.503	1	$4.78 \times 10^{-01}$	
POSITION	57.985	1	$2.64 \times 10^{-14}$	***
YOB:POSITION	0.390	1	$5.32 \times 10^{-01}$	

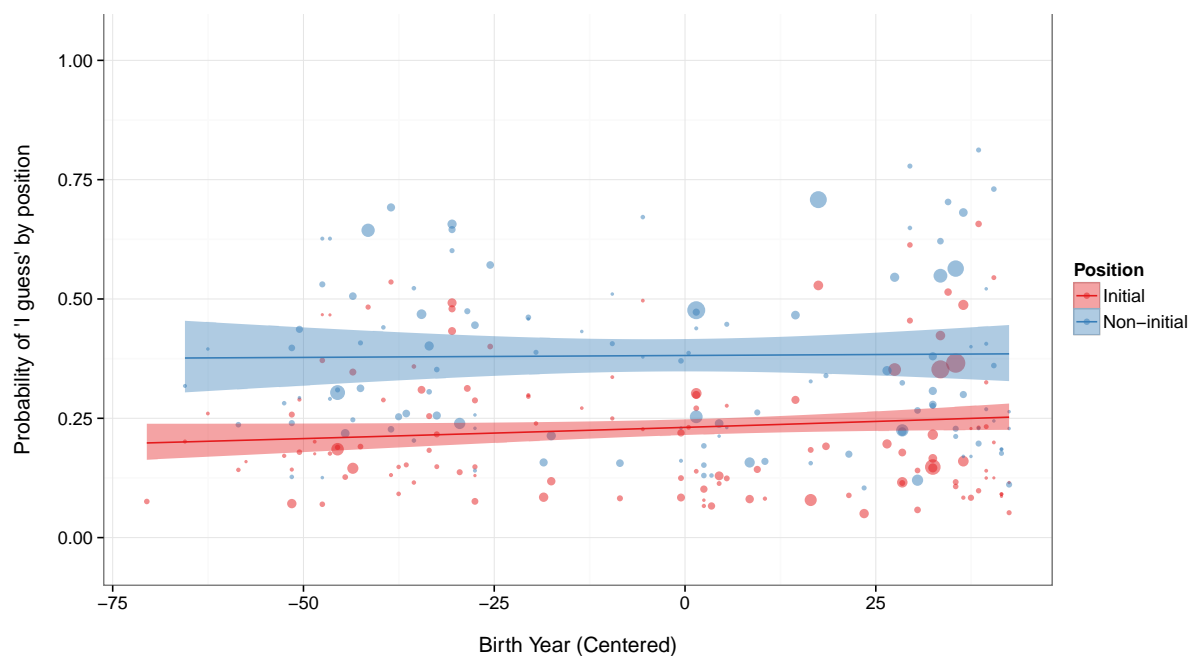


Figure 4.7: Probability of *I guess* (vs. all other EPs) by syntactic position over apparent time. Fitted values from the model in Table 4.6. Dot size represents the number of tokens at those coordinates.

dicator of *I guess*. Thus, the waxing and waning of *I guess* in Figure 4.2 is best interpreted as stability of the variant. There is a main effect of POSITION, such that *I guess* is favoured in non-initial position (compared to initial position) as indicated by the positive coefficient. The interaction of these two main effects is not significant. The significance and non-significance of these factors is confirmed in the analysis of deviance presented in Table 4.9. The model is visualized in Figure 4.7.

### Summary: Syntactic Position

Taken together, these two models help to disentangle the variable EP system. *I think* increases in frequency across the twentieth century, while *I guess* is stable. This must mean that the rise of *I think* is at the expense of all other EP variants, including *I suppose*, as suggested in Figure 4.2. Furthermore, while *I think* rises, it does so at a constant rate in initial and non-initial position, although it is favoured in initial position. Likewise, the favouring of *I guess* in non-initial position is constant across the twentieth century, though the overall frequency of this secondary variant remains stable. I suggest that this is another case of lexical replacement. Unlike with the GE system in which an innovative variant, *and stuff (like that)* has taken over the whole variable system, here the leading EP *I think* has simply replaced the obsolescing variants (*I suppose* etc.).

I now turn to the last diagnostic of grammaticalization: epistemic/doxastic strength of the verb.

#### 4.4.4 Epistemic/Doxastic Strength

In this section, I consider the hitherto untested claim that there exist differences in the epistemic (or doxastic) strength of the verb of the EP. Several researchers have argued that different EPs express different degrees of the commitment to the truth of propositions. These differences have been attributed to the retention and persis-

tence of the earlier lexical meaning of the verbs. For example, Thompson and Mulac (1991:325) observe that:

*'I think is a stronger assertion of belief than I guess. This is traceable to the difference between think and guess as verbs: guess implies an assertion based on little or no evidence, and hence less commitment to a proposition than think does.'*

Indeed, examining the OED definitions of the verbs of the three most frequent variants reveals a cline of epistemic/doxastic strength, as in (25).

- (25) a. **think**: The most general verb for expressing internal mental activity. To form or hold in the mind (an idea, image, or intuition).
- b. **guess**: To form an approximate judgement of (size, amount, number, distance, etc.) without actual measurement or calculation.
- c. **suppose** To assume (without reference to truth or falsehood) as a basis of argument. Often conveying diffidence, hesitancy, reluctance, or uncertainty in making an observation.

Of the top three variants, *think* has the strongest epistemic/doxastic strength. It is the most general term and implies that the speaker has a thought/belief. *Guess* and *suppose* both imply having weaker evidence.

The epistemic/doxastic strength of different EPs have been implicated in work on evidentiality in languages of the world. Givón (1989:134) gives the list of predicates in (26) ranked in terms of what he calls subjective/epistemic certainty.<sup>40</sup>

- (26) a. *I know* she was here.
- b. *I am sure* she was here.
- c. *I think* she was here.
- d. *I believe* she was here.

---

<sup>40</sup>Givón (1989:134) notes that these verbs often grammaticalize into evidential markers cross-linguistically.

- e. I *see* she was here.
- f. I *hear* she was here.
- g. I *guess* she was here.
- h. They *say* she was here.

Crucially, *I think* is ranked higher than *I guess*.

Aijmer (1997:18) provides a similar ranking of EPs on a scale of the degree of reliability of belief as in (27).<sup>41</sup>

- (27)
- a. *I am sure* (high degree of reliability)
  - b. *I believe*
  - c. *I think*
  - d. *I suppose*
  - e. *I guess* (low degree of reliability)

Thompson and Mulac (1991:325) implicate the epistemic/doxastic strength of EPs as a diagnostic of grammaticalization. They argue that Hopper's (1991:22) principle of PERSISTENCE predicts that throughout the grammaticalization of EPs *I think* will remain a stronger expression of belief than *I guess* and *I suppose*, and this is not only a reflection of, but indeed due to, the previous lexical meanings of the verbs.

---

<sup>41</sup>Aijmer (1997:18–19), working with Chafe's (1986) taxonomy of evidentials notes that *I think* is not restricted to expressing the "belief" mode of knowledge, but can also be analyzed as marking the "inductive" mode of knowledge. This seems to be true for the other main EP variants as well. For example, if a speaker is observing wet commuters with umbrellas detraining a London outbound train in Oxford, the propositions "I think/I guess/I suppose it is raining in London" and "it must/seems to be raining in London" could equally be analyzed as expressions of inductive knowledge. In fact, *I guess* and *I suppose* seem to be particularly suited for marking inductive knowledge. However, the line between belief and inductive knowledge is fuzzy and when and how inductive knowledge becomes a belief is a matter of epistemology. Regardless, Aijmer (1997:18) notes that within each mode of knowledge there are a range of linguistic expressions for a range of reliability.



### Operationalizing Subject of Complement Clause

As far as I am aware, this hypothesis has not been quantitatively tested in the literature. The problem is that there is no objective way to directly code a speaker's degree of strength of commitment to a proposition. In some cases, the analyst might have an intuition about the speaker's commitment, but this is bound to be subjective and different analysts might have different interpretations for any given token. As a first attempt at an objective measure of epistemic/doxastic strength, I operationalize the grammatical person of the subject of the complement clause of the EP. The hypothesis is that speakers will express a stronger commitment to propositions about themselves than about others. In other words, on average, complement clauses with first person subjects should co-occur with stronger EPs (*I think* or *I believe* according Thompson and Mulac [1991], Givón [1989], and Aijmer [1997]), while complement clauses with non-first person subjects should co-occur with weaker EPs (*I guess* or *I suppose*).

A further distinction can be made between complement clauses with third-person subjects and complement clauses with second-person subjects. Since speakers risk losing face by committing too strongly to a proposition about their interlocutor who is immediately available to deny the truth of that proposition, they may be more likely to use an EP that expresses the weakest commitment to the proposition.<sup>42</sup> Speakers might make stronger commitments to propositions about other people or other things. This theorized cline of commitment is shown in (28).

(28) First person subjects > third person subjects > second person subjects

The specific hypothesis is that *I think* with its strong commitment will be favoured with complement clauses with first person subjects, as in (29a), while *I guess* and

<sup>42</sup>This is similar to Overstreet's (1999:98) hypothesis that disjunctive GEs mark negative politeness. See Brinton (1996:238–9) for an overview of EPs and politeness.

*I suppose* will be favoured with complement clauses with third and second person subjects, as in (29b) and (29c).<sup>43</sup> Thus, to test the effect of epistemic/doxastic strength on the realization of EPs, I coded each token for the grammatical person of the subject of the complement clause: first person, second person, and third person.<sup>44</sup>

(29) a. *I think* I put in a good eight or ten hour day then.

(NIA/M/1906)

b. I never was around and *I guess* he was glad that I wasn't.

(NIA/M/1907)

c. Well *I suppose* you've seen cheese... maybe you haven't.

(EON/M/1904)

### Distributions

Figure 4.8 presents the distribution of the top three EP variants (and all others) by the grammatical person of the subject of the complement clause to which the EP is attached. The data is divided by the four main age groups.

Regardless of the grammatical person of the subject, *I think* is the majority variant across time. The one exception is in EOE where both *I suppose* and *I guess* are more frequently used with second person subjects.<sup>45</sup> This is in line with the present hypothesis: *I guess* and *I suppose* are weaker EPs and they are more frequent in contexts where speakers risk losing the most face—when they are speaking about their interlocutor, as in (30).

<sup>43</sup>Although Aijmer (1997:18) ranks *I guess* as having a weaker commitment than *I suppose*, it is not immediately obvious where these two variants should fall along this scale.

<sup>44</sup>I initially used a more fine-grained coding scheme that separately coded for NP vs. pronominal subjects, generics vs. specifics, and animacy but simple grammatical person provided the best account. Note also that many EPs occurred with sentence fragments. These tokens were coded as such and excluded from the analysis, along with any tokens in which the subject was unclear.

<sup>45</sup>There is one possible fixed expression, *I suppose/guess you could say*, which occurs three times in total.

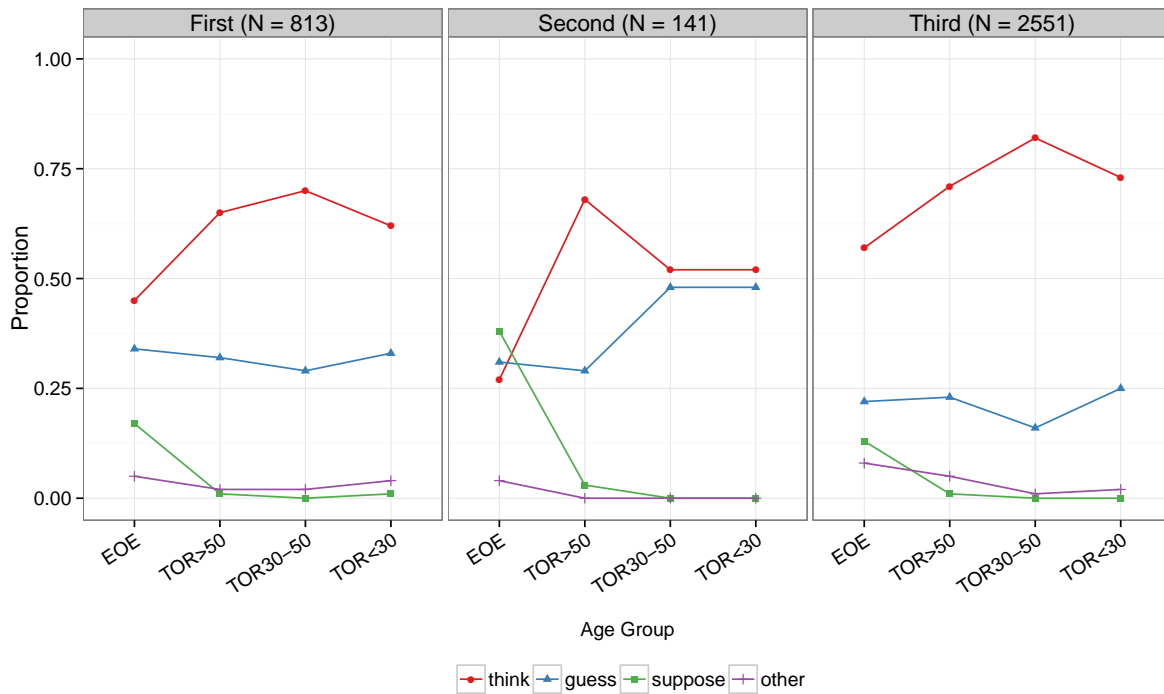


Figure 4.8: Distribution of EPs by subject of complement over apparent time.

- (30) a. I did have a position in Toronto at the Isolation Hospital. *I guess* you got that.

(BLV/F/1884)

- b. *I suppose* you know that organists play with their feet as well as their hands.

(BLV/M/1902)

In fact, before the obsolescence of *I suppose*, this variant seems to have been specialized for second person complement clause subjects, as above in (29c) and (30b): Among tokens with second person subjects, the most frequent EP is *I suppose*, a variant that otherwise is a third place contender. In TEA, although *I think* increases in frequency, regardless of the grammatical person of the complement clause subject, it is consistently the most frequent with third person subjects, then first person subjects,

and least frequent with second person subjects. There is a spike in the frequency of second person subjects with *I think* for the oldest age group in Toronto, but the frequency lowers again by the next age group.<sup>46</sup> *I guess* is stable for first and third persons. However, while the variant is more frequent with both first and second person subjects than with third person subjects in the older two age groups, it rises in frequency with second person subjects in the youngest two generations.

### Regressions

To assess the significance of these trends I turn to mixed-effects logistic regression. Because Figure 4.8 suggests that the progression of variants across time is not strictly linear, different ways of slicing the time dimension were tested to determine what provided the best statistical model.<sup>47</sup> For each, a binary split between time periods (EOE vs. TEA) provided a better model than models that included the four age group categorization or models that included a continuous predictor of speaker year of birth.<sup>48</sup> In essence, the models below test the persistence of epistemic/doxastic strength of EP variants in real time Ontario English.

Before examining the models, consider Table 4.10 which presents a distributional analysis of complement subjects in the two time periods for the top three EPs.

Comparing across these two points in real time reveals a slightly clearer picture than Figure 4.8. In EOE, *I think* is most frequent with third person complement subjects, than first person, and then second person. The same pattern holds in TEA but in each context the frequency of *I think* is higher. With *I guess*, there is a possible indication of a shift in the distribution across contexts. In EOE, first person complement subjects have the highest frequency of *I guess* followed by second persons and then

<sup>46</sup>The fluctuations for second person subjects may be due to small Ns.

<sup>47</sup>The model with the lowest AIC (Akaike Information Criteria) was chosen as the best model. See Tagliamonte and Denis (2014:102) for discussion of model comparison.

<sup>48</sup>The continuous predictor was tested as linear, quadratic (i.e., allowing for one change in the slope), and cubic (i.e., allowing for two changes in the slope).

Table 4.10: Distributional analysis of EPs by subject of the complement in EOE and TEA

	EOE						TEA					
	First		Second		Third		First		Second		Third	
	%	N	%	N	%	N	%	N	%	N	%	N
<i>I think</i>	44.7	97	26.9	7	56.9	232	63.6	287	57.9	55	74.1	1350
<i>I guess</i>	33.6	73	30.8	8	21.6	88	32.2	145	41.1	39	22.6	412
<i>I suppose</i>	17.1	37	38.5	10	13.2	54	0.9	4	1.1	1	0.5	10
Other	4.6	10	3.8	1	8.3	34	3.3	15	00	3	2.7	49

third persons. In TEA, first and second person complement subjects switch such that *I guess* is most common with second persons. As is visible in Figure 4.8, *I suppose* is most frequent with second person complement subjects in EOE with first and then third persons trailing. However, in TEA, the form has essentially obsolesced (N=15). To test the significance of these patterns, I turn to mixed-effects regression.

Table 4.11 presents the best model of the interaction between a binary main effect of time period (EOE vs. TEA; EOE as reference level) and a three-way main effect of grammatical person of the complement clause subject (first vs. second vs. third; first as reference level) on the realization of *I think* (vs. all other EPs). The significance and non-significance of these factors is confirmed in an analysis of deviance in 4.12.

Table 4.11: Mixed-effects logistic regression testing the fixed effects of PERIOD (reference level = EOE), SUBCOMP (reference level = first person), and their interaction and a random intercept for Speaker on the realization of EPs as *I think*. Treatment coding. Coefficients reported in log-odds. Correlations of fixed effects,  $r < |0.47|$ . N = 3018

	Estimate	Std. Error	z-value	p-value	
Intercept	-0.273	0.216	-1.266	$2.06 \times 10^{-01}$	
PERIOD (TEA)	1.180	0.262	4.509	$6.51 \times 10^{-06}$	***
SUBCOMP(Second)	-0.547	0.503	-1.086	$2.77 \times 10^{-01}$	
SUBCOMP(Third)	0.583	0.186	3.138	$1.70 \times 10^{-03}$	**
PERIOD:SUBCOMP (TEA:Second)	0.148	0.564	0.262	$7.94 \times 10^{-01}$	
PERIOD:SUBCOMP (TEA:Third)	-0.249	0.222	-1.124	$2.61 \times 10^{-01}$	
Random intercept:					
SPEAKER			Variance = 0.78, N = 126		

The significant, positive coefficient for the main effect of PERIOD indicates, as ex-

Table 4.12: Analysis of deviance,  $\chi^2$  test for model reported in Table 4.11.

	$\chi^2$	df	$p$ -value	
Intercept	1.601	1	$2.06 \times 10^{-01}$	
PERIOD	20.333	1	$6.51 \times 10^{-06}$	***
SUBCOMP	13.522	2	$1.16 \times 10^{-03}$	**
PERIOD:SUBCOMP	1.615	2	$4.46 \times 10^{-01}$	

pected given Table 4.10, that *I think* is more likely in TEA than EOE overall. The non-significant effect of SUBCOMP(Second) indicates that *I think* is neither more or less favoured with second person complement clause subjects than first person complement clause subjects (the reference level). However, the positive coefficient for SUBCOMP(Third) is significant, indicating that third person complement clause subjects favour *I think* more than first person complement clause subjects (and by transitivity second person complement clause subjects). The interactions between PERIOD and SUBCOMP are not significant. In other words, despite an overall increase in frequency across the twentieth century, there has been no change in the epistemic/doxastic strength of *I think* (as operationalized by the grammatical person of the subject of the complement clause).

Figure 4.9 visualizes the model in Table 4.11 by plotting the predicted probabilities. The boxplots are intended to help summarize the predicted probabilities and should not be interpreted as a direct representation of effect size or significance in the model.<sup>49</sup>

In sum, although *I think* has increased in frequency between EOE and TEA, the variant consistently favours third person complement clause subjects. I will return to this point below once the models of the other two variants have been discussed.

Table 4.13 presents the best model of the interaction of time and subject of the complement clause on the realization of *I guess* in the same way as Table 4.11. Again,

<sup>49</sup>The y-axis does not represent variable rule analysis style factor weights, which are centred around 0.5. Here, the weights are centred around the intercept (i.e., the input/correct mean). This is crucial for interpreting the figures below.

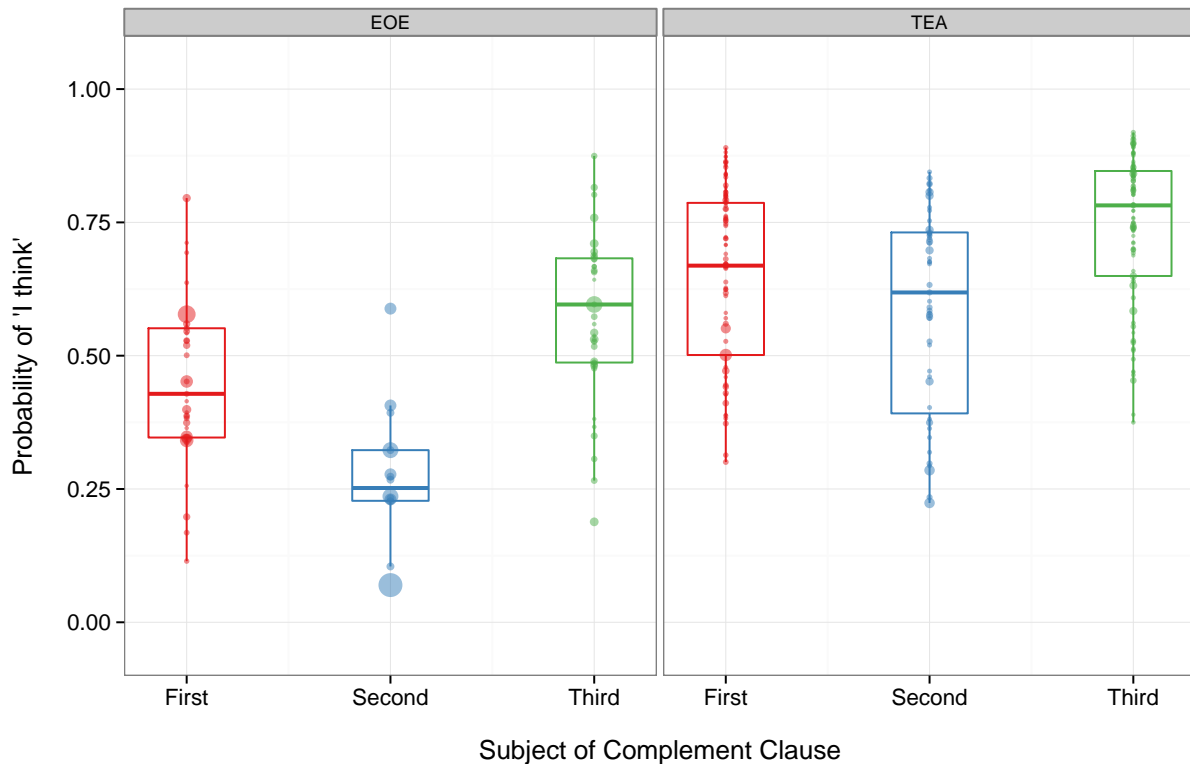


Figure 4.9: Probability of *I think* (vs. all other EPs) by subject of complement over real time. Fitted values from the model in Table 4.11. Dot size represents the number of tokens at those coordinates. EOE N = 651; TEA N = 2367.

the significance and non-significance of these factors is confirmed in an analysis of deviance in 4.14.

Table 4.13: Mixed-effects logistic regression testing the fixed effects of PERIOD (reference level = EOE), SUBCOMP (reference level = first person), and their interaction and a random intercept for Speaker on the realization of EPs as *I guess*. Treatment coding. Coefficients reported in log-odds. Correlations of fixed effects,  $r < |0.44|$ . N = 3018

	Estimate	Std. Error	z-value	p-value	
Intercept	-0.721	0.236	-3.058	$2.23 \times 10^{-03}$	**
PERIOD (TEA)	-0.489	0.285	-1.717	$8.60 \times 10^{-02}$	.
SUBCOMP(Second)	-0.638	0.504	-1.266	$2.05 \times 10^{-01}$	
SUBCOMP(Third)	-0.800	0.209	-3.826	$1.30 \times 10^{-04}$	***
PERIOD:SUBCOMP (TEA:Second)	1.202	0.568	2.116	$3.44 \times 10^{-02}$	*
PERIOD:SUBCOMP (TEA:Third)	0.496	0.244	2.028	$4.25 \times 10^{-02}$	*
Random intercept:					
SPEAKER	Variance = 0.96, N = 126				

The overall probability of *I guess* remains steady across the two time periods as

Table 4.14: Analysis of deviance,  $\chi^2$  test for model reported in Table 4.13.

	$\chi^2$	df	$p$ -value	
Intercept	9.349	1	$2.23 \times 10^{-03}$	**
PERIOD	2.948	1	$8.60 \times 10^{-02}$	.
SUBCOMP	14.709	2	$6.40 \times 10^{-04}$	***
PERIOD:SUBCOMP	6.659	2	$3.58 \times 10^{-02}$	*

indicated by the non-significance of the main effect of PERIOD. The main effect of SUBCOMP(Second) is not significant, indicating that there is no statistically significant difference in the probability of *I guess* with first or second person complement clause subjects (at least in EOE). There is however a significant difference between first and third person complement clause subjects such that first persons favour *I guess*, as indicated by the significant negative coefficient for SUBCOMP(Third). Importantly, the interactions between PERIOD and SUBCOMP are significant in the model, confirmed by the analysis of deviance. First, consider the interaction between PERIOD and third person complement clause subjects. To interpret the interaction, we can add together the coefficients of the two main effects and the interaction term ( $-.489 + -.800 + .496 = -.793$ ) and compare this to the main effect for SUBCOMP(Third) ( $-.800$ ). Although significant, the probability of *I guess* with third person subjects (relative to first person subjects) exhibits little change across time. The interaction with second person complement clause subjects suggests a more drastic change. Summing the coefficients of the main effects and the interaction term ( $-.489 + -.638 + 1.202 = 0.075$ ) suggests that relative to (the coefficient for) EOE ( $-0.638$ ), the probability of *I guess* with second person complement clause subjects (relative to first person subjects) has increased such that *I guess* is more favoured with second person subjects (than first person subjects) in TEA, opposite from the pattern in EOE (see Table 4.10).

To help visualize the model in Table 4.13, particularly the interactions, Figure 4.10 plots the predicted probabilities. It is clear that while the predicted probabilities of *I guess* remain steady with first and third person complement clause subjects, the



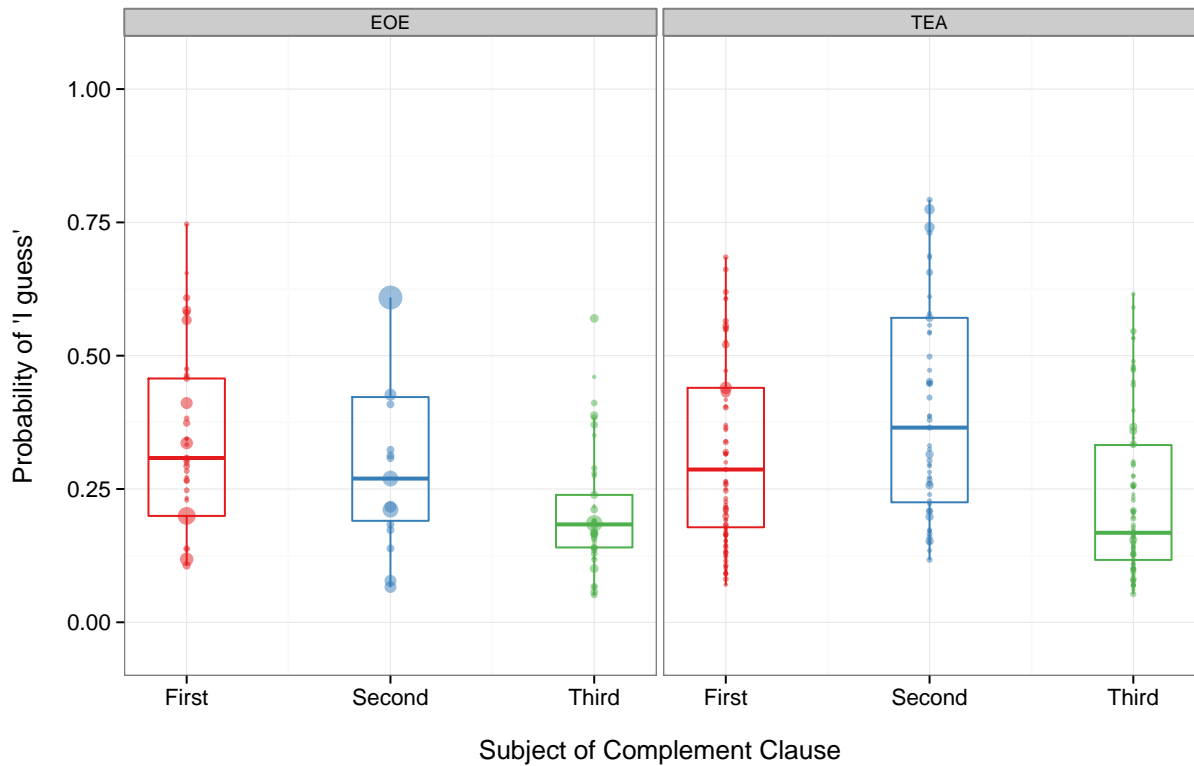


Figure 4.10: Probability of *I guess* (vs. all other EPs) by subject of complement over real time. Fitted values from the model in Table 4.13. Dot size represents the number of tokens at those coordinates. EOE N = 651; TEA = 2367.

predicted probabilities of this variant with second person complement clause subjects increases over real time. Thus, the model confirms the trend seen in Figure 4.8. Also of note is that *I guess* is disfavoured in the exact environment where *I think* is favoured (with third person complement clause subjects).

Lastly, I turn to a regression analysis of *I suppose*. However, since *I suppose* is very marginal in TEA, I only consider the EOE data. The model presented in Table 4.15 tests the effect of the grammatical person of the complement clause subject in EOE alone.

Recall that in Figure 4.8, *I suppose* was the most frequent variant with second person complement clause subjects, as in (29c) and (30b). The model above indicates that this trend is significant. There is a significant positive coefficient for `SUBCOMP(Second)` and a non-significant coefficient for `SUBCOMP(Third)`. Thus, in EOE, second person

Table 4.15: Mixed-effects logistic regression testing the fixed effect SUBCOMP (reference level = first person) and a random intercept for Speaker on the realization of EPs as *I guess* in EOE. Treatment coding. Coefficients reported in log-odds. N = 651

	Estimate	Std. Error	z-value	p-value	
Intercept	-2.123	0.292	-7.262	$3.80 \times 10^{-13}$	***
SUBCOMP(Second)	1.412	0.510	2.768	$5.65 \times 10^{-03}$	**
SUBCOMP(Third)	-0.184	0.262	-0.703	$4.82 \times 10^{-01}$	
Random intercept:					
SPEAKER			Variance = 1.19, N = 38		

Table 4.16: Analysis of deviance,  $\chi^2$  test for model reported in Table 4.15.

	$\chi^2$	df	p-value	
Intercept	52.743	1	$3.80 \times 10^{-13}$	***
SUBCOMP	10.607	2	$4.98 \times 10^{-03}$	**

complement clause subjects favour *I suppose* while first and third person subjects disfavour the form. Thus, the hypothesis that *I suppose* falls on the lower end of the epistemic/doxastic strength scale is confirmed. The model is visualized in Figure 4.11.

Table 4.17 summarizes the hierarchy of constraints of the three main variants in EOE and TEA.

Table 4.17: Summary of the results for the effect of the grammatical person of the complement clause subject in EOE and TEA.

	EOE	TEA
<i>I think</i>	$3 > \{1, 2\}$	$3 > \{1, 2\}$
<i>I guess</i>	$\{1, 2\} > 3$	$2 > 1 > 3^{50}$
<i>I suppose</i>	$2 > \{1, 3\}$	–

Some of these results are unexpected given the hypothesis above. *I think*, as the EP with the strongest epistemic/doxastic strength should be favoured with first person complement clause subjects more so than third or second persons and *I guess* and *I suppose* should be favoured with second and third person complement clause subjects. However, *I think* is favoured in the context that should represent the middle of the scale of epistemic/doxastic strength and *I guess* is favoured by the ends of the

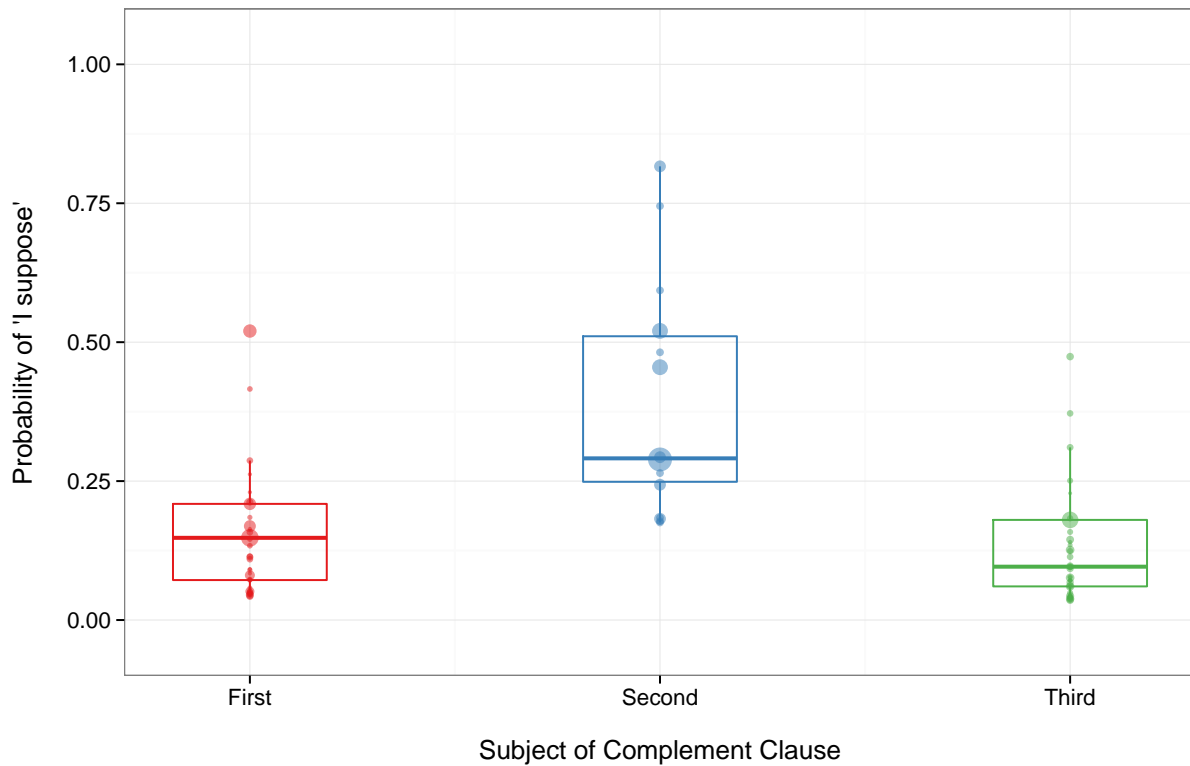


Figure 4.11: Probability of *I suppose* (vs. all other EPs) by subject of complement in EOE. Fitted values from the model in Table 4.15. Dot size represents the number of tokens at those coordinates.  $N = 651$ .

scale. Before its obsolescence, *I suppose* conformed to the hypothesis.

Thus, despite what has been suggested in the literature, these results may indicate that EPs do not nicely fit along a scale of epistemic/doxastic strength. This would undermine Thompson and Mulac's (1991) claim that EPs exhibit retention or persistence as predicted by grammaticalization theory. However, given the strong intuition that *I think* is indeed stronger than *I guess* and *I suppose*, let us explore two other possible explanations for these results. One possibility is that the way I have operationalized grammatical person of the complement clause subject to test epistemic/doxastic strength might not be accurate. Undoubtedly, there are other factors involved in how strongly a speaker is willing to commit to a proposition other than who or what a speaker is talking about. For example, we have different extents

<sup>50</sup>A three way significant difference in TEA between second, first, and third persons was confirmed by adjusting the reference level of the SUBCOMP factor in the model for *I guess*.

of knowledge about different individuals. I may be equally committed to the truth of propositions about my partner and her summer vacation plans (given that I will be involved in those plans<sup>51</sup>), as I am to the truth about propositions about me. On the other hand, I might be less likely to commit to the truth of propositions about someone I have never met. However, given the extent of the data, these other factors should even out and probabilistic trends should emerge regardless. For now, I do not wish to abandon my hypothesis about the correlation between the grammatical person of the complement clause subject and extent to which speakers commit to the truth about propositions. We are left with two facts 1) different EP variants intuitively express different degrees of commitment and 2) the models above have captured a real constraint on the realization of EPs (subject of the complement clause). This leads to another possible explanation for these results: the order in (28) is wrong. That is, perhaps first person complement clause subjects are not the propositions with which speakers will use the epistemically/doxastically strongest EP. Given our results, perhaps the order should be as in (31).

(31) Third person subjects > first person subjects > second person subjects

In fact, there is evidence from negated EPs that supports this hypothesis. Recall that EPs with a negative marker are typically the result of neg-raising, in which the negative marker of the embedded clause is raised to the main clause, as in (16c), repeated here in (32a). Neg-raised sentences are “felt as weaker and more tentative than their otherwise synonymous counterparts with lower-clause negation” as in (19a) repeated here in (32b) (Israel 2004:704).

(32) a. *I don't expect* he had too much money to bank as far as that went.

(EON/F/1906)

---

<sup>51</sup>Hopefully.

- b. *I think* they wouldn't have made a very good jam even. They were too sweet.

(NIA/F/1916)

Thus, if the hypothesis in (31) is correct, there should be a correlation between negated EPs and the grammatical person of the complement clause subjects because both have been implicated as ways in which speakers can weaken their commitment to propositions.<sup>52</sup> This can be tested with a cross-tabulation of negated EPs by complement clause subject types as in Table 4.18.

Table 4.18: Cross-tabulation of negated EPs and grammatical person of complement clause subject in EOE and TEA, showing the proportion of complement clause subject types that are negated.

	EOE		TEA	
	%	N	%.	N
Third person	14	509	7	2042
First person	18	285	13	528
Second person	22	36	9	105

In EOE, tokens with second person complement clause subjects occur most frequently with negated EPs, in the middle are first person complement clause subjects and occurring least frequently with negated EPs are third person complement clause subjects, suggesting that these are the propositions which speakers mitigate the least. The order proposed in (31) exactly lines up with the frequency of use of negated EPs. In TEA, although first person complement clause subjects occur more frequently with negated EPs than second person subjects, it is clear that speakers use more of the “weaker and more tentative” negated EPs with first person complement clause subjects than third person subjects.

<sup>52</sup>There may very well be other grammatical features that do the same. For example, the two major variants of future temporal reference in English can be thought of as expressing different levels of commitment. That is, *be going to* may express a stronger commitment than *will* because *will* “invites a conditional interpretation” (Huddleston and Pullum 2002:211).

We now have verification from the data itself that the order in (31) is correct. But what might underlie this order? Is there a theoretical reason that speakers might use weaker EPs when talking about themselves and their interlocutors than about non-participants in the discourse? The reason may lie with politeness strategies and face-saving.<sup>53</sup> I suggested above that the use of weaker EPs with propositions about one's interlocutor might be a politeness strategy that speaker's employ to avoid risking losing face. If a proposition about one's interlocutor is false, the interlocutor will most certainly know (in contrast to propositions about other people or things, the truth of which the interlocutor is not necessarily expected to know). To minimize the risk of losing face, speakers will mark the weakest commitment to that proposition. Consider again (30b), repeated here in (33).

(33) *I suppose* you know that organists play with their feet as well as their hands.

(BLV/M/1902)

If the speaker stated the utterance in (33) without the EP and his interlocutor did not know that organists play with their feet as well as their hands (i.e., that the speaker's proposition was false), the interlocutor could correct the speaker which would threaten the speaker's face. The *I suppose*, mitigates this potentially face threatening act by committing only weakly to the proposition about the interlocutor.

However, Brown and Levinson (1987), following Goffman (1967) observe that "facework involves the maintenance of *every participant's face* for the duration of the social interaction (as far as this is possible)" (Watts 2003:86, my italics). In addition to maintaining one's own face as a speaker, we are also aiming to reduce any threat to our interlocutor's face. The reason speaker's might use an EP that expresses a weaker commitment to propositions about themselves than to propositions about

---

<sup>53</sup>I assume a face-saving model of politeness such as Brown and Levinson (1987). See also Watts (2003) for an overview of this and other theories of politeness.

non-participants is the inverse reason for using a weaker EP when talking about an interlocutor. As experts on ourselves, we tend to know the truth about propositions about us. If a speaker states a proposition about him/her that a hearer had hitherto believe to be false, this could threaten the hearer's positive-face.<sup>54</sup> In particular, this is more of a face threatening act than stating a proposition about a non-participant in the discourse, since neither speaker nor hearer is the expert. Thus, by only weakly committing to propositions about us (although we are ultimately the experts), we minimize the risk of threatening our interlocutors' face. There is evidence for this function of weaker EPs in situations such as (34), in which interviewers assume something about the speaker which is false.

(34) **Interviewer:** What do you call this (points to couch).

**Speaker:** Sofa?

**Interviewer:** Sofa? Not chesterfield?

**Speaker:** Um, *I guess* I would have said chesterfield once upon a time.

(TOR/M/1930)

The interviewer asks the speaker what his generic term is for the piece of furniture that has cushioned seating for three people and, much to the interviewer's surprise, the speaker says *sofa*. The interviewer, expecting a Canadian born in 1930 to use the term *chesterfield*, seeks confirmation that the speaker would not actually say *chesterfield*. Using the EP *I guess*, the speaker weakly commits to the proposition that at some point in the past he would have said *chesterfield*. Given that his default term is *sofa*, his use of the weaker EP *I guess* allows him not to contradict himself, while at the same time, minimizes the threat to his interlocutor's face.

---

<sup>54</sup>This kind of face threatening act could fall under Brown and Levinson's (1987:66) category of "contradictions or disagreements, challenges (S[peaker] indicates that he [sic] thinks that H[earer] is wrong or misguided or unreasonable about some issue, such wrongness being associated with disapproval)."

### Summary: Epistemic/Doxastic Strength

In sum, I have explored the possibility that the realization of variants of EPs is constrained by the degree to which speakers commit to the truth of propositions. I have operationalized this epistemic/doxastic strength by the grammatical person of the complement clause subject. In one interpretation of the hierarchy of this constraint, this hypothesis flounders—there is no consistency. In a second interpretation of the hierarchy, *I think* is the strongest EP while *I guess* and *I suppose* are weaker, as claimed by Thompson and Mulac (1991) and Aijmer (1997). Given this second interpretation, the obsolescence of *I suppose* across the twentieth century, a variant once specialized as the weakest EP—it was strongly favoured with complements with second persons in EOE—has resulted in a change to the internal conditioning of the EP system. It appears that *I guess* has come to fill the role previously filled by *I suppose*: where in EOE *I guess* was favoured with both first and second person complement subjects, in TEA second person complement subjects come to favour this form significantly more than first persons, just as was the case for *I suppose* in EOE. This change occurred at the same time that the variable system generally shifted toward *I think*. That is, in each context, *I think* became more favoured between EOE and TEA. However, the internal conditioning remained the same over time.

How does this relate to the grammaticalization of these forms? Hopper's (1991:22) principle of PERSISTENCE predicts that some aspects of a grammaticalizing form's previous lexical meaning will persist as the form comes to serve a more functional role. If we follow the first interpretation of results, there is no evidence for persistence of the previous lexical meanings of the verbs of EPs. Following the second interpretation, this hypothesis is, in a sense, confirmed. The verb *think* is a general verb of cognition while *guess* implies estimation. Aspects of these meanings are preserved across time, as *I think* consistently expresses a stronger degree of commitment than *I guess* does. However, this fact alone cannot be used as a diagnostic of the continued,



gradual grammaticalization of EPs. Persistence may be predicted to occur during grammaticalization, but it is also necessary in cases of non-grammaticalization. That is, if EPs are not grammaticalizing further, we expect persistence of their meanings. Consider the fact that although *I think* has risen in frequency, the internal conditioning of the variant has remained constant. Thus, the rise in frequency was not because *I think* expanded its functioning to fill the gap left by *I suppose* (or vice versa). That is, *I think* did not undergo a bleaching of its meaning. The gap left by *I suppose* was primarily filled by the weaker *I guess*. The increase in frequency of *I think* on the other hand must have been a result of a decrease in frequency of the other low frequency variants in the system (e.g., *I believe*, *I imagine*, *I'd say*). In sum, under either interpretation of these results, there is no evidence from this diagnostic that *I think* or *I guess* has gradually grammaticalized further across the twentieth century.

## 4.5 The Development of EPs: Lexical Replacement, Specialization

Returning to Figure 4.2, it is clear that the increase in frequency of *I think* between EOE and TEA is at the expense of all other EPs, except for *I guess* which remains relatively stable. This chapter has shown that despite this rise in frequency, there is no substantive evidence for the ongoing grammaticalization of *I think*. A decrease in intervening material between the subject and the verb of the EP was used as a diagnostic of increasing fusion/decategorialization. Intervening modals and adverbials were both highly infrequent in all time periods. Although the presence of a negative marker between *I* and *think* decreased in real and apparent time, the overall frequency of negated EPs also decreased. We cannot rule out that these two observations are non-orthogonal, and thus, the evidence for increased fusion/decategorialization

is tentative at best. Changes in the syntactic position of EPs were also used as a diagnostic for further grammaticalization. One hypothesis in the literature is that increased use of EPs in non-initial position indicates further grammaticalization (Thompson and Mulac 1991). On the other hand, an increase in initial position has also been argued to indicate advanced grammaticalization of EPs to purely discourse-organization markers (Kaltenböck 2013). Regardless, a series of mixed effects logistic regressions shows that the rise of *I think* takes place at a constant rate in both initial and non-initial position. Although favoured in initial position, and thus potentially more advanced in Kaltenböck's (2013) sense, the fact that there was no change in the magnitude of this effect across time suggests that there was no further grammaticalization of the variant. *I guess* may be favoured in non-initial position, potentially more advanced in Thompson and Mulac's (1991) sense, but again, the lack of change in the magnitude of this effect suggests no ongoing grammaticalization. Lastly, the complement clause subject was used to operationalize the epistemic/doxastic strength of EPs. Another series of mixed effects logistic regressions showed that *I think* is consistently stronger than *I guess*. However, although *I think* rises in frequency, this change was not accompanied by a bleaching of this stronger meaning. Taken together, the changes in the EP system of Ontario English across the twentieth century are consistent with one possibility that Kroch (1994) predicts for morphological doublets: specialization. This observation will be the focus of the next chapter.

Finally, although this chapter has presented evidence against the idea that EPs have continued to gradually grammaticalize (as predicted by grammaticalization theory), I do not intend to suggest that EPs did not undergo a grammaticalization process at some point in the past. Pichler and Levey (2011) rightfully criticized Tagliamonte and Denis (2010) for dismissing the grammaticalization of GEs only on the basis of lack of evidence for *ongoing* grammaticalization. They suggest that synchronic stability of grammaticalization diagnostics does not indicate a lack of earlier,

but now arrested, grammaticalization. In the previous chapter, I tracked an innovative GE (*and stuff (like that)*) from its inception and found no evidence to indicate that the form underwent grammaticalization. In this chapter, I did not have the luxury of tracking an EP from its inception. As Brinton (1996) has shown, both *I think* and *I guess* have been used in English for centuries. Thus, although I did not find evidence for the ongoing grammaticalization of EPs throughout twentieth century Ontario English, this does not mean that these forms did not grammaticalize at some point in the past. In fact, it is highly likely that EPs, as adverbial adjunctives of some kind, changed from either main clauses (Thompson and Mulac 1991) or paratactic (relative) clauses (Brinton 1996, 2008; Fischer 2007) into pragmatic markers. In the concluding chapter of this thesis I will develop the idea that, crucially, this grammaticalization process was not gradual. Rather the development of EPs and GEs into the pragmatic markers that they are was abrupt and even predictable given the meaning of the lexical source.

# Chapter 5

## Conclusion

*syntax via pragmatic strengthening in discourse > syntax with a different function*

*Traugott (1995:15)*

### 5.1 Introduction

In this concluding chapter, I synthesize the major findings of chapters 3 and 4 by proposing a theory of the development of pragmatic markers that is consistent with generative views on grammaticalization and syntactic change. I then discuss the implications that stem from the results of this thesis for both grammaticalization theory and variationist work (particularly on pragmatic variables). Finally, I discuss the logical future extensions of this work.

### 5.2 The Non-Gradual Development of Pragmatic Markers

In chapter 3, I argued that grammaticalization theory, a theory that predicts the gradual change of lexical material into functional material via a number of mechanisms,

is not the right model for understanding the development of the general extenders system of Ontario English. A lack of evidence for phonetic reduction, decategorization, and semantic-pragmatic shift of the innovative variant *and stuff*, from its inception, undermines a grammaticalization theory account precisely because under that theory, these mechanisms of change only together define the phenomenon of grammaticalization. Evidence against any of these mechanisms is evidence against the grammaticalization of GEs, as conceived of by grammaticalization theory. Likewise, in chapter 4, I argued that although the epistemic parentheticals system of Ontario English reorganized through the twentieth century, these changes were not the result of ongoing, gradual grammaticalization. Again, I presented evidence inconsistent with grammaticalization theory: stability of the presence of intervening material (indicating a lack of ongoing fusion), a constant effect of syntactic position (indicating no change in the mobility of these pragmatic markers), and consistency with respect to the strength of the variants (indicating a lack of ongoing semantic bleaching).

While the results of both these case studies into the development of pragmatic markers run counter to the idea that pragmatic markers gradually grammaticalize over the course of their development according to grammaticalization theory, there is no denying that grammaticalization, *the phenomenon*, has taken place: lexical elements have become pragmatic markers. But if this phenomenon does not take place according to grammaticalization theory, how then, do pragmatic markers develop? Taken together, the two previous chapters allow for a particular conclusion: all the evidence points to grammaticalization as abrupt reanalysis of lexical material from one syntactic category to some other syntactic category (see Roberts and Roussou 2003).

In what follows I will flesh out the details of this conclusion, making reference to the relevant historical syntax literature. Furthermore, I conjecture that lexical material

does not develop a new pragmatic role out of the blue. Rather, lexical material that already triggers particular implicatures in particular syntactic positions and utterances is reanalyzed by speakers (learners) as belonging to a higher syntactic category and the implicatures that they trigger become conventionalized (cf. Brinton 1996; Waltereit 2002; 2006). Though the specifics will ultimately differ, the development of pragmatic markers is just as Traugott (1995:15) suggests: “syntax via pragmatic strengthening in discourse > syntax with a different function.”

Before fleshing out the details of this idea, it is necessary to review the concepts of parameters, competing grammars, gradualness, and abruptness.

### 5.2.1 Excursus on Variation, Change, and Competing Grammars

Kroch (1989:202) observes that syntactic change, while inherently linked with imperfect language acquisition (cf. Lightfoot 1979), is “generally gradual, with forms slowly replacing one another over centuries.” That is, the process of one variant replacing some other variant in a speech community takes place over several generations—it involves intergenerational transmission (Labov 2001; Janda 2001). Indeed, all cases of change necessarily involve some period of variability (Weinreich, Labov and Herzog 1968). In Kroch’s (1989) model, this period of variability—say, the ‘transition’ phase (cf. Weinreich, Labov and Herzog 1968)—is conceived of as GRAMMAR COMPETITION. In essence, where speakers have multiple options for expressing some expression, these multiple options are instantiations of separate, minimally different, grammars of their language. A change from language state  $\mathcal{A}$  to language state  $\mathcal{B}$  takes place when one grammar replaces another grammar. Before moving on, I must make clear what is meant by grammar and grammar competition.

Within Chomsky’s (1981, 1986) principles and parameters framework (the model of grammar with regard to which Kroch [1989] first discussed grammar competition), all languages are subject to a set of invariant PRINCIPLES and diversity arises “by

means of assigning different values to a finite set of options, called *PARAMETERS*” (Roberts and Roussou 2003:9). For example, different settings of one such parameter in Italian and English, the null subject parameter, are responsible for the fact that in sentences like (1), Italian lacks a subject pronoun but English must contain a subject pronoun.<sup>1</sup>

- (1) a. Parla italiano  
       speak<sub>3,SG</sub> Italian  
       ‘He/she speaks Italian
- b. \*Speaks Italian

(Roberts and Roussou 2003:9)

The null subject parameter in Italian is set such that the grammar of an Italian speaker allows for null subjects (see Rizzi 1982 for full details). In the grammars of English speakers, this parameter is set such that it (typically) disallows null subjects.

Contemporary generative syntax within the framework of Minimalism (Chomsky 1995 *et seq.*) has localized parameter settings to properties of functional heads (e.g., Biberauer, Holmberg, Roberts, and Sheehan 2010). In brief, Minimalist approaches to syntax generally agree that 1) the input to syntactic structure is lexical items, including functional heads, located in the lexicon, 2) lexical items are themselves composed of a set of features, 3) features provide instructions to the syntactic derivation, and 4) syntactic operations are invariant, given the lexical items (i.e., their features). That is, Italian and English differ not with respect to an amorphous null subject parameter but because of the differing properties (i.e., features) of some particular functional head in each language.<sup>2</sup> Another example is the case of Icelandic and Danish, which are parametrically different with respect to the raising of *V(v)* to *T* in embedded

<sup>1</sup>English does allow for variable null subjects in certain contexts (Harvie 1998). However, Haegeman (1990) argues that instances of ‘diary drop’ in English are not the result of different settings of the null subject parameter.

<sup>2</sup>The idea that linguistic diversity is contained within the lexicon is known as the Borer-Chomsky conjecture as similar ideas were independently proposed in Borer (1984) and Chomsky (1995).

clauses. Consider (2).

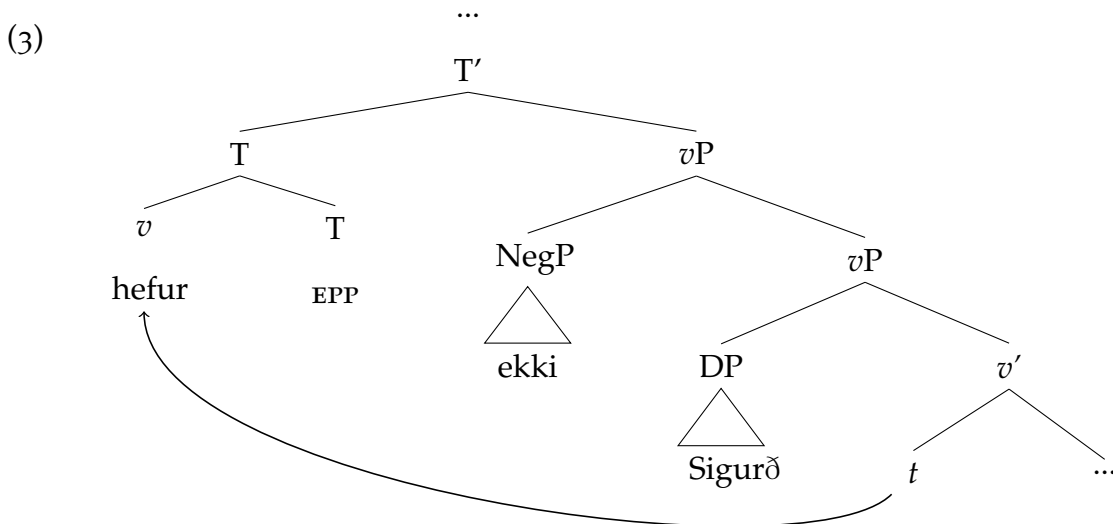
- (2) a. *Ég veit að [ Sigurð [ hefur ekki farið til Toronto ] ]*  
 I know that [ Sigurð [ has **not** gone to Toronto ] ]  
 'I know that Sigurd hasn't gone to Toronto.'

(Icelandic)

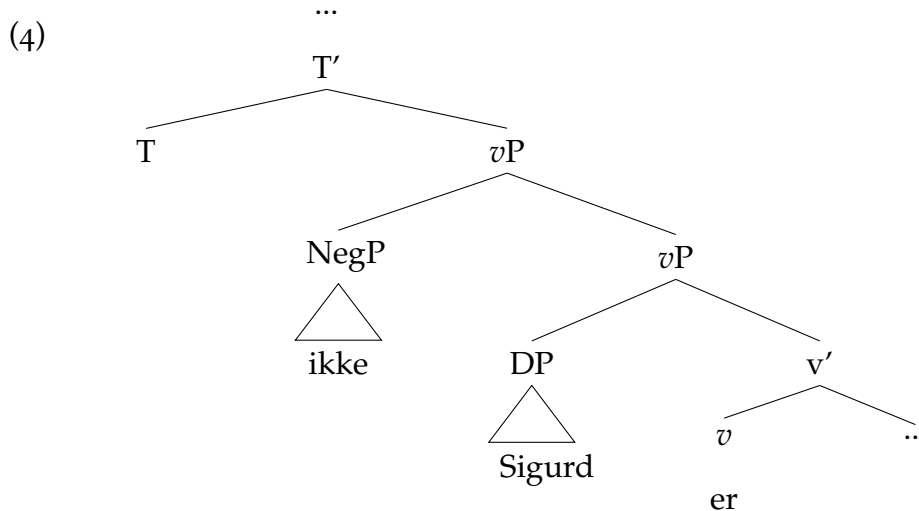
- b. *Jeg ved at [ Sigurd [ ikke er gået til Toronto ] ]*  
 I know that [ Sigurd [ **not** has gone to Toronto ] ]  
 'I know that Sigurd hasn't gone to Toronto.'

(Danish)

Assuming, based on independent evidence, that the negative marker in both of these languages (*ekki* and *ikke* respectively) is an adjunct to *vP*, these two sentences exhibit one minimal difference: in Icelandic, there is head-to-head movement such that *v* (*hefur*) moves from its merged position to above the negative marker (in *NegP*), arguably to head-adjoin with *T*, as in (3), while in Danish, no such head movement occurs and *v* (*er*) remains in situ, as in (4).







The crucial difference, while reflected in the movement (or not) of  $v$  to T, can be thought of as a result of different properties of the functional head T; the lexicon of Icelandic speakers contains a T that has some property that triggers the movement of  $v$  (perhaps an EPP feature) and the lexicon of Danish contains a T that does not. While Icelandic has one grammar and Danish has another, the locus of difference between these grammars is in the lexicon. Indeed, the upshot here is that cross-linguistic variation (at least morphosyntactic variation) is a matter of the properties of functional heads in the lexicon (Borer 1984; Chomsky 1995; Roberts and Roussou 2003; Biberauer, Holmberg, Roberts, and Sheehan 2010).

Returning to Kroch's (1989) model, if the variation that necessarily accompanies linguistic change is a result of competition between two grammars within a single speaker, this boils down to variation within the lexicon of a single speaker: multiple functional heads, with minimally different properties.<sup>3</sup> The concept of grammar competition in a Minimalist framework then is reducible to competition between near-identical functional heads. If a language were to change from an Icelandic verb raising grammar to a Danish verb in situ language, there must have been a period of variation when the speakers of that language had both an Icelandic-like T *and* a

<sup>3</sup>The existence of multilingualism is *a fortiori* evidence that speakers can have a command of multiple grammars (Kroch 1989:202).

Danish-like T. Indeed, one of Kroch's (1989:220) *prima facie* examples of grammar competition is the change in English from Middle English verb raising to Modern English non-verb raising. Consider the contrast in (5).

- (5) a. Quene Ester looked **never** with swich an eye.

(Chaucer, *Merchant's Tale*, line 1744, cited in Kroch 1989:226)

- b. Queen Esther **never** looked with such an eye.

(Kroch 1989:226)

In Middle English, the negative marker *never* occurs after the verb, while in Modern English it occurs before the verb. Analyzing these two periods of English in the same way as Icelandic and Danish above, Middle English is characterized as having a grammar with Icelandic-like T and Modern English has a grammar with Danish-like T. The transition between Middle English and Modern English involved a period of variability in which these two grammars were in competition and the grammar with Danish-like T 'won'. Indeed, an examination of this transition phase, between Middle and Modern English, shows just such variability, as in Figure 5.1, a simplified version of Kroch's (1989:228) Figure 7.

It is clear in Figure 5.1 that over two hundred years, there was variation in English in the order of main verbs and *never* as a result of competition between Icelandic-like T and Danish-like T. However, it is also clear that there was a trajectory toward decreasing use of post-verbal *never* (i.e., Icelandic-like T). Over these two hundred years, English gradually approached the categorical Modern English word order.

At this point, I must clarify what is meant by GRADUAL.

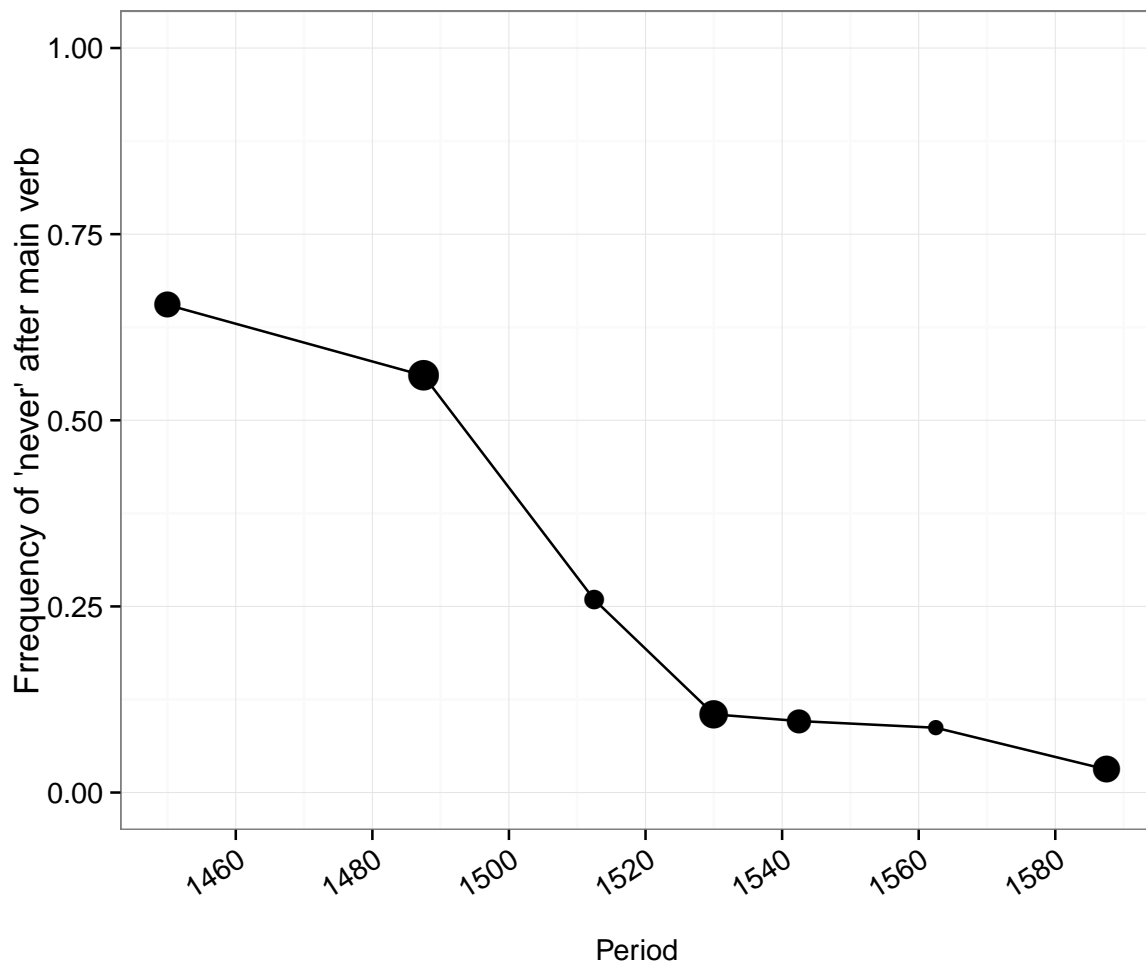


Figure 5.1: Simplified Figure 7 from Kroch (1989:228): “The decline of [*v* to T] raising in sentences with *never*.”

### 5.2.2 Gradualness, Discreteness, and the Constant Rate Effect

While in Kroch’s (1989) model of change, the process of transition is gradual, the transition phase itself should not be confused with grammatical change proper. Grammatical change should be thought of as a situation in which a community of speakers have acquired a grammatical system that is (parametrically) different from the grammatical system of the previous generation; in other words, through the regular process of intergenerational language transmission from parent-to-child, the set of parameters (functional heads) in a child’s system differs from the set of param-

eters (functional heads) in the parent's system (Roberts and Roussou 2003:11). As such, a change from one grammar to another is "necessarily instantaneous" (Kroch 1989:201).<sup>4</sup> Such a change is typically thought to be the result of REANALYSIS of the primary linguistic data in language acquisition, as triggered by some ambiguous context(s) (Lightfoot 1979; Roberts and Roussou 2003).<sup>5</sup> Fortson (2003:656) reminds us however that "[o]ne must not conflate the succession of diachronic events that precede reanalysis with the reanalysis itself." Indeed, the process of change in Kroch's (1989) model can be thought of as involving three discrete language states and two changes: state  $A$  with grammar  $\mathbb{A}$ , state  $AB$  with competing grammars  $\mathbb{A}$  and  $\mathbb{B}$ , and state  $B$  with grammar  $\mathbb{B}$ . On an abstract level, grammatical change is not gradual but involves two discrete changes: the addition of grammar  $\mathbb{B}$  (i.e.,  $A \rightarrow AB$ ) and the loss of grammar  $\mathbb{A}$  (i.e.,  $AB \rightarrow B$ ).

More concretely, the change in English discussed in the previous section involved a Middle English state in which the grammar of its speakers contained Icelandic-like  $T$  (language state  $A$ ). A change happened such that a new grammar (i.e., a new functional head) became available to speakers of a subsequent generation of English speakers; this is the grammar which contained Danish-like  $T$ . This is the starting point of the transition phase, when the probability of the new option increases from "zero to some small positive value in a temporal discontinuity" (Kroch 1989:205). This would have occurred prior to the earliest point in Figure 5.1 when the frequency of *never* after main verbs would have changed from one hundred percent to less than one hundred percent. For several centuries, the speakers of English had both options

---

<sup>4</sup>The instantaneity of change is a result of the fact that the grammar of one generation is either the same or different from the previous one. Once a child (or adolescent, see below) acquires something new and different from their parent into their grammar, a discrete change has occurred. With respect to the whole transition phase, this instantaneous reanalysis can be thought of as Weinreich, Labov, and Herzog's (1968) 'actuation'.

<sup>5</sup>Of course, two other situations lead to change of this sort. First, new variants can enter a community via diffusion from (contact with) some other community. Diffusion, unlike transmission, typically results in imperfect replication (Labov 2007; Tagliamonte and Denis 2014). The second situation, less recognized in the diachronic literature, is change occurring in adolescence, from peer to peer (Labov 2001; Tagliamonte and D'Arcy 2009).

(as shown in all of Figure 5.1), which competed (language state  $AB$ ). At some later point, another change occurred such that Icelandic-like T was no longer available in the grammars of some further subsequent generation (language state  $B$ ). This would have occurred after the latest point in Figure 5.1, when the frequency of *never* after main verbs became zero percent.

While the transition phase from language state  $A$  to language state  $B$  occurred over several centuries, the two substantive grammatical changes occurred abruptly. What changed gradually was usage patterns. In the model of grammar and change that I am discussing here, there is a clear distinction between two components of language, the GRAMMAR and USAGE.<sup>6</sup> Adger (2007) presents a schematic of such a system of grammar and usage that directly addresses the presence of variability from within a generative framework, as in (6).

- (6) a.  $\mathbf{G} \rightarrow \{v_1, \dots, v_i, \dots, v_n\}$  (=PoV)  
 b.  $\mathbf{U}(\mathbf{PoV}, \mathbf{C}) = v_i \in \mathbf{PoV}$

In this schematic of grammar and usage, grammar ( $\mathbf{G}$ ) is, as all generative grammars are, “a device that generates all of the grammatical sentences of a language and none of the ungrammatical ones” (Chomsky 1957:13). However, Adger’s (2007) system explicitly recognizes that the grammar of a language has ways of producing multiple grammatical options of saying the same thing (for example, multiple, minimally different functional heads). The set of these minimally different options is the pool of variants ( $\mathbf{PoV}$ ).<sup>7</sup> This is the extent of the grammatical system proper. Variation arises in usage ( $\mathbf{U}$ ), thought of as a choice function that takes the  $\mathbf{PoV}$  produced by the grammar and given the (sociolinguistic and linguistic) context ( $\mathbf{C}$ ) chooses a variant

<sup>6</sup>This has been the working assumption of most modern theoretical linguistics since Saussure’s (1916/1966) distinction between *langue* (the grammatical system) and *parole* (the social use of language) (Chambers 2009:26).

<sup>7</sup>In most cases, the  $\mathbf{PoV}$  is a singleton set (i.e., no variation).

(*v*).

Given that competing grammars can be relativized to competition between functional heads, this is essentially the same system as Kroch (1989): Adger's (2007) grammar (*G*) that produces multiple variants ( $v_1...v_n$ ) can be thought of as Kroch's (1989:202) "repertoire of grammatical knowledge" that contains a set of competing grammatical options. Both these systems place the probabilistic selection variants/competing grammars in the usage component of language.

The study of language use is the study of the choices that people make among alternative forms in their repertoire of grammatical knowledge in formulating utterances [and] [...] variation often reflects choices that are not categorically determined by linguistic principles at any level but instead are only probabilistically influenced by features of context and situation. In the case of replacement of one form by another, this is the expected circumstance [...]. To study such replacement is to determine the nature and weight of these probabilistic factors and to trace their temporal evolution [...].

(Kroch 1989:202)

The eventual source of variability is the structure of the lexicon, which is derived from a very general algorithm that links syntactic features with morphological forms. This set of feature-form mappings, together with the invariant syntactic operations, leads to a potential pool of variants from which the systems of use select a particular item on any occasion of utterance. [...] The choice of variant is [...] subject to pressures of processing and sociolinguistic status, neither of which are part of the grammatical system proper.

(Adger and Smith 2010:1133)

In this view of variation, one of the fundamental features of a linguistic variable, structural heterogeneity, is split across these modules of language: the 'heterogeneity' comes from multiple options produced by the grammar and the 'structure' comes from the probabilistic choice function, which is subject to social, linguistic, and cognitive constraints.<sup>8</sup> The clear divide between grammar and usage means that grammatical change is abrupt, while its consequences (the competition between multiple

<sup>8</sup>The extent to which certain linguistic constraints are part of the grammar or a condition of the

options in usage) play out probabilistically and gradually. As Pintzuk (2003:510) puts it “[t]he gradual nature of syntactic change is thus simply a reflex of the gradual nature of grammatical competition.”

The evidence for this modularity comes from the study of usage data and the discovery of the CONSTANT RATE EFFECT, as discussed in the previous two chapters (Kroch 1989). It is a fact about usage that a particular option might be highly favoured in some particular context or situation but much less favoured in some other context. That said, if the magnitude of such contextual or situational factors remains constant across the transition phase, as one option becomes more frequent than the other, we can assume that some underlying change is operating. Indeed, the constant rate effect is a prediction about cases of competing grammars. Because any case of grammatical competition is the result of a single underlying parametric change in a speaker’s grammar (i.e., the addition of a new functional head that is minimally different from some other functional head in the language), the rate of replacement during the transition phase will be constant, regardless of widely varying probabilistic factors of context and situation. As Kroch (1989:199) puts it “[c]ontexts change together because they are merely surface manifestations of a single underlying change in grammar.” Thus, wherever we find evidence of the constant rate effect, we can assume that there is a single underlying change in progress, i.e., grammatical competition. Conversely, I have argued in this thesis that certain mechanisms of grammaticalization, particularly semantic-pragmatic expansion, are diagnosable by evidence of different rates of change across contexts.

---

choice function is an open question. Adger and Smith (2005) and Adger and Smith (2010) make the case that at least some linguistic effects might be the result of the structure of the pool of variants and thus part of the grammar module. In particular, person and number constraints on *was/were* and verbal-*s* might be the result of multiple variants having slightly different  $\phi$  (i.e., person, number, gender) feature bundles.

### 5.2.3 Reconciling the Constant Rate Effect with Pragmatic Change

As just discussed, the constant rate effect is a prediction about competing grammars where competing grammars are conceived of as competing functional heads in the syntax. However, in chapters 3 and 4, several cases of a constant rate of change in the development of two sets of pragmatic markers were observed. While for many years pragmatic markers hovered along “the edge of grammar” (Massam, Starks, and Ikiua 2006), receiving only minimal attention from generative grammarians, a recent surge in interest has resulted in a growing understanding that pragmatic markers are syntactic objects proper (Bayer and Obenauer 2011; Davis 2011; Massam, Starks, and Ikiua 2006; Speas and Tenny 2003; Lam, Thoma, and Wiltschko 2013). Since pragmatic markers tend to operate with scope over whole propositions (as we have seen in this thesis), the majority of generative analyses consider pragmatic markers as functional heads or in dependencies with functional heads of the left periphery, above CP (or within an expanded CP) (Rizzi 1997).<sup>9</sup> A full review of the syntactic evidence for the status of pragmatic markers as syntactic is beyond the scope of this thesis, but if we begin with the assumption that GEs and EPs are in a dependency with some functional heads above CP, it is no surprise that the changes discussed in chapters 3 and 4 conform to the predictions of the constant rate effect. That is, these pragmatic markers, as syntactic objects act like any other syntactic objects through diachrony. Competition between *stuff* type GEs and other GEs or between *I think* and *I guess* and *I suppose* are instantiations of grammar competition in the same way that a change in word order from Middle English and Modern English was the result of competition between Icelandic-like T and Danish-like T. More specifically, the results of this thesis conform to each of the two possible outcomes of competing grammars:

---

<sup>9</sup>Much of this recent work takes as a starting point an earlier observation from Lewis (1970). Lewis (1970:55) proposed analyzing different sentence types as composed of two parts, the sentence radical “that specifies the state of affairs” and the mood or force that determines the sentence type. See Davis (2011) for more on Lewis’ proposal.



obsolescence and specialization.

Kroch (1994:8) argues that grammatical competition between morphosyntactic doublets will result in one of two possible outcomes:

1. In the absence of further linguistic change, one form eventually disappear[s] through disuse, just because of stylistic preferences or random statistical fluctuations. [...]
2. The doublet pair [becomes] stable due to differentiation in meaning and grammatical properties.

In chapter 3, I argued that the GE *and stuff (like that)* arose in Ontario English in the late nineteenth century and took over the previous variation in the system of adjunctive general extenders. This is consistent with Kroch's (1994) first possibility. In Tagliamonte and Denis (2010), we referred to such a change as lexical replacement: one variant of a set of pragmatic markers serving the same/similar functions replaced the other options. In the expanded diachrony of chapter 3, I tracked *and stuff (like that)* from its inception in the grammar of the Ontario English speech community. As it rose in frequency, it did so at a constant rate in different contexts (see Figure 3.13).

The changes that have taken place in the EP system are more in line with Kroch's (1994) second possibility. Although one form, *I suppose*, went to obsolescence, the top two forms (*I think* and *I guess*) remain in variation. Thus, we do not have lexical replacement. Rather, the results of the previous chapter suggest a trend toward the specialization of both *I think* and *I guess*. Each variant fills a probabilistic niche as in Table 5.1.

Table 5.1: Summary of the probabilistic niches of *I think* and *I guess*

	<i>I think</i>	<i>I guess</i>
Negation	Specialization of <i>I don't think</i>	Never negated
Syntactic position	Favoured in initial position	Favoured in non-initial position
E/D strength	Stronger commitment	Weaker commitment

In Kroch's (1994:8) model, stability can only result from clear "differentiation in meaning and grammatical properties." However, Wallenberg (2013) and Fruehwald and Wallenberg (2013) have argued that this will happen only if forms specialize along a categorical dimension (e.g., *fit* as the past tense and *fitted* as an adjectival passive). Stable *variation* can occur if forms specialize along some continuous dimension. For example, following Speyer (2010), Wallenberg (2013) discusses the case of topicalization word order in English, which has been stable since late Middle English but also constrained by prosody, with prosodically weak pronominal subjects favouring topicalization word order, as in (7).

- (7) a. [The first]<sub>i</sub> she'll feed  $t_i$  mouse chow, [the second]<sub>j</sub> she'll feed  $t_j$  veggies, and [the third]<sub>l</sub> she'll feed  $t_k$  junk food.
- b. ?[The first]<sub>i</sub> Anders will feed  $t_i$ , [the second]<sub>j</sub> Joel will feed  $t_j$ , and [the third]<sub>k</sub> Wim will feed  $t_k$ . Maggie will pay  $t_k$ .

In this way, topicalization movement is constrained by the continuous factor of the prosodic strength of the subject with movement specializing along the weak end of the scale and non-movement along the stronger end of the scale. The end result is stability.

In the case of EPs, *I think* and *I guess* have specialized along the scale of epistemic/doxastic strength, also a continuous scale. *I think* has specialized along the stronger end of the scale while *I guess* has specialized along the weaker end. This specialization has led to contemporary stability. Consider Figure 4.2, where after 1920, there is general stability of the two forms with *I think* hovering around seventy-five percent and *I guess* around twenty-five.

All told, the results of this thesis suggest that GEs and EPs are present in the syntax in a dependency with some functional heads.<sup>10</sup> In each case, I have found

---

<sup>10</sup>To determine which particular functional heads would require further investigation, but see below.

evidence of the constant rate effect operating which is indicative of grammar competition (i.e., variation between two minimally different functional heads). In addition, the two different outcomes in the two previous chapters, obsolescence and specialization, are consistent with what is expected in cases of such grammar competition.

In Forston's (2003:656) words, the previous section has discussed "the succession of diachronic events" that follow from grammatical change. So what then were the initial grammatical changes that took place with respect to these two sets of pragmatic markers and what role does the phenomenon of grammaticalization play?

#### **5.2.4 A Schematic for the Development of Pragmatic Markers as (Abrupt) Reanalysis**

While grammaticalization theory characterizes the change of lexical material to functional material as a gradual process associated with multiple interconnected mechanisms, the system I have described so far only allows abrupt, instantaneous changes to arise. That said, there are many documented cases of lexical material becoming functional material (or functional material becoming even more functional) in historical linguistics. How then, do we account for the phenomenon of grammaticalization in this model of language change? Very simply, following Roberts and Roussou (2003) among others, these grammaticalization phenomena must be understood not as some special form of language change, but rather, the opposite, as just like any other kind of change. In particular, Roberts and Roussou (2003:2) make a strong case for grammaticalization as "a regular case of parameter change not fundamentally different from other such changes." For Roberts and Roussou (2003:35) "grammaticalization involves reanalysis of functional categories [...] in such a way that new morphophonological realizations of functional features are created." While their mechanism of change involves the reanalysis of MOVE-based to MERGE-based instantiations of functional features, reanalysis of any sort could presumably result

in grammaticalization phenomena.<sup>11,12</sup> In what follows I make two conjectures about the conditions necessary for the reanalysis of lexical material into pragmatic markers using EPs and GEs as an example.

### **Conjecture 1: Position Matters**

The first condition under which lexical material may be reanalyzed as a pragmatic marker is for the lexical material to be positioned on either periphery of the utterance. In other words, the lexical material must occur utterance-initially or utterance-finally.<sup>13</sup> Given the hierarchical structure of syntax, these two (linear) positions, under certain conditions, may be parsed by learners as taking scope over the entire utterance. The ambiguity caused by the linearization of hierarchical structure is what leads to such reanalysis. In a generative framework, the outer edges of the utterance might contain material above CP (or within an expanded CP). As discussed above, there is precedence in the literature for assuming that functional heads that express interpersonal meaning, as pragmatic markers do, are located above CP. Bayer and Obenauer (2011) associate German modal particles (e.g., *denn* ‘then’; *bloß* ‘barely’) that “express the speaker’s attitude about him-herself or about the hearer with respect to the propositional content of the utterance” with the FORCE head of an expanded CP structure (see Rizzi 1997). Likewise, Lam (2014) associates similar pragmatic markers in Cantonese with a complex FORCEP system. Wiltschko and Heim (2014) propose a layer of syntax above CP that deals explicitly with meanings that involve GROUNDING: the process by which speakers and hearers negotiate their sets of beliefs to establish mutual understanding. There is also precedence from variationist work. Tagliamonte

---

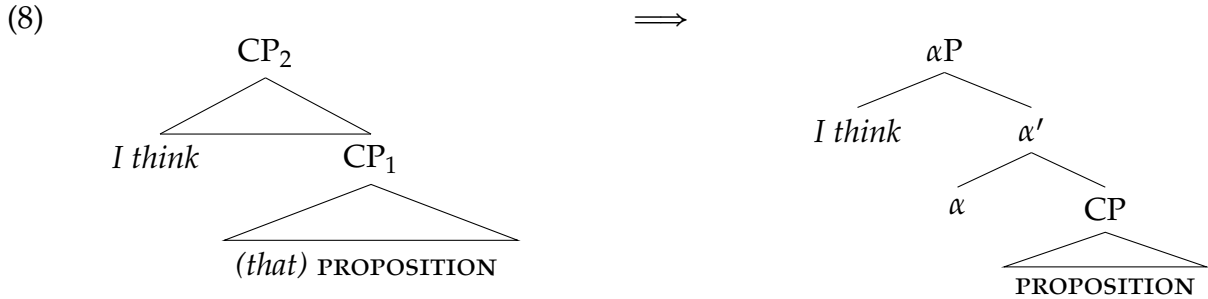
<sup>11</sup>Note that Roberts and Roussou (2003) survey changes that fall under the traditional definition of grammaticalization. They do not consider the ‘pragmatization’ side of grammaticalization. The extent to which these changes differ structurally is a question left for future research.

<sup>12</sup>For a general critique of the MERGE OVER MOVE economy principle, see Motut (2010).

<sup>13</sup>There are a number of documented pragmatic markers that are limited to sentence medial position, such as German modal particles. The syntax of these features and how they relate to functional material above the CP will not be discussed here but see Bayer and Obenauer (2011) and Lam, Thoma, and Wiltschko (2013).

(2014a) links utterance initial pragmatic markers with an expanded left periphery.

Assuming that this is case, the structural reanalysis of EPs might have been along the lines of (8).<sup>14</sup>

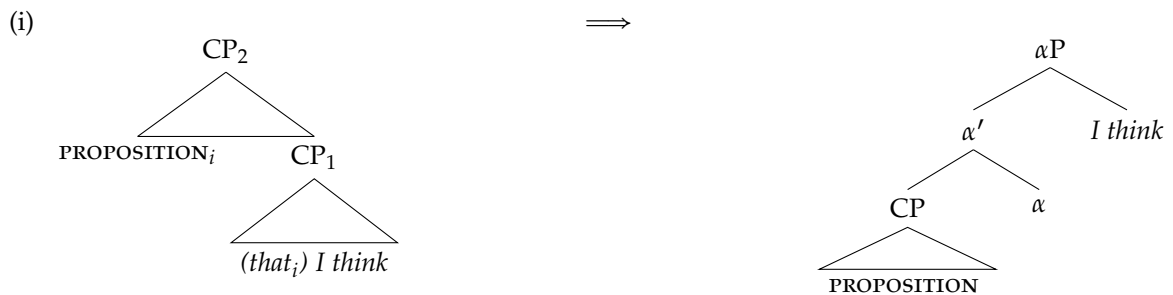


I remain agnostic about the details of the functional heads above CP but the development of EPs proposed here involves the subject and predicate of a CP with scope over another CP being reanalyzed as a specifier of a functional head above CP ( $\alpha$ ).<sup>15</sup> This is similar to what Thompson and Mulac (1991) propose, though for them the EP is reanalyzed as an adverbial. The adjunctive status of adverbials allows for the kind of positional mobility observed with EPs, as in (9).

(9) a. I don't know if it was Gerrard or below Gerrard. *I think* it was Gerrard.

(TOR/F/1917)

<sup>14</sup>In this example, I assume the trajectory of change proposed by Thompson and Mulac (1991) but the trajectory observed by Brinton (1996) is also consistent with my conjecture about peripherality. Here, the relative clause is on the right end of the utterance but this linear position is also reanalyzable as in a structural position above CP as below.



<sup>15</sup>The possibility of a non-overt complementizer in the embedded CP might have been necessary for the ambiguity to arise.

- b. Everyone has their bad apples *I guess*.

(TOR/F/1980)

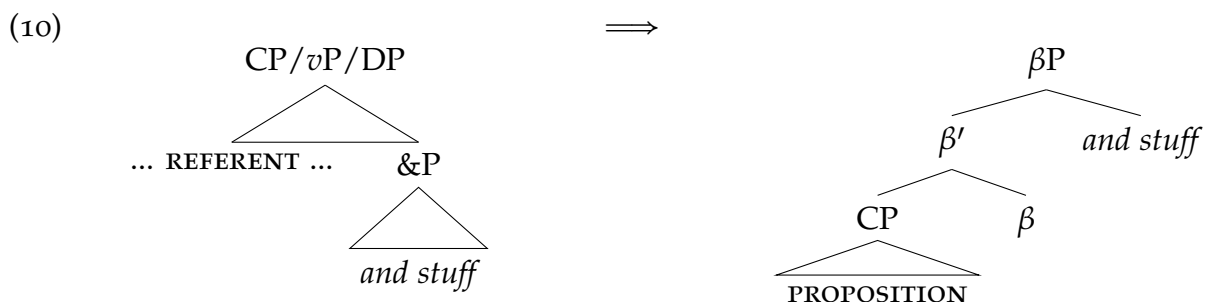
- c. People in Toronto *I think* are a lot more stressed.

(TOR/F/1980)

In the syntax of the pragmatic marker *I think* proposed here, such positional mobility could be achieved via movement of material within the CP above  $\alpha$ P. This could include movement of the whole CP, resulting in the utterance final position of the EP or some constituent within the CP, resulting in utterance medial position.

Crucially, competition between variants of EPs can be interpreted as competition between minimally different  $\alpha$  heads. Each  $\alpha$  has some minimally different feature, selecting a different EP variant just in the same way that Icelandic-like T has a feature that triggers movement to its specifier and Danish-like T does not have that feature.<sup>16,17</sup>

For GEs, the change involves the reanalysis of utterance-final lexical material as located in a higher functional phrase, as in (10).



<sup>16</sup>There are two possible sources of competition: 1) the multiple reanalysis of epistemic/doxastic complement taking verbs with first person subjects, or 2) the reanalysis of one, followed by analogical change for the others. Either option is possible and both would presumably result in a process of lexical replacement or specialization (Kroch 1994).

<sup>17</sup>This view of pragmatic markers as syntactic objects has an interesting consequence for Torres Cacoullos and Walker's (2009a:21) discussion of highly frequent "discourse formulas". There is no longer a dichotomy between "belonging to the lexicon as fixed or frozen (discourse-pragmatic) units" and belonging to "a productive grammar as instantiations of a construction with open-class positions" (Torres Cacoullos and Walker 2009a:21); being part of a productive grammar means being part of the lexicon and interacting with functional heads which are also part of the lexicon.

Again, I remain agnostic about the specific details of the functional heads above CP. However, since GEs are typically utterance-final, it is possible for them to be reanalyzed as part of a phrase higher than the CP of which they were earlier a part.<sup>18</sup> For example, the semantically bleached GEs discussed in §3.4.3 seem to take scope over the whole CP. We can see this when another pragmatic marker intervenes between the the end of the proposition and the GE as in (11).<sup>19</sup>

- (11) I'm excited to see Blink because they're a band I grew up listening to so I'm really stoked for that. Cypress Hill will be pretty cool to see I guess and stuff like that. Slayer and Metallica are rad bands so it'll be cool.

If we assume that the EP *I guess* is hierarchically above the CP, then the GE must be as well.

### Conjecture 2: Implicata Matter

The second conjecture is that only instances of a specific type of lexical material on the peripheries of utterances are reanalyzable as pragmatic markers—only lexical material that triggers particular implicatures consistent with the meanings conventionalized by functional heads above CP (see Bayer and Obenauer 2011, Lam 2014, Wiltschko and Heim 2014 etc.). Recall the discussion of semantic-pragmatic shift of GEs in §3.4.4. I argued that while advanced stage GEs only function interpersonally (e.g., to mark assumed shared knowledge, see (11) above), GEs at an earlier stage inherently expressed this interpersonal function in addition to the propositional function of marking a set. That is, the propositional meaning of GEs themselves (to

<sup>18</sup>The only difference between EPs and GEs is that  $\beta$  is right-headed where  $\alpha$  is left-headed (though see note 14 above)

<sup>19</sup>This example was found through a google search of the phrase “I guess and stuff”. There were millions of hits and many on the first ten pages of results represented examples just like (11). There are no examples like this in TEA however. The example itself comes from an interview with a twenty-something year old musician, native to Vancouver <<http://www.theaureview.com/interviews/chase-brenneman-from-living-with-lions-vancouver>>.

indicate a set of which its referent is a member) triggers the implicature that the speaker assumes her interlocutor has sufficient shared knowledge to construct the intended set. In a bleaching model of semantic change, the propositional meaning is lost and the interpersonal meaning remains; the GE is reanalyzed as belonging to a phrase above CP.

This proposal has an interesting consequence for the grammaticalization of GEs. While most research on GEs assumes that grammaticalization from lexical material to pragmatic markers occurred at some early point in the history of English, this analysis suggests that only the semantically bleached GEs that no longer function propositionally have grammaticalized, and that this change is much more recent. Recall that in Table 3.13 only the more recent TEA contained semantically bleached GEs.

With respect to EPs, when functioning as the main verb and subject of an utterance, the set of epistemic/doxastic, complement taking, matrix verbs that co-occur with first person pronominal subjects (e.g., *I think, I guess, I suppose*) trigger the same implicature as their corresponding, reanalyzed pragmatic markers express conventionally: weak commitment to what follows. Consider (12).

- (12) A: Where is the cat?  
 B: I think that she is hiding under the couch.

Speaker A is seeking information about the location of the cat (let's call her Jam). Speaker A expects a response from speaker B about Jam's location. B does not give a direct response to A's question. Instead B informs A about one of B's beliefs (the belief that Jam is hiding under the couch). Given Grice's (1975) Maxim of Relevance, A interprets B's response as necessarily relevant to A's question: B must not know for certain the whereabouts of Jam, otherwise B's response would have been "she is hiding under the couch". Thus, main clauses like this (i.e., early stage, lexical elements



along Thompson and Mulac's [1991] cline of grammaticalization) already implicate the weak commitment meaning that is conventionalized by (advanced stage) EPs.

Taken together these two conjectures add up to essentially the same trajectory of change that Traugott (1995:15) suggests for pragmatic markers: "syntax via pragmatic strengthening in discourse > syntax with a different function." The system envisioned here differs from Traugott's in two fundamental ways. First, the change happens through abrupt reanalysis. Second, pragmatic markers are syntactic objects above CP. What remains the same is that both stages belong to syntax proper and reanalysis is driven by the conventionalization of inherently implicated meanings.

## 5.3 Implications

In the short sections that follow I discuss some of the major implications of this thesis. First, I discuss the consequences for grammaticalization theory and second, I consider the significance for variationist work on pragmatic variables.

### 5.3.1 Implications for Grammaticalization Theory

This thesis has argued against some of the core tenets of the standard view of grammaticalization theory (e.g., Bybee, Perkins, and Pagliuca 1994; Heine 2003; Heine, Claudi and Hünemeyer 1991; Hopper and Traugott 1993; Lehmann 1982; Traugott 1982, 1995). In particular, I have argued that the development of GEs and the development of EPs did not occur gradually. That said, some recent research from grammaticalization theorists has come to a very similar conclusion as I have. In particular, Heine (2014:1205–1206) following Kaltenböck, Heine, and Kuteva (2011) argues on conceptual grounds that the development of pragmatic markers is not the result of grammaticalization or pragmatization, but rather a spontaneous process of "cooptation, which has the effect that information units such as clauses, phrases,

or words are taken from the domain of sentence grammar and deployed for the purposes of discourse organization.” Heine (2014) and Kaltenböck, Heine, and Kuteva’s (2011) proposal is situated within a theory of language that includes a Discourse Grammar composed of a Sentence Grammar and Thetical Grammar: the development of pragmatic markers involves the cooptation by Thetical Grammar of Sentence Grammar elements during discourse.<sup>20</sup> Evidence for instantaneous cooptation come from spontaneously created thetical material. Heine (2014) suggests that all of the options in (13), while functioning as theticals, are an open class (i.e., not formulaic) composed of elements from Sentence Grammar.

- (13) This may {*it need hardly be said/and this is not really surprising/would you believe it/if you please forgive me saying that/...*} lead to compromise over the patient’s best medical treatments to promote personal and commercial interests.

(Heine 2014:1221)

Material that is frequently coopted “may subsequently develop from an instantaneous to formulaic thetical” via constructionalization (Heine 2014:1223). While the specifics differ, this proposal is in line with what I have just proposed: lexical material within a CP (i.e., Sentence Grammar) that expresses particular implicatures (i.e., coopted instantaneous thetical) has been reanalyzed (i.e., constructionalization) as pragmatic markers (i.e., formulaic theticals) that are located above CP (Thetical Grammar). While Heine (2014) and Kaltenböck, Heine, and Kuteva (2011) recognize that some part of the development of pragmatic markers is abrupt, they maintain that the constructionalization process is gradual. From the perspective of this thesis, such gradualness is an empirical question and the evidence presented herein sug-

---

<sup>20</sup>Note that the separation of these components is similar to Lewis (1970) and the recent generative work on pragmatic markers that has followed in that vein. This (perhaps unexpected) concurrence of two distinct theoretical approaches to language is worth celebrating!

gests that to the extent that Kaltenböck, Heine, and Kuteva's (2011) cooptation and constructionalization are different, both are abrupt, not gradual.

Whether or not one agrees that the evidence I present undermines the gradualness assumption, I would like to challenge all researchers working on grammaticalization phenomena. If we are asking quantitative questions (e.g., about gradualness vs. abruptness), we must use quantitative methods. In particular, we must use quantitative analysis that is accountable (see my discussion in §4.2). As such, the methods of variationist sociolinguistics, refined for the last forty years, are best suited to such analyses.

### 5.3.2 Implications for Variationist Work on Pragmatic Variables

While there has indeed been less research on variation at the level of pragmatics from a variationist perspective, a consistent theme is that this variation must be given special treatment, different from other domains of language (Pichler 2010). However, I hope to have shown that both methodologically and theoretically, pragmatic variables can be treated in just the same way as phonological, morphological, lexical, or syntactic variables. Methodologically, if researchers follow the main tenets of variationist research (e.g., an appropriate circumscription of the variable context, conforming to the principle of accountability, and quantitative hypothesis testing) their analyses will be replicable, familiar, and accountable. Theoretically, pragmatic variables can be treated just like lexical and morphosyntactic variables, all syntactic objects belonging to the lexicon.

## 5.4 Next Steps

While this thesis has focussed on issues specifically related to the development of pragmatic markers in Ontario English, there are a number of future directions that

follow naturally from this work.

### Variables Beyond GEs and EPs

The EOE data is ripe for critical, real-time analysis of the multitude of variables that have been previously examined in the TEA and elsewhere. These include grammatical variables such as relativizers (D'Arcy and Tagliamonte 2010), modals (Tagliamonte and D'Arcy 2007b; Tagliamonte and Denis 2014), complementizers (Tagliamonte 2013), future temporal reference (Tagliamonte and Denis 2008; Denis and Tagliamonte 2014b), the genitive (Jankowski and Tagliamonte 2014), possessive verbs (Tagliamonte, D'Arcy, and Jankowski 2010), comparative complementizers (Brook 2014), non-standard agreement (Tagliamonte 2001), and adverb placement (Waters 2013); pragmatic variables such as utterance final tags (Denis 2013; Denis and Tagliamonte 2014a; Tagliamonte 2006b), quotatives (Tagliamonte and D'Arcy 2007a, 2009), and intensifiers (Tagliamonte 2008); and sociophonetic variation including the Canadian Shift (Roeder and Jarmasz 2010). Such a *longue durée* view can help to resolve unanswered questions in sociolinguistics, such as what exactly is the relation between language change and the sociolinguistic meaning of variables, how do innovations arise, and what social motivations are most important in the development of a speech community's grammar.

A second empirical extension of this work is to work toward an even better real-time comparison of the data. In an optimal world, this thesis would have been able to examine both earlier vernacular data that represents the Toronto speech community specifically and newer vernacular data from Belleville, Eastern Ontario, and Niagara region. A first step that is immediately available is to examine Belleville, as data from both time periods is housed at the University of Toronto LVC Lab (Tagliamonte and Denis 2014). A collection of synchronic data from Eastern Ontario and Niagara in addition to the diachronic data gathered for this dissertation would be a unique

complement to similar real-time data from Northern Ontario (Tagliamonte and Denis 2014; Tagliamonte 2014). Looking outside of Ontario, the EOE influence spread westward across Canada with Loyalist transplantation from Southern Ontario all the way to Vancouver Island (Denis 2009; Denis and D'Arcy 2014). Comparative analysis of these features through diachrony will further illuminate the history of Canadian English.

### **Pragmatic Markers: The Social and the Structural**

Other logical extensions of this work relate to pragmatic markers more specifically. One tangential finding in this thesis that I wish to make explicit now is that pragmatic markers are of great dialectological importance. By virtue of the sheer number of variants of both GEs and EPs available to speakers, these systems are ripe for developing local social meanings in speech communities. It's no surprise that many pragmatic markers have become enregistered local, regional features. For example, the GE *n'at* ('and that') is recognized as a feature of Pittsburgese (Johnstone 2009:171–172) just in the same way as the utterance-final tag *eh* is a stereotypical feature of Canadian English (Denis 2013), utterance-final *like* is a stereotype of Irish English (Diskin 2014), and the EP *I reckon* stereotypes Australian English (Rodríguez Louro and Harris 2013).<sup>21</sup> Likewise, *I guess* was used in early Canadian literature as a way of linguistically stereotyping Loyalist settlers (often referred to as 'yankees' by British authors, emphasizing their American heritage), as in (14) from Susanna Moodie's memoir *Roughing it in the Bush*.

- (14) The man turned his knowing, keen eye upon me, and smiled, half-humorously, half-maliciously, as he said, "You were raised in the old country, *I guess*; you have much to learn, and more, perhaps, than you'll like to know, before the

<sup>21</sup>*I reckon* may also be associated with southern American speech <<http://grammatically-speaking.blogspot.ca/2008/06/well-i-reckon-so.html>>.

winter is over.”

(Susana Moodie, *Roughing it in the Bush*, 1852/1871)

The use of *guess* was also a favourite pet peeve of early Canadian English pedant Rev. A. Constable Geikie. His 1857 address to the Canadian Institute included the following discussion:

When an Englishman speaks at random or without sufficient authority, he guesses. When he expresses an opinion, he thinks. Guess and think are not synonyms, but refer to two opposite states of mind. Far otherwise is it in the neighbouring republic, and with too many here; for, with Americans and their imitators, guess and think have an identical signification. A “Clear-grit” guesses that the person beside him who does not spit on the floor, is a tory and a contemptible aristocrat, while a tobacco-moistening “Hoosier” guesses, and for like reasons, that a Boston merchant must be a federalist. Now if they only knew it, neither of these discerning and refined individuals guesses at all. Contrariwise, each feels confident in the matter pronounced upon. The general conduct of the persons of whom they thus judge, together with the subdued action of their salivary glands, has satisfied both that the political tendencies of the others must be the antithesis of their own. They are in no uncertainty, and a guess is impossible. The ordinary American use of this word justly subjects its users to ridicule, unless the precision which our English tongue once boasted of is no longer a feature worth preserving.

(Rev. A. Constable Geikie, *Canadian English*)

While I have mostly focussed on the similarities across the speech communities considered here in an effort to investigate change, a serious considerations of the dialectological differences would certainly prove fruitful. For example, while Appendix A lists one hundred and fifteen different GE types that appear in EOE and TEA, in an analysis of GEs in York, UK, eighty seven different types were found, many of which were not found in Ontario (Denis 2011:63). The potential of pragmatic markers to be indicators, markers and stereotypes of dialects, further illustrates that pragmatic variation should be treated just like any other sociolinguistic feature.

Another avenue of research in a decidedly different direction is a serious theoretical examination of the syntax and semantics of GEs, EPs, and other pragmatic markers. Indeed, the literature still has no satisfying answer to the question ‘what is a pragmatic marker?’ (Brinton 1996:32). I have suggested in this chapter that pragmatic markers are syntactic units that form dependencies with functional heads above CP. Just exactly which functional heads those are and what those dependencies look like remains to be seen though research in this direction is beginning to elucidate these questions (Bayer and Obenauer 2011; Lam 2014; Wiltschko and Heim 2014).

### **As Above, So Below**

Finally, let us return to the start of this thesis. If what I propose in this thesis is an accurate description of the development of pragmatic markers, these findings should be replicable in other speech communities, with other languages, and with other pragmatic markers. If nothing else, I hope the present work will light a spark for future work of this sort.

# Appendix A

## GEs Appendix

Table A.1: Complete list of adjunctive GE forms, by raw frequency in EOE

GE	General Type	BLV	EON	NIA
and a few things like that	thing	0	1	0
and all	other	5	7	0
and all manner of things	thing	0	0	1
and all that	that	6	3	4
and all that kind of	that	0	0	1
and all that kind of stuff	stuff	2	0	1
and all that kind of thing	thing	4	0	2
and all that sort of thing	thing	1	0	0
and all that stuff	stuff	2	0	0
and all that type	that	0	0	1
and all that type of thing	thing	1	0	4
and all the rest of it	other	0	1	1
and all them	that	0	1	0
and all this	that	0	0	1
and all this kind of stuff	stuff	0	4	0
and all this sort of thing	thing	0	1	0
and all this stuff	stuff	0	0	1
and all those	that	1	0	0
and all those kind of things	thing	0	1	0
and all those things	thing	2	0	0
and anything	anything	0	0	1
and anything like that	anything	0	0	1
and associated things	thing	1	0	1
and different things like that	thing	0	0	2



Table A.2: Complete list of adjunctive GE forms, by raw frequency in EOE, *cont.*

GE	General Type	BLV	EON	NIA
and every	everything	0	0	1
and everything	everything	5	2	17
and everything else	everything	0	0	1
and everything like that	everything	3	0	1
and everything that goes with it	everything	0	0	1
and like	like	0	0	1
and like of that	like	5	0	1
and like that	like	1	0	4
and one thing and another	other	0	1	0
and one thing another	other	0	1	1
and other such things	thing	1	0	0
and so	so	0	1	0
and so forth	so	0	1	9
and so on	so	16	38	38
and so on like that	so	1	6	2
and somethings like that	something	0	1	0
and sort	other	0	0	1
and stuff	stuff	1	2	3
and stuff like that	stuff	0	4	4
and such like	like	0	2	0
and that	that	5	10	17
and that kind of thing	thing	3	0	0
and that sort of thing	thing	4	1	1
and that type of thing	thing	2	0	2
and the like of that	like	1	0	0
and the likes of that	like	0	1	0
and the whole bit	other	0	0	1
and these kind of things	thing	0	1	0
and things	thing	5	2	10
and things like that	thing	2	12	32
and things like this	thing	1	0	0
and things of that type	thing	1	0	0
and this	that	1	0	1
and this sort of thing	thing	4	0	0
and those things	thing	0	0	1
and various other people	other	1	0	0
and what have you	what	0	0	1
and whatever	what	0	1	0
and whatever else	what	1	0	0
and whatnot	what	1	0	6

Table A.3: Complete list of disjunctive GE forms, by raw frequency in EOE

GE	General Type	BLV	EON	NIA
or a something	something	0	1	0
or anything	anything	3	10	13
or anything else	anything	0	0	1
or anything like that	anything	6	8	12
or anything with it	anything	1	0	0
or different types like that	other	0	0	1
or nothing like that	nothing	0	1	0
or so	so	1	14	9
or so on	so	0	1	0
or some such thing	thing	0	0	2
or somebody	some	0	1	1
or someplace	some	0	1	0
or someplace like that	some	0	0	1
or something	something	10	31	16
or something around there	something	0	0	1
or something else	something	0	2	0
or something like	something	0	1	0
or something like that	something	13	26	15
or something like this	something	0	1	0
or something of that nature	something	1	0	0
or somewhere	some	0	1	0
or the likes of that	like	0	1	0
or things like that	thing	0	0	1
or what	what	0	3	0
or whatever	what	0	3	1
or whatnot	what	0	1	1

Table A.4: Complete list of connectorless GE forms, by raw frequency in EOE

GE	General Type	BLV	EON	NIA
all that	that	2	0	0
all that kind of thing	thing	1	0	0
all that stuff	stuff	1	1	0
all those	that	1	0	0
along there	other	0	0	1
around there	other	0	1	2
et cetera	etcetera	2	0	1
everything like that	everything	1	1	0
everything that went with it	everything	0	1	0
like that	like	2	0	3
nor that sort of thing	thing	1	0	0
nothing like that	nothing	0	0	2
so on	so	0	0	2
some such thing	thing	0	0	1
someplace	some	1	0	0
something like that	something	1	1	9
somewhere around there	some	0	0	1
somewhere in that vicinity	some	0	0	1
somewheres around there	some	0	0	1
somewheres in there	some	0	0	1
the like of that	like	0	1	0
things like that	thing	0	0	2
things of that type	thing	1	0	0
whatever	what	0	1	0
you name it	other	2	0	1

Table A.5: Overall distribution of all general extender types. Frequency of all GEs (N).

STUFF	25.7% (n=699)			
Form	Belleville	E. Ontario	Niagara	Toronto
and stuff	0.7(1)	0.9(2)	1.0(3)	17.3(376)
and stuff like that	0.0(0)	1.8(4)	1.4(4)	10.1(219)
<i>other</i>	3.5(5)	2.3(5)	0.7(2)	3.5(78)
SOMETHING	22.1% (n=626)			
Form	Belleville	E. Ontario	Niagara	Toronto
or something	7.1(10)	14.6(32)	5.5(16)	16.1(351)
or something like that	9.9(14)	13.2(29)	8.2(24)	6.4(140)
<i>other</i>	0.7(1)	1.4(3)	1.4(4)	0.4(9)
THING	11.9% (n=336)			
Form	Belleville	E. Ontario	Niagara	Toronto
and things	3.5(5)	0.9(2)	3.4(10)	2.0(44)
and things like that	2.1(3)	5.6(12)	12.0(35)	4.4(96)
<i>other</i>	20.0(28)	7.7(17)	18.1(51)	8.2(180)
WHAT	11.7% (n=332)			
Form	Belleville	E. Ontario	Niagara	Toronto
or whatever	0.0(0)	2.3(5)	0.3(1)	13.8(300)
and whatnot	0.7(1)	0.4(1)	2.4(7)	0.2(4)
<i>other</i>	0.7(1)	1.4(3)	0.4(1)	0.4(8)
EVERYTHING	7.4% (n=208)			
Form	Belleville	E. Ontario	Niagara	Toronto
and everything	3.5(5)	0.9(2)	6.2(18)	6.8(147)
and everything like that	2.8(4)	0.5(1)	0.3(1)	1.1(26)
<i>other</i>	0.0(0)	0.5(1)	0.7(2)	0.1(1)
ANYTHING	6.5% (n=183)			
Form	Belleville	E. Ontario	Niagara	Toronto
or anything	2.1(3)	4.5(10)	4.8(14)	4.0(86)
or anything like that	4.3(6)	3.7(8)	4.5(13)	1.8(40)
<i>other</i>	0.7(1)	0.0(0)	0.3(1)	0.1(1)
AND THAT	6.2% (n=176)			
Form	Belleville	E. Ontario	Niagara	Toronto
and that	4.3(6)	4.5(10)	6.2(18)	1.8(41)
and all that	9.2(13)	5.0(11)	1.7(5)	3.2(69)
<i>other</i>	0.0(0)	0.0(0)	0.0(1)	0.1(3)
SO	6.2% (n=175)			
Form	Belleville	E. Ontario	Niagara	Toronto
and so on	11.3(16)	17.8(39)	13.7(40)	1.7(36)
and so on like that	0.7(1)	2.7(6)	0.7(2)	0.0(0)
<i>other</i>	0.7(1)	7.3(16)	6.4(18)	0.0(0)
OTHER	3.0% (n=86)			
Form	Belleville	E. Ontario	Niagara	Toronto
<i>other</i>	12.1(17)	5.9(13)	9.9(29)	1.2(27)

# Bibliography

- Aaron, J. E. (2010). Pushing the envelope: Looking beyond the variable context. *Language Variation and Change*, 22(1):1–36.
- Aarts, B., Close, J., and Wallis, S. A. (2010). Recent changes in the use of the progressive construction in English. In Cappelle, B. and Wada, N., editors, *Distinctions in English grammar, offered to Renaat Declerck*, pages 148–167. Kaitakusha, Tokyo.
- Adger, D. (2006). Combinatorial variability. *Journal of Linguistics*, 42(3):503–530.
- Adger, D. (2007). Variability and modularity: A response to Hudson. *Journal of Linguistics*, 43(3):695–700.
- Adger, D. and Smith, J. (2005). Variation and the Minimalist Programme. In Cornips, L. and Corrigan, K., editors, *Syntax and variation: Reconciling the biological and the social*, pages 149–178. John Benjamins, Amsterdam/Philadelphia.
- Adger, D. and Smith, J. (2010). Variation in agreement: A lexical feature-based approach. *Lingua*, 120(5):1109–1134.
- Aijmer, K. (1985). What happens at the end of our utterances? The use of utterance-final tags introduced by *and* and *or*. In Togeby, O., editor, *Papers from the 8th Scandinavian Conference of Linguistics*, Copenhagen. Institut for Nordisk Philologie, Kopenhavens Universitet.

- Aijmer, K. (1997). *I think* – an English modal particle. In Swan, T. and Westvik, O. J., editors, *Modality in Germanic languages: Historical and comparative perspectives*, pages 1–48, Berlin. Mouton de Gruyter.
- Aijmer, K. (2002). *English discourse particles: Evidence from a corpus*. John Benjamins, Amsterdam.
- Aijmer, K. (2013). *Understanding pragmatic markers: A variational pragmatic approach*. Edinburgh University Press, Edinburgh.
- Alonso-Ovalle, L. (2006). *Disjunction in alternative semantics*. PhD thesis, University of Massachusetts Amherst.
- Andersen, H. (1972). Diphthongization. *Language*, 48(1):11–50.
- Anthony, L. (2011). Antconc (version 3.2.4m). Computer software.
- Anttila, A. (2002). Variation and optimality theory. In de Lacy, P., editor, *The Cambridge handbook of phonological theory*, pages 519–536. Cambridge University Press, Cambridge.
- Archive of Ontario (1987). Farm work and farm life in Ontario since 1890 oral history project. Ontario Government Record Series RG 16-200.
- Avis, W. S. (1967). Canadian English. In Avis, W. S., Drysdale, P. D., Gregg, R. J., and Scargill, M. H., editors, *Dictionary of Canadian English: The senior dictionary*. W. J. Gage Limited, Toronto.
- Baayen, H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press, Cambridge.
- Bailey, C.-J. N. (1973). *Variation and linguistic theory*. Center for Applied Linguistics, Arlington, Virginia.

- Bailey, G., Maynor, N., and Cukor-Avila, P. (1991). *The emergence of Black English: Text and commentary*. John Benjamins Publishing, Amsterdam/Philadelphia.
- Bailey, G., Wikle, T., Tinnery, J., and Sand, L. (1991). The apparent time construct. *Language Variation and Change*, 3(3):241–264.
- Bates, D., Maechler, M., and Bolker, B. (2011). *lme4: Linear mixed-effects models using Eigen and syntax*. R package version 0.999375-42.
- Bayer, J. and Obenauer, H.-G. (2011). Discourse particles, clause structure, and question types. *The Linguistic Review*, 28:449–491.
- Bayley, R. J. (2002). The quantitative paradigm. In Chambers, J. K., Trudgill, P., and Schilling-Estes, N., editors, *The handbook of language variation and change*. Blackwell, Malden.
- Beal, J. C., Corrigan, K. P., and Moisl, H. L. (2007). Taming digital voices and texts: Models and methods for handling unconventional synchronic corpora. In Beal, J. C., Corrigan, K. P., and Moisl, H. L., editors, *Creating and digitizing language corpora*, pages 1–16. Palgrave Macmillan, Hampshire.
- Biber, D., Finegan, E., Atkinson, D., Beck, A., Burges, D., and Burges, J. (1993). The design and analysis of the archer corpus: a progress report. In Kytö, M., Rissanen, M., and Wright, S., editors, *Corpora across the centuries: Proceedings of the First International Colloquium on English Diachronic Corpora*, pages 3–6. Rodopi, Amsterdam.
- Biberauer, T., Holmberg, A., Roberts, I., and Sheehan, M. (2010). *Parametric syntax: Null subjects in minimalist theory*. Cambridge University Press, Cambridge.
- Blechsmidt, E. (1977). *The beginnings of human life*. Springer-Verlag, New York.
- Blondeau, H. and Nagy, N. (2008). Subordinate clause marking in Montreal Anglophone French and English. In Meyerhoff, M. and Nagy, N., editors, *Social lines in*

- language – sociolinguistic and multilingual speech communities: Celebrating the work of Gillian Sankoff*. John Benjamins, Amsterdam/Philadelphia.
- Bloomfield, M. W. (1975). Canadian English and its relation to eighteenth century American speech. In Chambers, J. K., editor, *Canadian English: Origins and Structures*. Methuen, Toronto.
- Boersma, P. and Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, 32(1):45–86.
- Borer, H. (1984). *Parametric syntax*. Foris, Dordrecht.
- Boyce, G. (2008). *Belleville: A popular history*. Dundurn Press, Toronto.
- Brinton, L. J. (1996). *Pragmatic markers in English: Grammaticalization and discourse functions*. Mouton de Gruyter, Berlin.
- Brinton, L. J. (2008). *The comment clause in English: Syntactic origins and pragmatic development*. Cambridge University Press, Cambridge.
- Brinton, L. J. and Traugott, E. C. (2005). *Lexicalization in language change*. Cambridge University Press, Cambridge.
- Brook, M. (2014). Comparative complementizers in Canadian English: Insights from early fiction. *University of Pennsylvania Working Papers in Linguistics*, 20(2).
- Brown, P. and Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press, Cambridge.
- Bucholtz, M. (1999). “why be normal?”: Language and identity practices in a community of nerd girls. *Language in Society*, 28(2):203–223.
- Buchstaller, I., Rickford, J. R., Traugott, E. C., Wasow, T., and Zwicky, A. (2010). The sociolinguistics of a short-lived innovation: Tracing the development of quotative



- all across spoken and internet newsgroup data. *Language Variation and Change*, 22(2):191–219.
- Bybee, J. (1988). The emergent lexicon. *Chicago Linguistic Society*, 34:421–435.
- Bybee, J. (2003). Mechanisms of change in grammaticization: The role of frequency. In Joseph, B. D. and Janda, R. D., editors, *The handbook of historical linguistics*. Blackwell, Malden, Mass.
- Bybee, J., Perkins, R., and Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. University of Chicago Press, Chicago.
- Campbell, L. (2001). What's wrong with grammaticalization? *Language Sciences*, 23(2/3):113–161.
- Campbell, L. and Janda, R. (2001). Introduction: conceptions of grammaticalization and their problems. *Language Sciences*, 23(2/3):93–112.
- Campbell, M. F. (1966). *A mountain and a city: The story of Hamilton*. McClelland and Stewart Limited, Toronto/Montreal.
- Carnochan, J. (1914). *History of Niagara (in part)*. William Briggs, Toronto.
- Cedergren, H. (1973). *The interplay of social and linguistic factors in Panama*. PhD thesis, Cornell University, Ithaca, NY.
- Chafe, W. (1986). Evidentiality in English conversation and academic writing. In *Evidentiality: The linguistic coding of epistimology*, pages 261–273. Ablex, Norwood, N.J.
- Chambers, J. K. (1992). Dialect acquisition. *Language*, 68:673–705.

- Chambers, J. K. (1998). Canadian English: 250 years in the making. In Barber, K., editor, *The Canadian English dictionary*, pages ix–x. Oxford University Press, Toronto/Oxford.
- Chambers, J. K. (2002). Patterns of variation including change. In Chambers, J. K., Trudgill, P., and Schilling-Estes, N., editors, *The handbook of language variation and change*, pages 349–372. Blackwell, Malden.
- Chambers, J. K. (2009). *Sociolinguistic Theory: Revised Edition*. Blackwell, Malden, Mass.
- Channell, J. (1994). *Vague language*. Oxford University Press, Oxford.
- Cheshire, J. (1982). *Variation in an English Dialect: A Sociolinguistic Study*. Cambridge University Press, Cambridge.
- Cheshire, J. (2005a). Syntactic variation and beyond: Gender and social class variation in the use of discourse-new markers. *Journal of Sociolinguistics*, 9(4):479–507.
- Cheshire, J. (2005b). Syntactic variation and spoken language. In *Syntax and Variation: Reconciling the Biological and the Social*, pages 81–106. John Benjamins, Amsterdam/Philadelphia.
- Cheshire, J. (2007). Discourse variation, grammaticalisation and stuff like that. *Journal of Sociolinguistics*, 11(2):155–193.
- Cheshire, J., Kerswill, P., and Williams, A. (2005). Phonology, grammar, and discourse in dialect convergence. In Auer, P., Hinskens, F., and Kerswill, P., editors, *Dialect change, convergence and divergence in European languages*, pages 135–167. Cambridge University Press, Cambridge.
- Chomsky, N. (1957). *Syntactic structures*. Mouton & Co., The Hague.

- Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press, Cambridge.
- Chomsky, N. (1980). *Rules and representations*. Columbia University Press, New York.
- Chomsky, N. (1981). *Lectures on government and binding*. Foris, Dordrecht.
- Chomsky, N. (1986). *Barriers*. MIT Press, Cambridge, Mass.
- Chomsky, N. (1995). *The Minimalist Program*. MIT Press, Cambridge, Mass.
- Compton, R. (2011). Polysynthetic word-internal adjectives and verb-like adjectives in Inuit. In *Proceedings of the 2011 annual conference of the Canadian Linguistic Association*. Canadian Linguistic Association.
- Couch, W. T. (1939). Preface. In *These are our Lives*. The University of North Carolina Press, Chapel Hill.
- Cowper, E. and Hall, D. C. (2002). The syntactic manifestation of nominal feature geometry. In *Proceedings of the 2002 annual conference of the Canadian Linguistic Association*.
- Crerar, A. (1999). *Ties that bind: Farming, agrarian ideals, and life in Ontario, 1890–1930*. PhD thesis, University of Toronto.
- Croft, W. (1990). *Typology and universals*. Cambridge University Press, Cambridge.
- Cutler, C. A. (1999). Yorkville crossing: White teens, hip hop and African American English. *Journal of Sociolinguistics*, 3(4):428–442.
- D'Arcy, A. (2005). *Like: Syntax and Development*. PhD thesis, University of Toronto.
- D'Arcy, A. (2012). The diachrony of quotative: Evidence from New Zealand English. *Language Variation and Change*, 24(3):343–369.

- D'Arcy, A. and Tagliamonte, S. A. (2010). Prestige, accommodation and the legacy of relative *who*. *Language in Society*, 39(3):389–410.
- Davis, C. (2011). *Constraining interpretation: Sentence final particles in Japanese*. PhD thesis, University of Massachusetts, Amherst.
- Denis, D. (2009). Transmission and diffusion above the level of phonology: Evidence from Thunder Bay. Master's Forum paper, University of Toronto.
- Denis, D. (2010). Grammaticalization of general extenders in York English. Unpublished manuscript, University of Toronto.
- Denis, D. (2011). Innovators and innovations: Tracking the innovators of *and stuff* in York. *University of Pennsylvania Working Papers in Linguistics*, 17(2):61–70.
- Denis, D. (2013). The social meaning of *eh* in Canadian English. In Luo, S., editor, *Proceedings of the 2013 Canadian Linguistics Association*.
- Denis, D. and D'Arcy, A. (2014). Homogeneity, convergence, mega-trends, and stuff like that. Talk presented at New Ways of Analyzing Variation 43. Chicago, Illinois, October 24, 2014.
- Denis, D. and Tagliamonte, S. A. (2014a). Innovation, *right?* change, *you know?* Utterance-final tags in Canadian English. Manuscript to appear in 2015 in *Discourse-pragmatic variation and change*, Heike Pichler, editor.
- Denis, D. and Tagliamonte, S. A. (2014b). Stability out of grammaticalization? future temporal reference in North American English. Talk presented at the eighty-eighth annual meeting of the Linguistic Society of America. Minneapolis, Minnesota.
- Diertani, C. E. A. (2011). *Morpheme boundaries and structural change: Affixes running amok*. PhD thesis, University of Pennsylvania.

- Diewald, G. (2010). On some problem areas in grammaticalization theory. In Stathi, K., Gehweiler, E., and König, E., editors, *Grammaticalization: Current views and issues*. John Benjamins, Amsterdam/Philadelphia.
- Diewald, G. (2011). Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics*, 49(2):365–390.
- Dines, E. (1980). Variation in discourse—“and stuff like that”. *Language in Society*, 9(1):13–33.
- Dion, N. and Poplack, S. (2005). I’m like, I know, I do the same thing: Can minority speakers keep up with the Joneses? Paper presented at New Ways of Analyzing Variation 34, New York University, Oct. 2005.
- Diskin, C. (2014). ‘I sound Irish when I use *like!*’ Acquisition of discourse-pragmatic markers by non-native speakers of Irish English. Paper presented at Discourse-Pragmatic Variation and Change 2014. Newcastle University. Newcastle, UK. April 7–9, 2014.
- Dollinger, S. (2008). *New-dialect formation in Canada: Evidence from the English modal auxiliaries*. John Benjamins Publishing, Amsterdam/Philadelphia.
- Dubois, S. (1992). Extension particles, etc. *Language Variation and Change*, 4:179–203.
- Eckert, P. (1989). *Jocks and burnouts: Social categories and identity in the high school*. Teachers College Press, New York.
- Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, 41:87–100.
- Erman, B. (1995). Grammaticalization in progress: The case of *or something*. In Moen, I., Simonsen, H. G., and Lødrup, H., editors, *Papers from the XVth Scandinavian Con-*

- ference of Linguistics, Oslo, January 13–15 1995*, Oslo. University of Oslo, Department of Linguistics.
- Erman, B. (2001). Pragmatic markers revisited with a focus on *you know* in adult and adolescent talk. *Journal of Pragmatics*, 32(9):1337–1359.
- von Stechow, K. (1999). NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics*, 19(2):97–148.
- Fischer, K. (1998). Discourse particles, turn-taking, and the semantics-pragmatics interface. *Revue de Sémantique et Pragmatique*, 8:111–137.
- Fischer, K. (2000). *From cognitive semantics to lexical pragmatics: The functional polysemy of discourse particles*. Mouton de Gruyter, Berlin.
- Fischer, O. (2007). The development of English parentheticals: A case of grammaticalization? In Dollinger, S., Hüttner, J., Kaltenböck, G., and Lutzky, U., editors, *Tracing English through time. Explorations in language variation*. Braumüller, Vienna.
- Fortson, B. W. (2003). An approach to semantic change. In Joseph, B. D. and Janda, R. D., editors, *The Handbook of Historical Linguistics*. Blackwell, Oxford.
- French, O., editor (2006). *Heritage Atlas of Hastings County*. Wallbridge House Publishing, Belleville.
- Fronek, J. (1982). *Thing* as a function word. *Linguistics*, 20(9–10):633–654.
- Fruehwald, J. (2013). *The phonological influence on phonetic change*. PhD thesis, University of Pennsylvania, Philadelphia.
- Fruehwald, J., Gress-Wright, J., and Wallenberg, J. C. (2013). Phonological rule change: The constant rate effect. In Kan, S., Moore-Cantwell, C., and Staubs, R., editors, *Proceedings the fortieth annual meeting of the North East Linguistic Society*.

- Fruehwald, J. and Wallenberg, J. C. (2013). Optionality is stable variation is competing grammars. Paper presented at twenty-fifth Scandinavian Conference of Linguistics, Formal Ways of Analyzing Variation Workshop, Háskóli Íslands, May 15, 2013.
- Fryer, M. B. (1980). *King's men: The soldier founders of Ontario*. Dundurn Press Limited, Toronto/Charlottetown.
- Fuller Medina, N. (2012). Processing versus dialect specific grammars: Relativization in early african american english. Paper presented at Change and Variation in Canada 6, UQÀM/McGill University, June 2, 2012.
- Gardner, M. H., Denis, D., Brook, M., and Tagliamonte, S. A. (2013). Apparent time, real time, and the end of the s-curve: *be like* at the saturation point. Paper presented at New Ways of Analyzing Variation 42, Pittsburgh, Pennsylvania, Oct. 19, 2013.
- Givon, T. (1989). *Mind, code and context: Essays in pragmatics*. Psychology Press, Taylor & Francis, East Sussex.
- Goffman, E. (1967). *Interaction ritual: Essays on face-to-face behavior*. Aldine Publishing Co., Chicago.
- Grice, H. P. (1975). Logic and conversation. In Cole, P. and Morgan, J., editors, *Syntax and semantics 3: Speech acts*, pages 225–242. Seminar Press, New York.
- Gries, S. T. (2009). *Statistics for linguistics with R: A practical introduction*. Walter de Gruyter, Berlin.
- Günthner, S. and Mutz, K. (2004). Grammaticalization vs. pragmaticalization? the development of pragmatic markers in German and Italian. In Bisang, W., Himmelmann, N. P., , and Wiemer, B., editors, *What makes grammaticalization? A look from its fringes and its components*. Mouton de Gruyter, Berlin.

- Haegeman, L. (1990). Non-overt subjects in diary contexts. In Mascaró, J. and Nespor, M., editors, *Grammars in progress: GLOW Essays for Henk van Riemsdijk*. Foris, Dordrecht.
- Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language*, 10:41–53.
- Harkness, J. G. (1946). *Stormont, Dundas, and Glengarry: A history 1784–1945*. Mundy-Goodfellow Printing Co. Limited, Oshawa.
- Harvie, D. (1998). Null subjects in English: Wonder if it exists? *Cahiers Linguistiques D'Ottawa*, 26(1):15–26.
- Haspelmath, M. (1998). Does grammaticalization need reanalysis? *Studies in Language*, 22:49–85.
- Hastings County Historical Society (1975). Belleville oral history archive. City of Belleville, Ontario, Canada.
- Heffernan, K. (2007). *Phonetic distinctiveness as a sociolinguistic variable*. PhD thesis, University of Toronto.
- Heim, I. and Kratzer, A. (1998). *Semantics in generative grammar*. Blackwell, Malden, Mass.
- Heine, B. (1997). *Possession: Cognitive sources, forces, and grammaticalization*. Cambridge University Press, Cambridge.
- Heine, B. (2003). Grammaticalization. In Joseph, B. D. and Janda, R. D., editors, *The handbook of historical linguistics*. Blackwell, Malden, Mass.
- Heine, B. (2014). On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics*, 51(6):1205–1247.



- Heine, B., Claudi, U., and Hünnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. University of Chicago Press, Chicago.
- Heine, B. and Kuteva, T. (2005). *Language contact and grammatical change*. Cambridge University Press, Cambridge.
- Henry, A. (1995). *Belfast English and Standard English: Dialect variation and parameter setting*. Oxford University Press, Oxford.
- Herring, S. C. (1991). The grammaticalization of rhetorical questions in Tamil. In Traugott, E. C. and Heine, B., editors, *Approaches to grammaticalization, volume I: Focus on theoretical and methodological issues*, pages 253–284. John Benjamins, Amsterdam/Philadelphia.
- Himmelmann, N. P. (2004). Lexicalization and grammaticalization: Opposite or orthogonal? In Bisang, W., Himmelmann, N. P., and Wiemer, B., editors, *What Makes Grammaticalization? A Look from its Fringes and its Components*. Mouton de Gruyter, Berlin.
- Hoffmann, S. (2004). A tentative proposal: Grammaticalization by analogy? In Lindquist, H. and Mair, C., editors, *Corpus approaches to grammaticalization in English*. John Benjamins, Amsterdam/Philadelphia.
- Hooper, J. B. (1975). On assertive predicates. In Kimball, J. P., editor, *Syntax and semantics*. Academic Press, New York.
- Hopper, P. J. (1987). Emergent grammar. In Tomasello, M., editor, *The new psychology of language: Cognitive and functional approaches to language structure*, pages 155–176. Lawrence Erlbaum, Mahwah, N.J.
- Hopper, P. J. (1991). On some principles of grammaticization. In Traugott, E. C. and

- Heine, B., editors, *Approaches to grammaticalization, volume I: Focus on theoretical and methodological issues*, pages 17–35. John Benjamins.
- Hopper, P. J. and Traugott, E. C. (1993). *Grammaticalization*. Cambridge University Press, Cambridge.
- Horn, L. R. (1989). *A natural history of negation*. University of Chicago Press, Chicago.
- Hothorn, T., Hornik, K., and Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, 15(3):651–674.
- Hubbell, A. F. (1950). *The pronunciation of English in New York City*. Columbia University Press, New York.
- Huddleston, R. and Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge University Press, Cambridge.
- Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, 61(1):23–62.
- Janda, R. D. (2001). Beyond “pathways” and “unidirectionality”: On the discontinuity of language transmission and the counterability of grammaticalization. *Language Sciences*, 23(2/3):265–340.
- Jankowski, B. (2005). “We’ve got our own little ways of doing things here”: Cross-variety variation, change and divergence in the English stative possessive. Generals Paper, University of Toronto.
- Jankowski, B. (2013). *A variationist approach to cross-register language variation and change*. PhD thesis, University of Toronto.
- Jankowski, B. and Tagliamonte, S. A. (2014). On the genitive’s trail: Data and method from a sociolinguistic perspective. *English Language and Linguistics*, 18(2):305–329.

- Johnson, D. E. (2009). Getting off the Goldvarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass*, 3(1):359–383.
- Johnstone, B. (2009). Pittsburghese shirts: Commodification and the enregisterment of an urban dialect. *American Speech*, 84(2):157–175.
- Jones, G. M. and Schieffelin, B. B. (2009). Enquoting voices, accomplishing talk: Uses of *be+like* in instant messaging. *Language & Communication*, 29(1):77–113.
- Joseph, B. D. (2001). Is there such a thing as “grammaticalization?”. *Language Sciences*, 23(2/3):163–186.
- Joseph, B. D. (2003). Morphologization from syntax. In Joseph, B. D. and Janda, R. D., editors, *The handbook of historical linguistics*, pages 472–492. Blackwell, Malden, Mass.
- Joseph, B. D. (2004). Rescuing traditional (historical) linguistics from grammaticalization theory. In Fischer, O., Norde, M., and Perridon, H., editors, *Up and down the cline – the nature of grammaticalization*. John Benjamins, Amsterdam/Philadelphia.
- Kaltenböck, G. (2007). Spoken parenthetical clauses in English: A taxonomy. In Dehé, N. and Kavalova, Y., editors, *Parentheticals*, pages 25–52. John Benjamins, Amsterdam/Philadelphia.
- Kaltenböck, G. (2013). The development of comment clauses. In Aarts, B., Close, J., Leech, G., and Wallis, S., editors, *The verb phrase in English*, pages 286–317. Cambridge University Press, Cambridge.
- Kaltenböck, G., Heine, B., and Kuteva, T. (2011). On thetical grammar. *Studies in Language*, 35(4):848–893.
- Kärkkäinen, E. (2003). *Epistemic stance in English conversation: A description of interactional functions, with a focus on ‘think’*. John Benjamins, Amsterdam/Philadelphia.

- Karttunen, L. (1971). Implicative verbs. *Language*, 47(2):340–358.
- Kiesling, S. F. (1998). Men's identities and sociolinguistic variation: The case of fraternity men. *Journal of Sociolinguistics*, 2(1):69–99.
- Kiparsky, P. (1968). Linguistic universals and linguistic change. In Bach, E. and Harms, R., editors, *Universals in Linguistic Theory*, pages 171–202. Holt, Rinehart & Winston, New York.
- Kratzer, A. (2006). Decomposing attitude verbs. Paper presented at Honoring Anita Mittwoch on her 80<sup>th</sup> birthday at The Hebrew University of Jerusalem. July 4, 2006.
- Kratzer, A. and Shimoyama, J. (2002). Indeterminate pronouns: The view from Japanese. Paper presented at the 3rd Tokyo Conference on Psycholinguistics.
- Kroch, A. S. (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change*, 1(3):199–244.
- Kroch, A. S. (1994). Morphosyntactic variation. In *Proceedings of the thirtieth annual meeting of the Chicago Linguistics Society*, volume 2, pages 180–201.
- Kroch, A. S. (2001). Syntactic change. In Baltin, M. and Collins, C., editors, *The handbook of contemporary syntactic theory*. Blackwell, Malden, Mass.
- Laberge, S. (1978). *Étude de la variation des pronoms sujets définis et indéfinis dans le français parlé à Montréal*. PhD thesis, Université de Montréal.
- Labov, W. (1963). The social motivation of a sound change. *Word*, 29:273–309.
- Labov, W. (1966a). The linguistic variable as a structural unit. *Washington Linguistic Review*, 3:4–22.
- Labov, W. (1966b). *The social stratification of English in New York City*. Center for Applied Linguistics., Washington.

- Labov, W. (1972). *Sociolinguistic Patterns*. Blackwell, Malden, Mass.
- Labov, W. (1994). *Principles of linguistic change, volume I: Internal factors*. Blackwell, Malden, Mass.
- Labov, W. (2001). *Principles of linguistic change, volume II: Social factors*. Blackwell, Malden, Mass.
- Labov, W. (2007). Transmission and diffusion. *Language*, 83(2):344–387.
- Labov, W. (2012). *Principles of linguistic change, volume III: Cognitive and cultural factors*. Blackwell, Malden, Mass.
- Labov, W., Rosenfelder, I., and Fruehwald, J. (2013). One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language*, 89(1):30–65.
- Lam, Z. W.-M. (2014). A complex FORCEP for speaker- and addressee-oriented discourse particles in Cantonese. *Studies in Chinese Linguistics*, 35(2):61–80.
- Lam, Z. W.-M., Thoma, S., and Wiltschko, M. (2013). The syntax of grounding. Poster presented at the 2013 annual meeting of the Canadian Linguistic Association. University of Victoria, Victoria, British Columbia. June 2, 2013.
- Lass, R. (1997). *Historical linguistics and language change*. Cambridge University Press, Cambridge.
- Lavandera, B. R. (1978). Where does the sociolinguistic variable stop? *Language in Society*, 7(2):171–182.
- Lehmann, C. (1982). *Thoughts on grammaticalization*. Lincom Europa, republished in 1995, Munich.

- Lemieux, M., Fontaine, C., and Sankoff, D. (1987). Quantifieur et marqueur de discours. In Sankoff, D., editor, *Diversity and diachrony.*, pages 381–390. John Benjamins, Amsterdam/Philadelphia.
- Levey, S. (2007). *The next generation: Aspects of grammatical variation in the speech of some London preadolescents.* PhD thesis, Queen Mary, University of London, London.
- Levey, S. (2012). General extenders and grammaticalization: Insights from London preadolescents. *Applied Linguistics*, 33(3):257–281.
- Levin, M. (2013). The progressive verb in modern American English. In Aarts, B., Close, J., Leech, G., and Wallis, S., editors, *The verb phrase in English*, pages 187–216. Cambridge University Press, Cambridge.
- Levitin, D. J. (2014). *The organized mind: Thinking straight in the age of information overload.* Allen Lane/Penguin, Toronto.
- Lewis, D. (1970). General semantics. *Synthese*, 22(1):18–67.
- Macaulay, R. (2002). Discourse variation. In Chambers, J. K., Trudgill, P., and Schilling-Estes, N., editors, *The handbook of language variation and change.* Blackwell, Malden, Mass.
- MacDonell, J. A. (1893). *Sketches illustrating the early settlement and history of Glengarry in Canada.* Wm. Foster, Brown & Co., Montreal.
- Mair, C. (2004). Corpus linguistics and grammaticalization theory: Statistics, frequency, and beyond. In Lindquist, H. and Mair, C., editors, *Corpus approaches to grammaticalization in English.* John Benjamins, Amsterdam.
- Mair, C. (2006). *Twentieth-century English: History, variation, and standardization.* Cambridge University Press, Cambridge.

- Margerie, H. (2007). From downgrading to (over) intensifying: a pragmatic study of English and French. In *Explorations in pragmatics: Linguistic, cognitive, and intercultural aspects*, pages 287–311. Mouton de Gruyter, Berlin.
- Massam, D., Starks, D., and Ikiua, O. (2006). On the edge of grammar: Discourse particles in Niuean. *Oceanic Linguistics*, 45(1):191–205.
- Meillet, A. (1912). L'évolution des formes grammaticales. *Scientia (Revista di Scienza)* 12, 26(6):384–400.
- Meyerhoff, M. (1994). Sound pretty ethnic, eh? A pragmatic particle in New Zealand. *Language in Society*, 23(3):367–388.
- Milroy, L. (1980). *Language and Social Networks*. Basil Blackwell, Oxford.
- Motut, A. (2010). Merge over Move and the empirical force of economy in Minimalism. *Toronto Working Papers in Linguistics*, 33:1–54.
- Mufwene, S. (1995). The founder principle in creole genesis. *Diachronica*, 13:83–134.
- Mufwene, S. S., Rickford, J. R., Bailey, G., and Baugh, J. (1998). *African American English: Structure, history, and use*. Routledge, London.
- Mullan, K. (2010). *Expressing opinions in French and Australian English discourse: A semantic and interactional analysis*. John Benjamins, Amsterdam/Philadelphia.
- Nelles, C. M. (1996). *Niagara! The first generation of a new country, Canada*. Pandora, Victoria.
- Nevalainen, T., Raumolin-Brunberg, H., and Mannila, H. (2011). The diffusion of language change in real time: Progressive and conservative individuals and the time depth of change. *Language Variation and Change*, 23(1):1–43.

- Nevins, A. and Parrott, J. K. (2010). Variable rules meet impoverishment theory: Patterns of agreement leveling in English varieties. *Lingua*, 120(5):1135–1159.
- Newmeyer, F. J. (2001). Deconstructing grammaticalization. *Language Sciences*, 23(2/3):187–229.
- Norrby, C. and Winter, J. (2002). Affiliation in adolescents' use of discourse extenders. In *Proceedings of the 2001 conference of the Australian Linguistic Society*.
- O'Keeffe, A. (2004). 'Like the wise virgins and all that jazz': Using a corpus to examine vague categorization and shared knowledge. *Language and Computers*, 52(1):1–26.
- Östman, J.-O. (1982). The symbiotic relationship between pragmatic particles and impromptu speech. In Enkvist, N. E., editor, *Impromptu speech: A symposium. Papers contributed to a symposium on problems in the linguistic study of impromptu speech*, pages 147–177. Åbo Akademi, Åbo.
- Overstreet, M. (1999). *Whales, Candlelight and Stuff Like That*. Oxford University Press, Oxford.
- Overstreet, M. (2014). The role of pragmatic function in the grammaticalization of English general extenders. *Pragmatics*, 24(1):105–129.
- Overstreet, M. and Yule, G. (1997). On being inexplicit and stuff in contemporary American English. *Journal of English Linguistics*, 25(3):250–258.
- Overstreet, M. and Yule, G. (2002). The metapragmatics of 'and everything'. *Journal of Pragmatics*, 34(6):785–794.
- Pagliuca, W. (1994). Introduction. In *Perspectives on grammaticalization*, pages ix–xx. John Benjamins, Amsterdam.



- Palacios Martínez, I. M. (2011). "I might, I might go I mean it depends on money things and stuff". a preliminary analysis of general extenders in British teenagers' discourse. *Journal of Pragmatics*, 43(9):2452–2470.
- Palander-Collin, M. (1997). A medieval case of grammaticalization, *methinks*. In Rissanen, M., Kytö, M., and Heikkonen, K., editors, *Grammaticalization at work: Studies in the long-term developments of English*, pages 371–403. Mouton de Gruyter, Berlin.
- Paolillo, J. (2011). Independence claims in linguistics. *Language Variation and Change*, 23(2):257–274.
- Parvaresh, V., Tavangar, M., Rasekh, A. E., and Izadi, D. (2012). About his friend, how good she is, *and this and that*: General extenders in native Persian and non-native English discourse. *Journal of Pragmatics*, 44(3):261–279.
- Pesetsky, D. (2013). What is to be done? Plenary talk at 87<sup>th</sup> annual meeting of the Linguistic Society of America. Boston, Massachusetts. Jan. 4, 2013.
- Pichler, H. (2008). *A qualitative-quantitative study of negative auxiliaries in northern English dialect: I don't know and I don't think, innit?* PhD thesis, University of Aberdeen.
- Pichler, H. (2010). Methods in discourse variation analysis: Reflections on the way forward. *Journal of Sociolinguistics*, 14(5):581–608.
- Pichler, H. and Levey, S. (2010). Variability in co-occurrence of discourse features. *University of Reading Language Studies Working Papers*, 2:17–27.
- Pichler, H. and Levey, S. (2011). In search of grammaticalization in synchronic dialect data: General extenders in northeast England. *English Language and Linguistics*, 15(3):441–471.

- Pintzuk, S. (1999). *Phrase structures in competition: Variation and change in Old English word order*. Garland Pub., New York.
- Pintzuk, S. (2003). Variationist approaches to syntactic change. In Joseph, B. D. and Janda, R. D., editors, *The handbook of historical linguistics*. Blackwell, Malden, Mass.
- Poplack, S., editor (2000). *The English history of African American English*. Blackwell, Malden, Mass.
- Poplack, S. (2007). Foreword. In Beal, J. C., Corrigan, K. P., and Moisl, H. L., editors, *Creating and digitizing language corpora*, pages ix–xiii. Palgrave Macmillan, Hampshire.
- Poplack, S. and Malvar, E. (2007). Elucidating the transition period in linguistic change: The expression of the future in Brazilian Portuguese. *Probus*, 19(1):121–169.
- Poplack, S. and Meechan, M. (1998). How languages fit together in codemixing. *International Journal of Bilingualism*, 2(2):127–38.
- Poplack, S. and St-Amand, A. (2007). A real-time window on 19th-century vernacular French: The *Récits du français québécois d'autrefois*. *Language in Society*, 36(5):707–734.
- Poplack, S. and Tagliamonte, S. A. (2001). *African American Vernacular English in the diaspora*. Blackwell, Malden, Mass.
- Poplack, S. and Turpin, D. (1999). Does the FUTUR have a future in (Canadian) French? *Probus*, 11(1):133–164.
- Potts, C. (2005). *The logic of conventional implicatures*. Oxford University Press, Oxford.
- Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1985). *A comprehensive grammar of the English language*. Longman, London.

- R Development Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ramchand, G. C. (1997). Questions, polarity, and alternative semantics. In *Proceedings of NELS27*, pages 383–396, Amherst. GLSA.
- Rissanen, M. (1991). On the history of *that/zero* as object clause links in English. In Aijmer, K. and Altenberg, B., editors, *English corpus linguistics: Studies in honour of Jan Svartvik*. Longman, London.
- Rizzi, L. (1982). *Issues in Italian syntax*. Foris, Dordrecht.
- Rizzi, L. (1997). The fine structure of the left periphery. In Haegeman, L., editor, *Elements of grammar*. Kluwer, Dordrecht.
- Roberts, I. and Roussou, A. (2003). *Syntactic change: A minimalist approach to grammaticalization*. Cambridge University Press, Cambridge.
- Rodríguez Louro, C. and Harris, T. (2013). Evolution with an attitude: The grammaticalisation of epistemic/evidential verbs in Australian English. *English Language and Linguistics*, 17(3):415–443.
- Roeder, R. V. and Jarmasz, L.-G. (2010). The Canadian Shift in Toronto. *Canadian Journal of Linguistics*, 55(3):387–404.
- Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1(1):75–116.
- Sankoff, D. (1988). Sociolinguistics and syntactic variation. In Newmeyer, F. J., editor, *Linguistics: The Cambridge survey*, pages 140–161. Cambridge University Press, Cambridge.

- Sankoff, D. and Sankoff, G. (1973). Sample survey methods and computer assisted analysis in the study of grammatical variation. In Darnell, R., editor, *Canadian languages in their social context*, pages 7–64. Linguistic Research Inc., Edmonton.
- Sankoff, D., Tagliamonte, S. A., and Smith, E. (2012). Goldvarb lion: A multivariate analysis application for Macintosh. Computer software.
- Sankoff, D. and Thibault, P. (1981). Weak complementarity: Tense and aspect in Montreal French. In *Syntactic change*, pages 205–216. University of Michigan Press, Ann Arbor.
- Sankoff, D., Thibault, P., and Bérubé, H. (1978). Semantic field variability. In Sankoff, D., editor, *Linguistic variation: Models and methods*, pages 23–44. Academic Press New York.
- Sankoff, G. (1974/1980). Above and beyond phonology in variable rules. In Sankoff, G., editor, *The social life of language*, pages 81–93. University of Pennsylvania Press, Philadelphia.
- Sankoff, G. (2005). Cross-sectional and longitudinal studies in sociolinguistics. In Ammon, U., editor, *Sociolinguistics: An international handbook of the science of language and society*, pages 1003–13. Walter de Gruyter, Berlin.
- Sankoff, G., Thibault, P., Nagy, N., Blondeau, H., Fonollosa, M.-O., and Gagnon, L. (1997). Variation in the use of discourse markers in a language contact situation. *Language Variation and Change*, 9(2):191–217.
- Santorini, B. (1992). Variation and change in Yiddish subordinate clause word order. *Natural Language and Linguistic Theory*, 10(4):595–640.
- Schneider, E. W. (2002). Investigating variation and change in written documents.

- In Chambers, J. K., Trudgill, P., and Schilling-Estes, N., editors, *The handbook of language variation and change*, pages 67–96. Blackwell, Malden, Mass.
- Schwenter, S. A. and Cacoulios, R. T. (2008). Defaults and indeterminacy in temporal grammaticalization: the ‘perfect’ road to perfective. *Language Variation and Change*, 20(1):1–39.
- Smith, N. and Leech, G. (2013). Verb structures in twentieth-century British English. In Aarts, B., Close, J., Leech, G., and Wallis, S., editors, *The verb phrase in English*, pages 68–98. Cambridge University Press, Cambridge.
- Speas, P. and Tenny, C. (2003). Configurational properties of point of view roles. In Di Sciullo, A. M., editor, *Asymmetry in grammar, volume 1: Syntax and semantics*. John Benjamins, Amsterdam/Philadelphia.
- Speyer, A. (2010). *Topicalization and stress clash avoidance in the history of English*. de Gruyter Mouton, Berlin.
- Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy*, 25:701–725.
- Stenström, A.-B., Andersen, G., and Hasund, I. K. (2002). *Trends in teenage talk: Corpus compilation, analysis and findings*. John Benjamins, Amsterdam.
- Stubbe, M. and Holmes, J. (1995). *You know, eh* and other exasperating expressions: An analysis of social and stylistic variation in the use of pragmatic particles in a sample of New Zealand English. *Language and Communication*, 15(1):63–88.
- Tagliamonte, S. and Roberts, C. (2005). So weird; so cool; so innovative: The use of intensifiers in the television series *Friends*. *American Speech*, 80(3):280–300.
- Tagliamonte, S. A. (2001). *Come/came* variation in English dialects. *American Speech*, 76(1):42–61.

- Tagliamonte, S. A. (2002). Comparative sociolinguistics. In Chambers, J. K., Trudgill, P., and Schilling, N., editors, *The handbook of language variation and change*. Blackwell, Malden, Mass.
- Tagliamonte, S. A. (2005). So who? Like how? Just what? Discourse markers in the conversations of young Canadians. *Journal of Pragmatics*, 37:1896–1915.
- Tagliamonte, S. A. (2006a). *Analysing sociolinguistic variation*. Cambridge University Press, Cambridge.
- Tagliamonte, S. A. (2006b). “So cool, right?”: Canadian English entering the 21st century. *Canadian Journal of Linguistics*, 51(2/3):309–331.
- Tagliamonte, S. A. (2007–2010). Directions of change in Canadian English. Research Grant. Social Sciences and Humanities Research Council of Canada (SSHRCC). #410-070-048.
- Tagliamonte, S. A. (2008). So different and pretty cool! Recycling intensifiers in Toronto, Canada. *English Language and Linguistics*, 12(2):361–394.
- Tagliamonte, S. A. (2010–2013). Transmission and diffusion in Canadian English. Research Grant. Social Sciences and Humanities Research Council of Canada (SSHRCC). #410-101-129.
- Tagliamonte, S. A. (2012). *Variationist sociolinguistics: Change, observation, interpretation*. Wiley-Blackwell, Malden, Mass.
- Tagliamonte, S. A. (2013). *Roots of English: Exploring the history of dialects*. Cambridge University Press, Cambridge.
- Tagliamonte, S. A. (2014a). Obsolescence and innovation in discourse-pragmatic change: The view from Canada. Plenary talk at Discourse-Pragmatic Variation and Change 2014. Newcastle University. Newcastle, UK. April 7–9, 2014.

- Tagliamonte, S. A. (2014b). System and society in the evolution of change: The view from Canada. In Green, E. and Meyer, C., editors, *Faces of English*. de Gruyter Mouton, Berlin.
- Tagliamonte, S. A. and Baayen, R. H. (2012). Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change*, 24(2):135–178.
- Tagliamonte, S. A., Chen, J., Chin, J., and Maddeaux, R. (2011). Sociolinguistics as experiential learning: Curriculum development in the teaching / learning interface. Paper presented at New Ways of Analyzing Variation 40, Georgetown University, Oct 29, 2011.
- Tagliamonte, S. A. and D'Arcy, A. (2007a). Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change*, 19(2):199–217.
- Tagliamonte, S. A. and D'Arcy, A. (2007b). The modals of obligation/necessity in Canadian perspective. *English World-Wide*, 28(1):47–87.
- Tagliamonte, S. A. and D'Arcy, A. (2009). Peaks beyond phonology: Adolescence, incrementation, and language change. *Language*, 85(1):58–108.
- Tagliamonte, S. A., D'Arcy, A., and Jankowski, B. (2010). Social work and linguistic systems: Marking possession in Canadian English. *Language Variation and Change*, 22(1):149–173.
- Tagliamonte, S. A. and Denis, D. (2008). Linguistic ruin? LOL. Instant messaging and teen language. *American Speech*, 83(1):3–34.
- Tagliamonte, S. A. and Denis, D. (2010). The *stuff* of change: General extenders in Toronto, Canada. *Journal of English Linguistics*, 38(4):335–368.

- Tagliamonte, S. A. and Denis, D. (2014). Expanding the transmission/diffusion dichotomy: Evidence from Canada. *Language*, 90(1):90–136.
- Tagliamonte, S. A. and Molfenter, S. (2007). How'd you get that accent? Acquiring a second dialect of the same language. *Language in Society*, 36:649–675.
- Tagliamonte, S. A. and Smith, J. (2005). No momentary fancy! The zero 'complementizer' in English dialects. *English Language and Linguistics*, 9(2):289–309.
- Terraschke, A. (2010). *Or so, oder so, and stuff like that*—general extenders in New Zealand English, German and in learner language. *Intercultural Pragmatics*, 7(3):449–469.
- Thompson, S. A. (2002). "Object complements" and conversation: Towards a realistic account. *Studies in Language*, 26(1):125–164.
- Thompson, S. A. and Mulac, A. (1991). A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In Traugott, E. C. and Heine, B., editors, *Approaches to grammaticalization, volume II*, pages 313–329. John Benjamins, Amsterdam/Philadelphia.
- Torres Cacoullos, R. and Walker, J. A. (2009a). On the persistence of grammar in discourse formulas: a variationist study of *that*. *Linguistics*, 47(1):1–43.
- Torres Cacoullos, R. and Walker, J. A. (2009b). The present of the English future: Grammatical variation and collocations in discourse. *Language*, 85(2):321–354.
- Traugott, E. C. (1982). From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In Lehmann, W. P., editor, *Perspectives on Historical Linguistics*. John Benjamins, Amsterdam.
- Traugott, E. C. (1988). Pragmatic strengthening and grammaticalization. In Axmaker,



- S., Jaisser, A., and Singmaster, H., editors, *Berkeley Linguistics Society 14 (Parasession on Grammaticalization)*, pages 406–416.
- Traugott, E. C. (1995). The role of the development of discourse markers in a theory of grammaticalization. *International Conference on Historical Linguistics XII*.
- Traugott, E. C. (2003). Constructions in grammaticalization. In Joseph, B. D. and Janda, R. D., editors, *A handbook of historical linguistics*, pages 624–647. Blackwell, Oxford.
- Traugott, E. C. and Dasher, R. B. (2002). *Regularity in semantic change*. Cambridge University Press.
- Traugott, E. C. and König, E. (1991). The semantics-pragmatics of grammaticalization revisited. In Traugott, E. C. and Heine, B., editors, *Approaches to grammaticalization, volume I: Focus on theoretical and methodological issues*, pages 189–218. John Benjamins Publishers.
- Traugott, E. C. and Trousdale, G. (2010). Gradience, gradualness and grammaticalization: How do they intersect? In Traugott, E. C. and Trousdale, G., editors, *Gradience, gradualness and grammaticalization*. John Benjamins, Amsterdam.
- Trousdale, G. (2008). Constructions in grammaticalization and lexicalization: Evidence from the history of a composite predicate construction in English. In Trousdale, G. and Gisborne, N., editors, *Constructional approaches to English grammar*, pages 33–67. Mouton de Gruyter, Berlin.
- Trudgill, P. (1972). Sex, covert prestige, and linguistic change in the urban British English of Norwich. *Language in Society*, 1(2):179–196.
- Trudgill, P. (1974a). Linguistic change and diffusion: description and explanation in sociolinguistic dialect geography. *Language in Society*, 3(2):215–246.

- Trudgill, P. (1974b). *The Social Differentiation of English in Norwich*. Cambridge University Press, Cambridge.
- Urmson, J. O. (1963). Parenthetical verbs. In Caton, C. E., editor, *Philosophy and ordinary language*, pages 220–240. University of Illinois Press, Urbana.
- Van Bogaert, J. (2011). *I think* and other complement-taking mental predicates: A case of and for constructional grammaticalization. *Linguistics*, 49(2):295–332.
- Van Herk, G. and Poplack, S. (2003). Rewriting the past: Bare verbs in the Ottawa Repository of Early African American Correspondence. *Journal of Pidgin and Creole Languages*, 18(2):231–266.
- Vincent, D. (1992). The sociolinguistics of exemplification in spoken French in Montréal. *Language Variation and Change*, 4(2):137–162.
- Vincent, D. and Sankoff, D. (1992). Punctors: A pragmatic variable. *Language Variation and Change*, 4(2):205–216.
- Wagner, S. E., Hesson, A., and Bybel, K. (2014a). Reliability, accountability and stuff like that: Quantifying pragmatic function in general extenders. Unpublished manuscript, Michigan State University.
- Wagner, S. E., Hesson, A., and Little, H. (2014b). The use of referential general extenders across registers. Manuscript to appear in 2015 in *Discourse-pragmatic variation and change*, Heike Pichler, editor.
- Wallace, W. S. (1914). *The United Empire Loyalists: A chronicle of the great migration*. Glasgow, Brook & Company, Toronto.
- Wallenberg, J. C. (2013). A unified theory of stable variation, syntactic optionality, and syntactic change. Paper presented at fifteenth Diachronic Generative Syntax conference. University of Ottawa. August 2, 2013.

- Waltereit, R. (2002). Imperatives, interruption in conversation, and the rise of discourse markers: a study of Italian *guarda*. *Linguistics*, 40(5):987–1010.
- Waltereit, R. (2006). The rise of discourse markers in Italian: A specific type of language change. In Fischer, K., editor, *Approaches to discourse particles*, pages 61–76. Elsevier, Amsterdam.
- Ward, G. and Birner, B. J. (1993). The semantics and pragmatics of *and everything*. *Journal of Pragmatics*, 19(3):205–214.
- Warner, A. (1982). *Complementation in Middle English and the methodology of historical syntax*. Croom Helm, London/Canberra.
- Waters, C. (2013). Transatlantic variation in English adverb placement. *Language Variation and Change*, 25(2):179–200.
- Watts, R. J. (2003). *Politeness*. Cambridge University Press, Cambridge.
- Weiner, E. J. and Labov, W. (1983). Constraints on the agentless passive. *Journal of Linguistics*, 19(1):29–58.
- Weinreich, U., Labov, W., and Herzog, M. (1968). Empirical foundations for a theory of language change. In Lehmann, W. P. and Malkiel, Y., editors, *Directions for historical linguistics*, pages 95–188. University of Texas Press, Austin.
- Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer, New York.
- Wiltschko, M. and Heim, J. (2014). The syntax of sentence peripheral discourse markers: A neo-performative analysis. Paper presented at Outside the clause: Form and function of extra-clausal constituents, University of Vienna, July 2014.
- Winter, J. and Norrby, C. (2000). “Set marking tags” and stuff. In *Proceedings of the 1999 Australian Linguistics Society*.

- Wolfram, W. A. (1969). *A sociolinguistic description of Detroit Negro Speech*. Center for Applied Linguistics, Washington.
- Youssef, V. (1993). Marking solidarity across the Trinidad speech community: The use of *an ting* in medical counselling to break down power differentials. *Discourse & Society*, 4(3):291–306.
- Zabbal, Y. (2004). A compositional semantics of the french expression N'IMPORTE QU. Generals paper, University of Massachusetts Amherst.
- Zelinsky, W. (1992). *The cultural geography of the United States. A revised edition*. Prentice Hall, Englewood Cliffs, NJ.
- Zhang, Q. (2005). A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society*, 34:431–466.