

LIKE:
SYNTAX AND DEVELOPMENT

by

Alexandra Faith D'Arcy

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

© Copyright by Alexandra Faith D'Arcy 2005

Like: Syntax and Development

Alexandra Faith D’Arcy
Doctor of Philosophy 2005
Department of Linguistics
University of Toronto

Discourse LIKE is one of the most salient features of present-day vernacular English. It is overtly stigmatized and associated with adolescents, where it is perceived as a crutch for lexical indecision (e.g., Diamond 2000; Siegel 2000). In the literature, LIKE is sometimes characterized as a “meaningless interjection” (OED) that can be used “grammatically anywhere” (Siegel 2002:64). Descriptions such as these suggest that LIKE is unconstrained, yet language, despite inherent variability, is rule-governed (see also Underhill 1988; Andersen 2001).

LIKE has received much attention in the pragmatic literature (e.g., Schourup 1983; Andersen 1997 et seq.), but it has never been investigated from a variationist perspective. Consequently, this dissertation presents an accountable analysis of LIKE in a large corpus of contemporary English. The hypothesis developed in this work is that LIKE is not random, but interacts with syntactic structure in regular and predictable ways. To address this issue, the variable context is circumscribed according to structural criteria and the analyses are embedded within current Minimalist Theory (e.g., Chomsky 1995 et seq.). Over 20,000 structurally defined contexts are examined, comprising data from 97 speakers between the ages of 10 and 87.

This method reveals that LIKE is 1) highly constrained by the syntax and 2) occurs in specific positions among speakers of all ages. Indeed, examination of language-internal constraints reveals that the community shares a single variable grammar for LIKE (Poplack & Tagliamonte 2001). This feature is shown to have developed gradually and systematically, arriving at its current state through regular processes of language change. Using the grammaticalization models proposed by Traugott (1997 [1995]) and Brinton (forthcoming), it is argued that after initially developing as a discourse marker, where it occurs clause-initially and links sequences of dialogue (Fraser 1988, 1990), LIKE then begins to enter syntactic structure, spreading to one maximal projection at a time.

Acknowledgements

For me, this degree was so much more than long hours and lack of sleep. It has also been about the lessons learned and the friendships forged. The simple truth is that the journey could not have been completed without a great deal of help and support from friends, family, and faculty.

I have been blessed with an incredibly devoted committee. Sali, Jack, and Alana invested a lot of time and effort with comments, criticisms, and questions that ultimately sparked many great moments. I'm sincerely appreciative of their honesty, even when they saw things differently from me.

Alana Johns allowed me space and when appropriate, she provided invaluable insights that cut to the heart of the matter. Alana, you were exactly what I needed. Jack Chambers, my mentor, inspiration, and friend, you invited me to U of T and thankfully I heeded the call. From you I learned to appreciate life both inside and outside the lab. To do your poetic soul justice, I hope that my letters are truly 'plus belle' after our time together. And Sali Tagliamonte ... who generously welcomed me into her office, her life, and her mind for four memorable years. Sali, the learning curve was steep, but the ride was spectacular. You taught me to think critically and to always search for the deeper meaning. I am eternally grateful for our time together as student and teacher and also as collaborators and friends. Thank you.

It was also my good fortune to have worked with Keren Rice during my time here. The most rewarding relationships are multifaceted. Keren, thank you for being my teacher, advisor, friend.

I must also say a big thank you to those I affectionately call “the great ones.” How I ended up writing a dissertation that focuses on adjunction instead of segmental representations, for example, is one of life’s great mysteries. My sincere appreciation goes to all of you for guiding me through unfamiliar territory. Gabriela Alboiu, Elizabeth Cowper, and Milan Rezac gave up their time and answered my questions. Email is a blessed thing. And Diane Massam. Diane had never worked with me before but when I approached with questions, she was unfailingly eager, helpful, and supportive. What a pleasure. Two people whose doors I knocked on with a complete and total lack of shame are Mike Barrie and Jila Ghomeshi. I appreciated not only their patience, but also their genuine interest and infectious enthusiasm. Others offered insightful comments on this work throughout its various stages. I thank Patricia Cukor-Avila, Ronald Macaulay, Miriam Meyerhoff, Suzanne Romaine, and Haj Ross. I am also indebted to Laurel Brinton and Elizabeth Traugott for the help and discussion they offered. Your work has been an inspiration. Thank you.

Nothing in life is truly accomplished alone. I’ll be eternally grateful for the friends who shared this experience with me. Thank you Abdel-Khalig Ali, Susana Bejar, Mohammad Haji-Abdolhosseini, Midori Hayashi, Jonathon Herd, Manami Hirayama, Arsalan Kahnemuyipour, and Jacek Panster. A special thank you to my friend Bettina Spreng for the shop talk, the laughter when I so desperately needed it, and of course, the kitty-sitting.

And lastly, my family. Your support came in many ways, but the most cherished were the simplest gestures — a phone call, a kind word, a hug — offered for no other reason than love. Words can never express what it meant to know that I was never alone, no matter the

physical distance between us. To my parents, Heather and Richard, how do I express my gratitude? Thank you for everything and I love you so much. To my other mother, Linda, thank you for believing in me. As much as I may have learned in school, I've learned more from you about life. To my brothers Graeme and Cory, their families, and my entire extended family, thank you. And Bill, I'm keeping my promise to relax, not work too hard, and enjoy life's little gifts. We miss you.

The most important contributor to all of this, however, has been Kevin Galloway. Sweet boy, you are my partner in every sense of the word and I am so thankful that you have shared this journey with me. Your support has been undying, unfailing, and unquestioning. That it encompasses every aspect of our relationship is everything to me. I only hope I've been able to communicate this and return it in kind. I love that you share my passion just because you are passionate about me, about us, about our life together. You have not only encouraged my dreams and cleared the path for me to pursue them, but more importantly, you have helped me to find my balance and to recognize the joy that is simply life. Kevin, I could not have done this without you. I am who I am because I have you. Thank you.

Contents

List of Tables	viii
List of Figures	ix
1 INTRODUCTION	1
1.1 Introduction.....	2
1.2 Research Questions.....	11
1.3 The (Socio) Linguistic Issues.....	12
1.3.1 Social factors	17
1.3.2 Linguistic factors.....	20
1.4 Methodology.....	24
1.4.1 Data	24
1.4.2 Method	26
1.5 Summary	31
2 NUMERICAL EXPRESSIONS	33
2.1 The data	37
2.2 Results.....	40
2.3 Discussion.....	47
3 GRAMMATICALIZATION	51
3.1 Discourse Markers in a Theory of Grammaticalization.....	52
3.2 Exemplification from English.....	60
3.2.1 <i>Soðlice</i>	60
3.2.2 <i>Indeed</i>	61
3.3 Discourse LIKE	63
4 THE CLAUSE-INITIAL CONTEXT	74
4.1 Methodology.....	76
4.1.1 Circumscribing the variable context.....	76
4.1.2 Coding and analysis	81
4.2 Results.....	84
4.2.1 Matrix clauses.....	85
4.2.1.1 Distributional analysis.....	85
4.2.1.2 Multivariate analysis.....	96
4.2.1.3 Summary.....	100
4.2.2 Subordinate clauses.....	101
4.2.2.1 CP subordinate clauses.....	105
4.2.2.2 TP subordinate clauses.....	108
4.3 Discussion.....	111

5	THE DETERMINER PHRASE	116
5.1	Methodology.....	119
5.1.1	The variable context.....	119
5.1.2	Coding and analysis	126
5.2	Results.....	130
5.2.1	Distributional analysis.....	130
5.2.2	Multivariate analysis.....	140
5.3	Discussion.....	146
6	PREDICATE ADJECTIVES	149
6.1	The data	152
6.2	Results.....	153
6.3	Discussion.....	162
7	THE VERBAL DOMAIN	166
7.1	Methodology.....	167
7.1.1	The variable context.....	167
7.1.2	Coding and analysis	171
7.2	Results.....	175
7.2.1	Distributional analysis.....	175
7.2.2	Multivariate analysis.....	195
7.2.3	The evidence from adverbs	198
7.3	Discussion.....	202
8	THE DEVELOPMENT OF DISCOURSE ‘LIKE’	204
8.1	The Development of Discourse LIKE.....	206
8.1.1	From whence it came	206
8.1.2	Ongoing development.....	208
8.1.3	Implications for grammaticalization.....	218
8.2	Social Embedding.....	221
8.3	Closing Remarks.....	227
	References	231

List of Tables

1.1 The sample	25
2.1 Overall distribution of numerically quantified expressions	40
2.2 Two independent multivariate analyses of the effects of age and sex on the use of LIKE and adverbial approximators	44
4.1 Overall distribution of LIKE on matrix CPs	85
4.2 Distribution of LIKE on matrix clauses according to position in turn	90
4.3 Distribution of LIKE on matrix clauses according to the status of the CP	92
4.4 Contribution of external and internal factors on the probability of LIKE on matrix CPs	97
4.5 Comparison of results for LIKE within the TP subordinate context	111
5.1 Distribution of LIKE in the DP context across apparent-time	130
5.2 Distribution of LIKE in the NP context across apparent-time	133
5.3 Distribution of LIKE according to the status of the DP	136
5.4 Distribution of LIKE according to the status of the DP II	137
5.5 Contribution of external and internal factors on the probability of LIKE in the DP context	141
6.1 Distribution of LIKE with predicate adjectives across apparent-time	154
6.2 Overall distribution of LIKE in PredAPs according to sex	158
6.3 Contribution of external and internal factors on the probability of LIKE in the PredAP context	160
7.1 Criteria used to determine the class of a monadic verb	174
7.2 Distribution of LIKE in the VP context over apparent-time	176
7.3 Distribution of LIKE across different surface string types in the VP	185
7.4 Distribution of LIKE according to verb type	187
7.5 Distribution of LIKE according to semantic criteria	189
7.6 Contribution of external and internal factors on the probability of LIKE in the VP context	196
7.7 Classification of adverbs	200
8.1 Summary of analyses	205
8.2 Summary of linguistic constraints on LIKE	213

List of Figures

2.1	Distributions of numerical modifiers across apparent-time	41
3.1	Grammaticalization of LIKE	63
4.1	Distribution of LIKE and other discourse markers with matrix CPs across apparent-time	86
4.2	Cross-tabulation of LIKE with matrix CPs by age and sex	88
4.3	The use of LIKE in four age groups according to sex: matrix CPs	98
4.4	Cross-tabulation of LIKE on matrix CPs: internal factors and age	99
4.5	Distribution of LIKE on CP subordinates across apparent-time	105
4.6	The use of LIKE in four age groups according to sex: subordinate CPs	108
4.7	Distribution of LIKE on TP subordinates across apparent-time	110
4.8	Distribution of LIKE in the clause-initial context across apparent-time	113
5.1	Cross-tabulation of LIKE in the DP context by age and sex	135
5.2	Cross-tabulation of status by age	136
5.3	Cross-tabulation of definiteness by age	139
5.4	Cross-tabulation of modification by age	139
5.5	Distribution of LIKE according to modification of the DP	144
6.1	Distribution of LIKE in PredAP contexts according to verb	155
6.2	Cross-tabulation of sex by age in the PredAP context	158
6.3	Cross-tabulation of sex by verb in the PredAP context	162
6.4	Time-line for LIKE in adjectival contexts	163
7.1	Comparison of the distribution of LIKE in the PredAP and VP contexts	177
7.2	Type of surface string in which LIKE occurs according to age	186
7.3	Type of verb with which LIKE occurs according to age	191
7.4	Cross-tabulation of LIKE by speaker sex and age	194
8.1	Generalization of LIKE across maximal projections apparent-time	209
8.2	The rise and development of LIKE as a pragmatic device	212
8.3	Distribution of LIKE in four contexts according to speaker sex	223

“It is only through an analysis of variation
that the reality and meaning of a norm can be established at all.”

Edward Sapir, 1938

Chapter 1 INTRODUCTION

Despite a relatively short history of systematic research, dating only to the early 1980s (Fraser 1990:384), the body of literature dealing with discourse markers in English is vast. Perhaps no individual form, however, has received as much attention as has LIKE, exemplified in (1).

- (1) a. I love Carrie. **LIKE**, Carrie's **LIKE** a little **LIKE** out-of-it but **LIKE** she's the funniest, **LIKE** she's a space-cadet. Anyways, so she's **LIKE** taking shots, she's **LIKE** talking away to me, and she's like, "What's wrong with you?" (3/T/f/18)¹
b. Well you just cut out **LIKE** a girl figure and a boy figure and then you'd cut out **LIKE** a dress or a skirt or a coat, and **LIKE** you'd colour it. (N/8/f/75)

Works on LIKE have spanned a number of topics, with issues ranging from grammaticalization (Meehan 1991; Romaine & Lange 1991; Buchstaller 2001), syntactic distributions and constraints (Ross & Cooper 1979; Underhill 1988; Andersen 1997, 1998, 2001), sociolinguistic distributions and frequencies (Andersen 1997 et seq.; Dailey-O'Cain 2000; Tagliamonte 2005), cross-linguistic comparisons (Hasund 2003), to pragmatic function (Schourup 1983; Miller & Weinert 1995; Andersen 1997 et seq.; Hasund 2003; etc.). It is this last aspect that has been most emphasized, with LIKE generally considered to mark either focus (e.g., Underhill 1988; Meehan 1991; Miller & Weinert 1995; Romaine &

¹ The parenthetical information following examples marks the sub-corpus the datum was extracted from and the individual speaker code, followed by the speaker's sex and age. The sample used for this research is outlined in section 1.4.1.

Lange 1991) or ‘loose talk’ (e.g., Andersen 1997 et seq; Sharifian & Malcolm 2003). This latter approach, couched in relevance-theoretic notions of contextual effects (Sperber & Wilson 1995), proposes that LIKE signals a mismatch between the propositional form of the utterance and the thought it represents.

What makes LIKE so interesting is its high text frequency among adolescents and young adults. It is one of the most salient features of present-day vernacular English, due in part to being overtly stigmatized by older speakers as a purported indicator of incoherence and uncertainty. Where did this particle come from? The ubiquity of discourse LIKE provides an ideal opportunity to track its development in a large corpus of contemporary data. In this dissertation, I address questions such as: Can LIKE really go “anywhere” syntactically? Are some contexts more frequent than others? If so, how can this be accounted for? Is there evidence for how LIKE has attained its current state of use (i.e., are the patterns of use random or have they developed systematically)?

1.1 INTRODUCTION

It is important to be clear about exactly which uses of LIKE are discursive as opposed to those that are not. Putatively pragmatically-motivated occurrences of LIKE as in (1) are distinct from its homonymous grammatical functions as lexical verb (2a), noun (2b), preposition (2c), conjunction (2d), adverb (2e), and suffix (2f).²

² LIKE as a preposition and conjunction derives from OE gelic while LIKE as a lexical verb derives from ME lician (Romaine & Lange 1991:271; also OED).

- (2)
- a. I don't really **LIKE** her that much. (2/m/f/12)
 - b. He grew up with the **LIKES** of the company of all great fighters. (N/4/m/60)
 - c. It looks **LIKE** a snail; it just is a snail. (I/@/f/19)
 - d. It felt **LIKE** everything had dropped away. (ITM/m/40)
 - e. I don't want to talk **LIKE** that. (I/7/m/35)
 - f. He said I went "(mumbling)" or something like stroke-**LIKE**. (N/H/f/31)

Discourse **LIKE** must also be differentiated from the etymologically-related — though functionally distinct — use of **LIKE** as a quotative complementizer (Romaine & Lange 1991), as in (3). In such instances, the collocation be + like forms a unit with the specialized role of demarcating sequences of constructed dialogue (e.g., Butters 1982; Blyth, Recktenwald & Wang 1990; Romaine & Lange 1991; Ferrara & Bell 1995; Tagliamonte & Hudson 1999; Dailey-O'Cain 2000; Cukor-Avila 2002; Buchstaller 2004; Tagliamonte & D'Arcy 2004, under review).

- (3)
- a. They'**RE LIKE**, "Okay we'll pay you now." (I/¢/m/21)
 - b. He'**S LIKE**, "You have to calm down." (2/i/f/19)
 - c. It skinned us. It **WAS LIKE**, "(whoosh)." (3/P/f/16)

Although it has been argued that both quotative be like and the discourse marking **LIKE** may operate as textual focus markers (e.g., be like: Blyth et al. 1990; Meehan 1991; Romaine & Lange 1991; Buchstaller 2001; **LIKE**: Underhill 1988; Meehan 1991; Miller & Weinert 1995), it is nonetheless apparent that the former functions as a verbal element, inflectable for both tense and agreement, while the latter cannot be classified according to lexical diagnostics (i.e., it has little or no semantic content and can generally be omitted without affecting the grammaticality of the utterance).

Moreover, the rise of LIKE as a discourse element long pre-dates the development of be like as a verb of quotation, which is believed to have evolved during the final quarter of the twentieth century (Butters 1982; Blyth et al. 1990; Romaine & Lange 1991), a chronology that has recently been substantiated by apparent-time evidence (Tagliamonte & D’Arcy under review). Indeed, as outlined by Romaine and Lange (1991:261), despite the shared origin of these forms from the original conjunction, each represents a distinct development in the history of English. Consequently, the quotative is excluded from the current analysis; the focus here is the discourse particle.

Discourse LIKE has garnered attention from linguists and grammarians for quite some time, and among the latter group, it is generally condemned as colloquial and vulgar (see, for example, Jespersen 1942:417). Despite the overt stigmatization of LIKE, the written record indicates that it has been performing pragmatic functions for well over two centuries. Attestations in the OED date back to the late eighteenth and early nineteenth centuries, where LIKE generally occurs in clause-final position, as in (4):

- (4) a. Father grew quite uneasy, **LIKE**, for fear of his Lordship’s taking offence.
(1778 F. Burney Evelina II. xxiii. 222)
- b. Of a sudden **LIKE**. (1801 tr. Gabrielli's Myst. Husb. III. 252)
- c. In an ordinary way **LIKE**. (1826 J. Wilson Noct. Ambr. Wks. 1855 I. 179)
- d. If your honour were more amongst us, there might be more discipline **LIKE**.
(1838 Lytton Alice ii. iii)
- e. “Why **LIKE**, it’s gaily nigh like to four mile **LIKE**.”
(1840-41 De Quincey Style ii. Wks. 1862 X. 224)
- f. Might I be so bold as just to ax, by way of talk **LIKE**, if [etc.].
(1870 E. Peacock Ralf Skirl. I. 112)
- g. He hasn’t passed his examinations **LIKE**... He has that Mr. Karkeek to cover him **LIKE**. (1911 A. Bennett Hilda Lessways i. vi. 49)
- h. “You’re a chauvinist,” Danny said. “Oh, yeah. Is that bad **LIKE**?”
(1961 New Statesman 22 Sept. 382/2)

The OED describes this use as parenthetical, where it can be glossed as ‘as it were’ or ‘so to speak’.³ As such, LIKE in this construction is clearly pragmatic, providing metalinguistic commentary on the preceding statement (see also Brinton (2005) regarding (I say)).

Grant and Dixon (1921:142) cite examples as in (5) from the early nineteenth century that exhibit uses identical to contemporary practice. In (5a), LIKE occurs with a numerical expression, and in (5b), it takes scope over the following determiner phrase:

- (5) a. The three mile diminished into **LIKE** a mile and a bittock. (Scott 1815)
 b. She asked my wife what was **LIKE** the matter wi’ her. (Wilson 1835-1840)

Consider also (6), from The English Dialect Dictionary (Wright 1902), where LIKE introduces adverbial phrases:

- (6) He would not go **LIKE** through that. They are **LIKE** against one another as it is.

Since the spoken vernacular has a tendency to be more innovative than the written standard, we can assume that discourse LIKE originated well before these attested examples and has deep roots in the history of English. This is all the more salient given the belief that LIKE in final position qualifies the preceding statement, but its use in other positions constitutes a ‘meaningless interjection’ (OED) and crucially, that this function originated in the United States (e.g., Andersen 2001:209). The difference between final LIKE and other positions is typically distinguished as the ‘traditional’ (and obsolescing) British pattern, where LIKE takes backward scope, in contrast to the ‘innovative’ American one where LIKE

³ It will be argued in Chapter 3 that the examples in (4) represent a criterial stage in the development of LIKE as a discourse marker.

takes forward scope (Andersen 2001:216; see also Meehan 1991:45; Romaine & Lange 1991:249).⁴

An obvious problem with this view is that discourse **LIKE** is used in the innovative ‘American’ way by older speakers in Britain, as exemplified in (7).⁵ If this use were strictly an American innovation, it would be difficult to account for its occurrence in regional British varieties, particularly as it is unlikely that elderly residents of Ayrshire, for example, would have adopted a stigmatized feature of the (North) American youth culture and made productive use of it.

- (7)
- a. I couldn’t stand it, **LIKE** I just couldn’t. (MPT/a/m/63)
 - b. Northeast there was always a little bit of road, **LIKE** it was my thinking bit of road. (MPT/#/f/81)
 - c. In those days **LIKE** there was very little traffic. (CLB/r/m/88)
 - d. That was **LIKE** the visitors and we says we would nae mind ken. (AYR/G/f/78)
 - e. We were doing **LIKE** a nature study. (PVG/d/f/62)
 - f. We were **LIKE** walking along that Agohill Road. (CLB/l/f/86)
 - g. We were **LIKE** ready to **LIKE** mutiny. (YRK/p/f/74)
 - h. They were just **LIKE** sitting waiting to die. (AYR/c/m/75)
 - i. They didn’t go **LIKE** to Ireland like they do nowadays. (YRK/k/f/87)
 - j. Tied with bits of rope. It was **LIKE** up, up and across. (MPT/z/m/78)

Further, discourse **LIKE** is not restricted to British (Miller & Weinert 1995; Andersen 1997 et seq.) and North American varieties (Pei 1973; Underhill 1988; Meehan 1991; Romaine & Lange 1991; Jucker & Smith 1998; Dailey-O’Cain 2000; Wolgemuth 2003), but

⁴ Following the tradition in the discourse-pragmatic literature, I use the notion of scope here in the very general sense of ‘modification domain’ (e.g., Underhill 1988; Romaine & Lange 1991; Andersen 1997 et seq.; Traugott 1997 [1995], 1997; Brinton 1996, 2005, forthcoming; etc.). I do not intend it in the narrow sense in which scope refers to the domain in which an operator can affect the referential interpretation of some expression (e.g., quantifiers and negative polarity, pronouns, other quantifiers) (see, for example, Reinhart 1978; Horstein 1995; Beghelli & Stowell 1997; Szabolsci 2001).

⁵ The data in (7) are from Tagliamonte (to appear); AYR = Ayrshire, CLB = Cullybackey, MPT = Maryport, PVG = Portavogie, and YRK = York.

is also prevalent in varieties such as Australian Aboriginal English (Sharifian & Malcolm 2003). If this feature were not already present in some form in the English dialects brought to the Australian continent by its settlers, it becomes difficult to account for such parallel and simultaneous developments across the English-speaking world in any linguistically plausible way.

Thus, both the historical record and synchronic facts contradict the notion that American English is the sole source of ‘innovative’ uses of discourse LIKE, despite claims that its roots stem from counterculture groups (i.e., jazz, cool, and beat) of 1960s New York (see Andersen 2001:216, and references therein). Rather, current usage has likely evolved from an earlier embryonic state (Trudgill 2002:41) that was transplanted during Britain’s period of colonial expansion to the New World. This is not to preclude the possibility of the transnational spread of linguistic innovations but rather to suggest that a more elaborated account is likely in order.

The question remains as to how discourse uses of LIKE arose in the first place. Specifically, what is the path by which LIKE came to develop pragmatic functions? It has been suggested that both quotative be like and discourse LIKE have evolved from the use of LIKE as a preposition (2c) and a conjunction (2d) (Meehan 1991; Romaine & Lange 1991; Andersen 2001; Buchstaller 2001). According to Romaine and Lange (1991:261), once LIKE began to function as a conjunction, it was able to take clausal complements. This is critical to the further evolution of this morpheme as a quotative complementizer, since in this function, be like introduces full sentential clauses (see (3) above). I will argue in Chapter 3 that the path from preposition and conjunction to discourse marker is more complex,

involving a series of shifts in its syntactic, semantic, and pragmatic status, an issue I return to in Chapter 8.

This type of trajectory, where a lexeme develops a new function while simultaneously undergoing semantic and pragmatic change, may be indicative of grammaticalization (e.g., Meillet 1948; Heine, Claudi & Hünnemeyer 1991; Traugott & Heine 1991; Hopper & Traugott 2003; etc.).⁶ Several analyses hold that this is precisely what has occurred with LIKE (Meehan 1991; Romaine & Lange 1991; Ferrara & Bell 1995; Andersen 2001). Brinton (1996; also 2005, forthcoming) supports this view, arguing that several discourse markers can be best understood within the framework of grammaticalization theory, a position also asserted in recent work by Traugott (1997 [1995], 2001, 2003) and Schiffrin (1992) for English, Lehti-Eklund (1990) for Swedish, and Matsumoto (1988) and Onodera (1995) for Japanese. As this view of grammaticalization entails a departure from the traditional functionalist perspective, which does not make reference to pragmatic inferencing, I specifically address in detail the notion of grammaticalization as it pertains to discourse markers in general, and to LIKE in particular, in Chapter 3.

⁶ Jespersen (1942:417), however, suggests that discourse LIKE has evolved from the derivational suffix, —like. As Romaine and Lange (1991:272-273, fn. 15) point out, such a cline does not constitute grammaticalization but would instead be an example of ‘degrammaticalization’ or ‘regrammaticalization’, wherein a morphological affix evolves into a free morpheme with an independent syntactic or discourse function.

At the same time, discourse LIKE is sometimes said to carry with it a referential core of comparison or similarity, embodying meanings such as ‘for example’, ‘as if’, or ‘approximately’ (e.g., Schourup 1983; Underhill 1988; Meehan 1991; Jucker & Smith 1998). This leads to another possibility: discourse uses may represent pragmatically conditioned lexical replacement. If this were so, and LIKE functioned as a substitute for ‘approximately’ etc., we would expect it to maintain its meaning in all contexts, yet in some cases, such as those in (1), LIKE has little to no semantic content (Underhill 1988; Meehan 1991; Buchstaller 2001). Though diminished content is a hallmark of grammaticalized forms (as demonstrated by the paradigmatic example of the adverbializing suffix *-ly*, originally *liche* ‘body’), such bleaching is atypical of lexical replacement.

In cases such as those in (8) (see also (5a)), where LIKE precedes a numerical expression, its status is controversial. In this context, it seems to carry an approximative meaning.

- (8) a. It could have taken you all day to go LIKE thirty miles. (N/©/f/76)
 b. You-know, it was LIKE a hundred and four [degrees]. (N/w/m/84)
 c. They were LIKE eighteen years old; they were kids. (N/s/f/52)

Some researchers note that when LIKE modifies a quantified element, it affects truth conditions (Andersen 1997 et seq.; Jucker & Smith 1998). Siegel (2002:48) points out that such a function is at odds with LIKE purely as a discourse marker, since markers tend to fall into the post-compositional pragmatic component, when the semantic relations that hold over propositions have already been encoded. Consequently, others such as Schourup (1983:39), Underhill (1988:234), and Biber et al. (1999:858) argue that when it occurs in

numerical contexts, LIKE functions adverbially and carries with it the full semantic force of modifiers such as ‘approximately’ or ‘roughly’. I concur with their position, and provide corroborating evidence for it in Chapter 2. From that point forward, I exclude LIKE from the quantitative analyses when it modifies quantified elements.

Despite the extensive attention this particle has received, there are many aspects of LIKE which remain unclear. Foremost among these is the developmental trajectory that has led to what appears to be the haphazard use of this form by younger speakers. As the Los Angeles Times put it, “To the common ear, the word seems just flung in” (Diamond 2000:2). Thus, it is perhaps not surprising that a recurrent theme in the literature is the syntactic freedom of discourse LIKE, where it is characterized by its ability to “occur grammatically anywhere in a sentence” (Siegel 2002:64; emphasis my own). This seemingly arbitrary distribution has led to the notion that LIKE indicates the general inarticulateness of adolescents. Indeed, Diamond (2000:2) cites one professor who states that contemporary youth are “so unsure of the words to use that LIKE eases them over the decision.”

Linguistic theory, however, presupposes a model of language that is systematic and rule-governed despite its inherent variability. It is therefore antithetical to the design of the language faculty that LIKE should be random, *ad hoc*, or unconstrained, regardless of the appearance to the contrary.

In fact, attempts have been made to delimit the syntactic positions in which LIKE may surface (Underhill 1988; Andersen 2001). However, these analyses have largely relied on the places where LIKE does occur, while those where it does not have received little attention. Moreover, with the exception of Dailey-O’Cain (2000), no investigation of

discourse LIKE has considered the speech community in its entirety; rather, the focus has been on a specific sub-section of the population: younger, and generally adolescent, speakers (e.g., Miller & Weinert 1995; Andersen 1997 et seq.; Siegel 2002; Hasund 2003; Sharifian & Malcolm 2003). Until these groups are seen in relation to older segments of the population, patterns of use among those who are purported to use LIKE the most cannot be contextualized. Presumably, contemporary youth could not have invented this form *ex nihilo*. It is more plausible that LIKE has developed gradually, arriving at its current usage among younger speakers through regular processes of language change.

1.2 RESEARCH QUESTIONS

The aim of this research is to present a variationist analysis of discourse LIKE with a view to tracking its development in contemporary (Canadian) English. The theory to be defended in this dissertation is that current use is both systematic and rule-governed, implicating not only pragmatic constraints on its use, but syntactic ones as well. The main questions driving this investigation are:

- 1) What are the contexts in which LIKE occurs and how does this particle interact with the syntax? In other words, is there evidence for systematicity in use or can LIKE in fact occur “anywhere”?
 - If there are structural considerations that govern the distribution of LIKE, what are they?
 - How can we account for them?
- 2) Are these considerations constant across apparent-time or is there evidence for ongoing change?
- 3) Is LIKE simply age-graded, associated with younger speakers, or is there evidence that its use is rising in frequency and/or that it is spreading into new contexts?
- 4) Are the ‘innovative’ uses of LIKE in (1) related to the ‘traditional’ ones in (4)?
- 5) What can these results tell us about discourse particles more generally?

In short, this study aims to describe the synchronic distribution and development of LIKE and from there, to establish how it is constrained in the grammar. The dissertation is organized as follows. I first present evidence that in numerically quantified contexts, as in (8), LIKE is not a pragmatic device, but functions adverbially, signaling approximation (Chapter 2). I then set out the current view of the role discourse markers play in a theory of grammaticalization and examine whether LIKE fits into this model of language change (Chapter 3). In the following four chapters (Chapters 4-7), I apply variationist methodology to a corpus of contemporary data in order to provide an accountable analysis of the path(s) along which LIKE may be diffusing. I track the trajectory of LIKE across apparent-time, as this may shed light on the nature of this discourse feature (i.e., substitution, grammaticalization, generational change, etc.). In these chapters, I focus on individual, structurally circumscribed contexts (see §1.4). Finally, I discuss the relevance of these findings for the development of LIKE and its implications for the analysis of discourse particles (Chapter 8). This research will therefore provide critical insight into the development and function of discourse LIKE in English, as well as additional evidence for understanding the mechanisms that may underlie grammatical change more generally.

1.3 THE (SOCIO) LINGUISTIC ISSUES

Discourse-related forms constitute, in terms of frequency, the fourth largest word-class in spoken language; they are outnumbered by only verbs, pronouns, and nouns (Altenberg 1990:185). According to Aijmer (2002:2-3), they are not “meaningless decorations,” but are

used “with great precision as signposts in the interaction.” This raises the question: What exactly are discourse markers and what are their functions?

The answer to the former is by no means straightforward, since the definition of a marker seems to depend on which of several pragmatic functions is considered primary (Brinton 1996:30). Consequently, I will begin the discussion by outlining those that are seen as fundamental to discourse markers in general, as summarized in Brinton (1996:36-37). These functions can be divided into two broad categories: TEXTUAL and INTERPERSONAL. Within the first category, three sub-classes can be distinguished: 1) turn-taking, 2) discourse structure, and 3) information structure.

In negotiating turn-taking, markers help a speaker initiate discourse (e.g., Strang 1963; Quirk et al. 1985; Schiffrin 1987) or acquire the floor during discourse (e.g., Faerch & Kasper 1982; Levinson 1983; Schiffrin 1987). Markers can also serve as fillers, prolonging a turn or helping the speaker hold the floor (e.g., Brown 1977; Faerch & Kasper 1982; Östman 1982; Schiffrin 1987).

With regard to discourse structure, markers can signal a boundary such as the introduction of a new topic or a shift in topic (e.g., Brown 1977; Svartvik 1979; Brown & Yule 1983; Erman 1986; Schiffrin 1987). Further, they can indicate the relationship between two utterances within the discourse (e.g., Quirk et al. 1985; Schiffrin 1987; Fraser 1988, 1990). In this last function, Schiffrin (1987) refers to them as DISCOURSE DEICTICS, a subset of discourse markers more generally. Fraser (1988, 1990) and Traugott (1997 [1995], 1997), assuming a much more restrictive view, consider only expressions serving as deictics to be true discourse markers. In other functions, these expressions fall into the category of

discourse particles. The relevance of this distinction will become clear in Chapter 3. For now, no distinction between MARKER and PARTICLE is made; rather, the terms are used interchangeably.

Concerning information structure, discourse markers can denote new as well as old information (Quirk et al. 1985; Erman 1987; Schiffrin 1987).

In the interpersonal realm, discourse markers aid cooperative aspects of communication such as checking or expressing understanding, as well as establishing a sense of sharing or intimacy between interlocutors (e.g., Quirk 1972; Östman 1982; Schourup 1983, 1999; Quirk et al. 1985; Schiffrin 1987). Indeed, since interactions that are free of markers are perceived as unnatural, awkward, dogmatic, or even unfriendly (Brinton 1996:35; see Quirk 1972; Svartvik 1979; Schiffrin 1987), perhaps the foremost utility of these forms is that they help both speaker and hearer maintain cooperation through a verbal discourse (Even-Zohar 1982:180).

It is the textual and interpersonal functions of discourse markers that together constitute pragmatic meaning (Brinton, forthcoming). Returning then to the question of what they are, discourse markers have the following characteristics. They are high frequency features of spoken discourse (Quirk et al. 1985; Erman 1987; Watts 1989), occurring sometimes multiple times within a single sentence. They are stylistically stigmatized (Quirk 1972; Evan-Zohar 1982; Östman 1982; Erman 1987; Schiffrin 1987), characterized as “one of the most perceptually salient features of oral style” (Watts 1989:208). They are largely void of propositional content and are thus difficult to categorize lexically (Östman 1982; Schiffrin 1986), often classified as interjections (Brinton forthcoming). This semantic emptiness also

renders them awkward to gloss or translate (Brinton 1996; see also Hasund 2003 on English LIKE and Norwegian LIKSOM). They appear to be optional elements of discourse (Quirk 1972; Svartvik 1979; Brown & Yule 1983; Quirk et al. 1985; Schourup 1999, Schiffrin 1987; Fraser 1988). Markers can also be defined along phonological and prosodic parameters, occurring in a separate tone group where they are usually ‘short’, unstressed, and/or reduced (Östman 1982; Schiffrin 1987).

Frequently cited forms in Modern English include so, then, well, you know, actually, anyway, and I mean. Added to this list is LIKE, which shares the formal characteristics of other discourse markers. In terms of the pragmatic functions that have been associated with its use, some fall into the interpersonal realm, where LIKE may serve to indicate to the listener the speaker’s epistemic stance regarding the form of the utterance (e.g., Schourup 1983; Andersen 1997 et seq.), while others are clearly textual, such as marking focus (Underhill 1988; Miller & Weinert 1995) or serving as a pausal interjection (e.g., Schourup 1983).

Miller and Weinert (1995:372-374) dispute LIKE as a pausal interjection, arguing that in their data, it rarely occurs with the hesitations and syntactic reformulations that are typically associated with processing difficulties. Further, they point out that although only half of their tokens of LIKE (N = 40) are fully integrated into the discourse context, when LIKE does co-occur with a pause, it is two and a half times more likely to be fully realized phonetically. They maintain that this result is explicable on the basis that LIKE plays too important a role as a discourse organizer to be reduced.

Siegel (2002:46-47), however, argues that LIKE can in fact be accounted for as the product of lexical indecision, because its frequency is significantly correlated to processing time (see also Ross & Cooper 1979:345). She shows that as the time to plan an utterance before speaking increases, the use of LIKE decreases. Viewed from this perspective, the function of LIKE is primarily one of maintaining the floor, since it is used to fill the pause created by searching for words during a turn.

In assessing the various arguments in the literature regarding the pragmatic motivations for discourse LIKE, it is important to bear in mind that discourse markers are not limited to performing a single function. Indeed, it is typical for them to fill a variety of roles and even to fill two or more simultaneously (Schiffrin 1987:60-61). Thus we find numerous attestations that LIKE is multifunctional, performing a number of pragmatic duties (Andersen 2001; Buchstaller 2001). As discussed above, these functions may be distributed between the textual and interpersonal planes of the pragmatic component, yet discourse markers are not restricted to a solely pragmatic role. They may also operate within the morphophonemic, syntactic, and semantic modules (Schiffrin 1987; van Dijk 1982; Redeker 1990). Further, discourse markers have “highly constrained syntactic as well as intonational properties” (Traugott 1997 [1995]:5). As such, a growing number of scholars argue that pragmatic particles unambiguously belong to the grammar (e.g., Fraser 1988; Traugott 1997 [1995], 2003; Brinton 2005, forthcoming).

I will resume this last line of discussion in Chapter 3; first, in order to contextualize my research questions I outline previous work on LIKE that has taken social and non-pragmatic language-internal factors into account.

1.3.1 Social factors

Research on discourse LIKE has focused primarily on its pragmatic properties (e.g., Schourup 1983; Underhill 1988; Miller & Weinert 1995; Andersen 1997 et seq.; Buchstaller 2001; Sharifian & Malcolm 2003; etc.). Comparatively, its sociolinguistic distribution has received little attention. Even when the effects of external factors such as speaker sex and age are considered, discourse LIKE and quotative be like are often collapsed, treated as a single feature (e.g., Andersen 1997 et seq.; Hasund 2003; Levey 2004).⁷ While this does not invalidate the findings, it potentially obscures any differences that social factors may have on the diffusion of these two forms. The important point here, however, is that there is not always agreement across investigations concerning the role of social conditioning on the use of LIKE.

Consider speaker sex. While Andersen (2001), Siegel (2002), Hasund (2003) and Tagliamonte (2005) report that females produce more tokens of LIKE, Wolgemuth (2003:66,72-73) found that it is significantly more frequent in the speech of males. The results of Dailey-O’Cain (2000:66) also support a male association, though the sex effect fails to reach the level of significance.⁸

⁷ Since I address only sex and age in this dissertation, I limit the current discussion to these factors. The speakers selected for the analyses presented herein are largely middle class Caucasians, but for discussion of the effects of social class on the use of LIKE, see Andersen (2001:289-290), and for ethnic identity see Andersen (2001:290-291) and Ferrara and Bell (1995:277-278).

⁸ Interestingly, similar mixed results were reported in early analyses of quotative be like (e.g., Blyth et al. 1990; Romaine & Lange 1991; Ferrara & Bell 1995). More recent research, however, clearly demonstrates that as the frequency of this form rises, a strong and significant female association develops (Tagliamonte & Hudson 1999; Tagliamonte & D’Arcy 2004, under review).

The data for both Andersen's and Hasund's analyses come from the Bergen COLT corpus, collected in 1993, where even among the females, it occurs 3.24 times per 1000 words (Andersen 2001:288; Hasund [2003:218] reports a slightly higher mean, 3.5 times per 1000 words). The data used in Dailey-O'Cain's and Wolgemuth's research were individually collected in 1995, and as in COLT, LIKE is relatively infrequent. Methodological differences prohibit a direct comparison, but in over 3200 clauses, Wolgemuth found only 202 tokens of LIKE (2003:43), while Dailey-O'Cain reports that even at its most frequent, LIKE occurs with just 14 percent of focus elements (2000:66). Thus, it is possible that in these instances, the text frequency of LIKE was simply too low to generate reliable results according to sex.

A further possibility is that quantitatively, sex simply does not condition LIKE in any straightforward manner (see also Eckert 2003:395). Previous research of other discourse markers is equally mixed. For example, Östman (1981) finds that you know is more frequent in the speech of females than in that of males, Erman (1992) finds the exact opposite, and Holmes (1986) finds that it is equally distributed between the sexes. Questions tags have yielded no further insights. Contrary to Lakoff's (1973a) claim that these types of features are typical of female speech, Dubois and Crouch (1975) and Cameron, McAlinden and O'Leary (1989) report that males produce more tags than females. Holmes (1984, 1995), however, supports Lakoff's finding, reporting a higher frequency among her female participants. Dubois (1992:198-199) argues that differences in overall rates of use between the sexes derive from stylistic variation; she concludes that proportional variation is "at most weakly correlated with sociodemographic factors."

The more important effect of sex seems to relate instead to the way in which pragmatic forms are used. In other words, the role of sex is qualitative, not quantitative. For example, using data from Montreal French, Dubois (1992:197) reports that extension particles such as tout le kit ('everything'), which function as universal quantifiers (i.e., extending the scope of the sentence "to include the entire set containing the operand" (1992:194)), are favoured by women. In contrast, forms like affaires de même ('things like that'), which function generically and comparatively, are favoured by men. A more straightforward example is provided by Erman (1992:229-231), who demonstrates that females prefer to use you know and you see to mark discourse structure (e.g., foreground vs. background), while males typically use these forms for decoding and turn-regulating functions. In sum, it appears that function is more critical than frequency in distinguishing male-female uses of discourse features.

The results for age are less controversial. All studies agree that LIKE is more frequent in the interactions of younger speakers, with use peaking among adolescents (Andersen 1997 et seq.; Dailey-O'Cain 2000; Hasund 2003; Tagliamonte 2005). These analyses also demonstrate that LIKE is much less frequent among 10-12 year olds than among their teenage peers. Working from similar findings, Miller and Weinert (1995:380) postulate that discourse LIKE is acquired after the age of 10. Assuming that LIKE functions as a focus element, they suggest that this developmental pattern may be linked to the acquisition of discourse management skills, which develop relatively late.

What is unknown is whether all pragmatic functions of LIKE are acquired after this time, or whether some contexts of use are acquired sooner than others. Further, research shows

that children as young as three or four years of age are able to acquire the linguistic constraints on stable variation (Roberts 1994, 1997; Roberts & Labov 1995), and recent findings involving 10-12 year olds indicate that even in cases of ongoing change, it is the social constraints that are developmental, not the linguistic or pragmatic ones, which are already established by this time (Tagliamonte & D'Arcy 2004).

Another issue concerns the role age may play in the diffusion of discourse LIKE within the community. It remains unclear whether LIKE is an age-graded feature or whether its use persists after adolescence. Indeed, there is no evidence bearing on the issue of whether or not the frequency of LIKE will remain stable among those who currently use it. Consequently, the upper edge of discourse uses is unknown, raising the question of whether LIKE is maintained at adolescent levels, decreases, or disappears. In other words, is this an age-related usage or is it *bona fide* generational change (see also Cheshire 2002:27)?

1.3.2 Linguistic factors

A further result of the emphasis on pragmatics is that the effects of other linguistic factors that may bear on the use of LIKE remain virtually unknown. Siegel (2002) attempts to address the semantic aspects of LIKE, arguing that it affects the weak/strong distinction in sluicing sentences and existential 'there' constructions, and that it may affect truth conditions more generally. Syntactically, LIKE is characterized by its "syntactic detachability and positional mobility" (Romaine & Lange 1991:261), which is roughly equivalent to Siegel's (2002:64) claim that it can "occur grammatically anywhere in a sentence." These descriptions derive from the wide variety of surface strings with which

LIKE can co-occur, including noun phrases (9a), verb phrases (9b), adjective phrases (9c), prepositional phrases (9d), adverb phrases (9e), and full clauses or sentences (9f) (e.g., Underhill 1988; Andersen 1997 et seq.).

- (9) a. I couldn't get **LIKE** a peaceful sleep. (2/r/f/11)
 b. They don't want me to **LIKE** turn out to be some bum. (3/J/m/18)
 c. It was **LIKE** fake. (3/O/m/11)
 d. It's not **LIKE** for adults (2/n/f/19)
 e. People are doing it **LIKE** unbiasedly. (I/i/m/22)
 f. **LIKE** she was a maid, she worked at a bakery. (I/&/f/21)

In an attempt to circumscribe the syntactic distribution of LIKE, Underhill (1988:243) claims that it is “closely rule-governed. [...] it always or nearly always introduces a constituent.” But, ‘constituent’ is a generic category (e.g., Adger 2003:64); in terms of structure, this description is vague.

A decade before the publication of Underhill (1988), Ross and Cooper (1979) — in an extremely detailed analysis — outlined the syntactic restrictions on LIKE, comparing it to even, only, also, just and too. The theory has changed significantly since then, but the phenomena they were trying to explain remain largely unchanged. What is important for the present purposes is that Ross and Cooper propose a constraint called “LIKE as a Left-Bracket Condition” (LLBC), which states that LIKE has to open a constituent that dominates the focused element (1979:349). In short, it is associated with the left periphery of the element over which it scopes.⁹

⁹ Ross and Cooper later amend the LLBC to incorporate the notion of C-command, arguing that LIKE has to be adjacent to the constituent that C-commands the focused element (1979:358). As this line of argumentation not only relies heavily on LIKE performing strictly as a focus particle but also draws crucially on the notion of LIKE as a pausal interjection, I set aside this part of their analysis.

Andersen (2001:275), on the other hand, avoids the notion of constituency and distinguishes instead between clause-internal (9a-e) and clause-external (9f) contexts. In the former LIKE is “syntactically bound to and dependent on a linguistic structure,” while in the latter it is unbound. I will focus on two of his findings here. First, Andersen reports that approximately one-third of all instances of LIKE are parenthetical (i.e., unbound), while the remaining 66 percent occur clause-internally. He posits this to mean that LIKE displays a greater degree of syntactic boundedness than is generally the case for discourse particles (2001:274). Second, Andersen finds that whether LIKE is situated before a phrase or within one is dependent on the type of phrase being modified (2001:276). For example, LIKE occurs more often before a determiner than before the head noun, but in the case of verb and prepositional phrases, it is more typically found within the phrase. Based in large on the evidence from verb phrases, he proposes the PRINCIPLE OF LEXICAL ATTRACTION, which states that LIKE tends to occur before lexical rather than functional material (2001:284).

This type of approach is revealing in terms of the types of patterns we can expect to find with regard to various surface strings. What it cannot account for, however, is the distribution of LIKE on syntactic grounds. Specifically, it uncovers little about what types of structural constraints may be operative.

Consider the examples in (10), from Andersen (2001:277):

- (10) a. They should do them [in **LIKE** fishing shops].
 b. So they [can **LIKE** score] the equalise [sic] and win it.

When LIKE occurs within a prepositional phrase, as in (10a), it follows the preposition, taking scope over the following noun phrase. This type of example is therefore not

informative with regard to prepositional phrases, since LIKE is only part of this phrase through the coincidence of being adjoined to a noun phrase which forms its complement. Further, while we may talk of being ‘within’ a noun phrase in the sense of occurring between a head and its specifier, the same cannot be said with respect to verb phrases, since in phrases such as the bracketed sequence in (10b), the modal ‘can’ is not the specifier of the verb ‘score’, but rather is the head of its own functional category, tense. In short, ‘entering’ a noun phrase is not equivalent to ‘entering’ a verb phrase. Thus, a more informative line of inquiry may be to consider the distribution of LIKE with a view to structural considerations, and to ponder what, if anything, examples like (10a) have in common with those like (10b).

Despite the wide range of strings in which LIKE occurs and the issues inherent with attempting to account for this ‘mobility’, three structures repeatedly emerge as the most frequent slots for discourse uses: the verb phrase, the noun phrase, and the sentence/clause initial context (e.g., Schourup 1983; Underhill 1988; Romaine & Lange 1991; Andersen 1997; Levey 2004; Wolgemuth 2003). These analyses rely on occurrences of LIKE itself. Thus, statements of frequencies — of which there are many — give no indication as to the propensity of LIKE in a given context overall. Indeed, Wolgemuth (2003:43,66) finds that although the greatest number of LIKE tokens precede a noun phrase (N = 68), when the overall distribution is calculated for this context, LIKE accounts for a mere 3 percent of the data (N = 1773).

It is this last point that highlights the over-riding importance of the PRINCIPLE OF ACCOUNTABILITY (Labov 1972:72), a methodological underpinning of quantitative sociolinguistics. Accountability refers to the practice of considering not only all the

instances of the feature under investigation, but also those places where it could have occurred but did not. Notably, the contexts where LIKE is cited as being most frequent (i.e., VP, NP, sentence/clause initial) are also the most frequent syntactic slots overall (see, for example, Altenberg 1990:185), making it difficult to interpret previous results. Consequently, little is known about how discourse LIKE is actually distributed linguistically.

This situation is compounded by treating the contexts in which LIKE may occur as monolithic entities, undifferentiated by either their internal constituency or their role in the syntax. For example, taking the verb phrase as a baseline, where in the VP can LIKE occur? Is it distributed evenly across different types of VPs, or do certain contexts favour its use? Is verb type relevant? In other words, does syntactic structure play a role in the types of constructions in which LIKE appears?

As I will show, the custom adopted to this point in the analysis of discourse LIKE misses important generalizations about its distribution. Moreover, I will show that while the primary functions of LIKE may be pragmatic, this marker has internal linguistic constraints on its use. This last result clearly implicates the grammar in the account of discourse LIKE.

1.4 METHODOLOGY

1.4.1 Data

I propose to investigate the issues raised in the two previous sections by performing an accountable quantitative analysis of discourse LIKE. The data for this research come from 97

speakers between the ages of 10 and 87, stratified by age and sex (see Table 1.1), who were drawn from a corpus of contemporary Toronto English.¹⁰

TABLE 1.1 The sample

Age	Male	Female	Total
10-12	5	5	10
15-16	4	4	8
17-19	5	5	10
20-24	5	5	10
25-29	5	5	10
30-39	5	5	10
40-49	4	4	8
50-59	4	4	8
60-69	4	4	8
70-79	3	4	7
80+	4	4	8
TOTAL	48	49	97

These materials, which comprise standard Labovian sociolinguistic interviews (Labov 1966, 1972), were collected between the autumn of 2002 and the summer of 2004 using a combination of network and random sampling strategies. The prime criterion was that speakers be born and raised in Toronto. This, in combination with the broad age range encompassed by the corpus, allows for the examination of LIKE using the apparent-time construct (Labov 1963, 1966).

¹⁰ These data were made available through the Research Opportunities Program at the University of Toronto (ROP), SSHRC grant #410-03-0005, 'Linguistic Changes in Canada Entering the 21st Century' (TO), and the director of these projects, Sali Tagliamonte. The speakers used here represent a sub-sample of the full 1.5 million-word Toronto English Archive, housed in the Sociolinguistics Laboratory of the Linguistics Department at the University of Toronto. For further details regarding the sampling methodology and corpus construction, see Tagliamonte and D'Arcy (2004, under review).

1.4.2 Method

Within the traditional Labovian sociolinguistic paradigm, the application of quantitative methods to the examination of discourse-level features is relatively new (e.g., Tagliamonte 2005). Even more recent is the analysis of discourse LIKE within the variationist framework, which has either had to contend with a paucity of tokens (Wolgemuth 2003) or rely on distributions, thus setting aside the issue of language internal factors and their possible effects on use (Dailey-O’Cain 2000; Tagliamonte 2005). Perhaps the greatest challenge to the rigorous quantitative approach that is at the heart of the variationist tradition is what Stubbe and Holmes (1995:71) describe as the infeasibility of “delineating a universe of discourse for pragmatic devices which includes opportunities for potential occurrence as well as actual occurrences.” In other words, the envelope of variation for these types of features has been considered to be subjective, residing strictly in the motivations of the speaker and therefore not available for reconstruction by the analyst (e.g., Dines 1980; Vincent 1986; Dubois 1992; Dubois & Sankoff 2001; also Poplack & Tagliamonte 2000, 2001). As a result, investigations of discourse LIKE have relied on the raw number of occurrences, which are divided by the total number of words (providing a ratio per 1000 words) to provide a normalizing measure for gauging frequency (Andersen 1997 et seq.; Hasund 2003; Levey 2004). This methodological practice allows for accountable and valid comparisons of data cells both within as well as across corpora (e.g., Stubbe & Holmes 1995:71; Macauley 2002:299).

Nonetheless, there are at least two complications to this approach. First, in basing frequency on the number of words, the central issue — that the ‘universe of discourse’

cannot be established — remains unaddressed. Instead, the focus is simply shifted from the local context of individual utterances to the universal context of the discourse as a whole. Second, this method assumes that the entire discourse, and thus all of its component parts, presents equal opportunities for the use of the discourse feature under examination. It is impossible to exclude contexts where the feature could not occur for pragmatic, syntactic, or semantic reasons. This type of methodology therefore has the potential to obfuscate variation arising from the operation of constraints on use.

Consequently, I adopt the principle of accountability (Labov 1972) and circumscribe the variable context for discourse LIKE according to structural criteria. It is important to note that I make no attempt to ascertain the pragmatic motivations of the speaker, but rely instead on structural factors in determining the contexts of use. This method has never been applied to a discourse feature before now. However, discourse LIKE is unique in that it is used in a number of syntactically delimited positions. Thus, despite the inaccessibility of the ‘pragmatic universe’ that gives rise to the appropriate interactional context for its use, the linguistic universe in which LIKE occurs is fully accessible and objectively circumscribable. I will elaborate on the methodology in chapters 4-7; here I present a brief introduction to the issues.

As touched upon in section 1.3.2, although discourse LIKE can occur in a number of syntactic positions, the three most frequent slots have been reported to be the clause-initial context, before a noun or noun phrase, and before a verb (e.g., Schourup 1983; Underhill 1988; Andersen 1997; etc.). Since these are also the most frequent slots in general, and because these previous results are not based on accountable methods, it is uncertain whether

these alleged ‘hot’ spots are in fact preferential slots for LIKE. In order to test these claims, I focus on individual heads and their projections, such as determiner phrases, verb phrases, and predicate adjective phrases, as well as clause-level structures. In so doing, I delimit the variable context following well-defined syntactic considerations. Moreover, Roberts and Roussou (2003) argue that grammatical change progresses along a trajectory entailing movement from one functional head to another. This methodology thus enables me to address this issue, a topic I take up in Chapter 8.

Circumscribing the variable context in this way necessitates omitting those structures in which the position of LIKE cannot be syntactically disambiguated. The focus must be contexts which allow for the unequivocal determination of the slot occupied by LIKE, otherwise the validity of the generalizations is compromised. A common practice in diachronic studies of syntactic variation and change (e.g., Pintzuk & Kroch 1989; Pintzuk 1993 et seq.; Santorini 1993; Taylor 1994) is to disambiguate the position of the variable in question vis-à-vis some other element in the phrase, such as adverbs, particles, and pronouns, because these elements serve as a baseline for situating the locus and the nature of the variation. Pintzuk (1993, 1996, 1999:56), for example, argues for optional verb movement in Old English subordinate clauses on the basis of the distribution of the verb relative to particles: in verb-medial subordinates, the particle frequently appears after the finite verb (ahoff upp ‘lifted up’), explicable on the basis that these are the only clauses in which the verb could have moved leftward, while in verb-final clauses, the particle remains in situ, to the left of the finite verb (up ahof ‘lifted up’). Throughout the current

investigation, if structural diagnostics such as this cannot determine the syntactic position of LIKE, the context is removed from consideration.

In extracting the data from the interview materials, both contexts with and without LIKE were included in the analyses. Consider for example the passage in (11), which illustrates the clause-initial context. Wherever LIKE occurs in parentheses and unbolded, it was not uttered in actuality. Only the bolded form represents an actual occurrence.

- (11) (LIKE) I got quarantined,
 but (LIKE) there was a ten-day quarantine starting from the certain date,
 and (LIKE) there was already a week happened.
LIKE, the week had already gone by,
 so (LIKE) I only got quarantined for probably-
 (LIKE) it might have been four days I got quarantined. (N/Q/f/72)

Each of these clauses was extracted and coded for whether or not LIKE occurred, as well as for a number of other factors including language-external ones such as speaker age and sex, and language-internal ones like clause type and position in the discourse (e.g., utterance initial, medial, etc.).

To avoid having the results of an individual skew those of the group, the same amount of data was extracted for every speaker (e.g., 75 verb phrases, 60 clauses, etc.) whenever possible.

The focus of this analysis being the discourse function, non-discourse functions of LIKE as in (2) were ignored. Similarly, quotative be like forms no part of this investigation on the basis that it represents a separate development with its own specialized function as a quotative complementizer (§1.1). Only instances of LIKE as used in (1) and (5)-(7) were included (though the ‘approximative’ function of LIKE as in (8) is the topic of Chapter 2).

Whether LIKE is performing discursively or not in a given utterance is not always clear. For example, both uses of LIKE in (12) are ambiguous. The first is potentially prepositional, with the noun phrase [back-stabbing assholes] its complement (i.e., they were similar to back-stabbing assholes). The second LIKE may be adverbial, this time the argument of the verb phrase [were backstabbing] (i.e., they were back-stabbing in the manner of assholes). Because of these possibilities, and because the intended functions cannot be disambiguated, both tokens were excluded.

(12) He said that they were **LIKE** back-stabbing **LIKE** assholes. (3/S/m/18)

There are also instances when LIKE is not used, but were it to surface in a particular context, its function would be ambiguous, as in (13).

(13) No one is \emptyset a stranger there. (I/&/f/21)

Here, a likely — though not certain — interpretation of ‘No one is like a stranger here’ is that no one is similar to a stranger here. As a result, tokens such as this were also excluded.

Finally, there are a number of other exclusions that only became apparent through the application of accountable methods. Specifically, in being able to code for a number of contextual factors whether LIKE occurred or not, it was discovered that certain contexts are invariant. For example, as shown in (14), LIKE never precedes copular BE when the verb occurs alone without auxiliaries or modals and is finite (N = 384). Further, exemplified in (15), LIKE is never used clause-initially following a direct question (N = 155).

- (14) a. They \emptyset are just on my bed. (2/g/m/11)
 *They **LIKE** are just on my bed.
- b. I've caught **LIKE** trout that \emptyset are small. (3/F/f/17)
 *I've caught **LIKE** trout that **LIKE** are small.
- (15) a. Int: Are you guys close?
 FD: \emptyset When we see each other. (N/K/m/26)
 ***LIKE** when we see each other.

In short, this methodology reveals that a number of syntactic constructions do not constitute part of the variable context for **LIKE**. This discovery refutes Siegel's (2002:64) claim that **LIKE** is unconstrained syntactically, because there are certain structures in these data in which it never occurs. These categorical contexts were ultimately excluded, as were a number of constructions in which **LIKE** does occur, but its frequency falls below 5 percent. These were removed because extreme distributions of this type not only skew overall results, but more importantly, they distort the analytical results achieved by multivariate analysis (e.g., Guy 1988:132), performed here using Goldvarb (Rand & Sankoff 1990). In the relevant chapters I will explore the reasons behind these categorical (or nearly categorical) distributions, where it will be seen that the distinction between a context that is variable and one that is categorical is not stipulative, but often falls out from syntactic considerations (see also Smith 2000). All remaining contexts — which potentially allow **LIKE** — form the basis for the analyses that constitute the body of this dissertation.

1.5 SUMMARY

This dissertation addresses the rise of discourse **LIKE** in a contemporary variety of English, using apparent-time data from a single, cohesive speech community (see Chambers &

Trudgill 1998:151): Toronto, Canada. The approach I adopt is to examine the way(s) in which LIKE constructions have arisen by determining how they are used and consequently, how they shape the language (Hopper & Traugott 2003:1). In so doing, I systematically track not just those positions where LIKE does surface, but also those where it does not. As will be shown, this method provides a clear view of the development of this feature, enabling me to address the issue of whether LIKE represents grammaticalization, lexical replacement, neither, or possibly even both. Where previous research has argued for entirely pragmatic functions for discourse LIKE, this analysis will demonstrate that this feature systematically interacts with the syntax. It is the regularity of these interactions across the community that validates the methodological approach adopted here, which is novel to the analysis of a discourse-pragmatic feature. I will show how LIKE has gradually generalized not only from one phrase-level structure to another, but to new constructions within each projection. From the perspective afforded by accountable quantitative methodology, I will discuss the relevance of these findings for tracking the evolution of discourse-pragmatic features as well as their implications for understanding the more general mechanisms guiding grammatical change.

Chapter 2 NUMERICAL EXPRESSIONS

A characteristic trait of discourse features, both markers and particles, is their lack of lexical meaning (Östman 1982; Schiffrin 1986). This is one of the reasons that they are notoriously difficult to categorize, gloss or translate (Brinton 1996; see also Hasund 2003). This aspect of their semantic profile arguably derives from the grammaticalization process, which in the initial stages is characterized by pragmatic strengthening, a development that is generally followed by semantic bleaching or weakening (e.g., Hopper & Traugott 2003).

In the vast majority of its uses, LIKE meets the semantic emptiness criterion, imbued instead with pragmatic meaning(s). However, in the highly circumscribed context of quantifiers, and numerical expressions in particular, as in (16), LIKE is argued to carry propositional content, denoting an approximative meaning (Schourup 1983; Underhill 1988; Meehan 1991; Jucker & Smith 1998; Biber et al. 1999).

- (16) a. It could have taken you all day to go **LIKE** thirty miles. (N/©/f/76)
b. You-know, it was **LIKE** a hundred and four [degrees]. (N/w/m/84)
c. They were **LIKE** eighteen years old; they were kids. (N/s/f/52)

Historically, the use of LIKE in quantified contexts can be traced to the early nineteenth century. However, as the examples in (17) demonstrate, its content in this frame was not originally approximative. (17a) is drawn from Scottish dialect data, where Grant and Dixon

(1921:142) describe its function as an adverb of probability. This same function is present in (17b), from the OED, where it can be glossed as ‘more like(ly)’.

- (17) a. The three mile diminished into **LIKE** a mile and a bittock. (Scott 1815)
 b. No more at midway to heaven, but **LIKER**, midway to the pit. (Meredith 1898)

In fact, even though **LIKE** has embodied approximative meanings for well over four centuries in some of its uses as a preposition and conjunction (Meehan 1991; OED), the OED contains no examples in which **LIKE** can be construed as carrying this meaning before quantified phrases. Nonetheless, contemporary data indicate that **LIKE** is used in this way not only among younger speakers in London (Andersen 2001:277; see also Levey 2004), where it reportedly occurs quite frequently, but instances can also be found among older speakers of regional varieties across the United Kingdom, as seen in (18).

- (18) a. I was only **LIKE** forty-one or forty-two or something aye. (AYR/J/f/86)
 b. It used to be frae Cumnock to Ayr was **LIKE** eighteen miles. (AYR/c/f/78)
 c. Whenever I was wee, whenever I was **LIKE** ten or twelve year old. (PVG/d/f/62)
 d. That’ll have been **LIKE** thirty-five year I would say. (MPT/z/m/78)
 (Tagliamonte to appear)

The germane question is, is the use of **LIKE** in measurable, numeric contexts part of its function as a pragmatic feature or is it an entirely different feature? That is, is the **LIKE** in (16) the same **LIKE** that we find in (19)?

- (19) a. There was all **LIKE** dirt roads and all that. (N/©/f/76)
 b. It lasted **LIKE** through the night into the next morning. (N/w/m/84)
 c. I’ve discovered new reading material which is **LIKE** so interesting. (N/s/f/52)

Andersen (2001:260) argues yes, LIKE in (16) is a “genuine pragmatic marker” because it signals to the listener that “the utterance contains a loose interpretation of the speaker’s thought.” In other words, whereas adverbs such as ‘roughly’ and ‘approximately’ operate at the propositional level, LIKE operates metalinguistically, i.e., pragmatically (ibidem). By implication, LIKE is not functionally equivalent to the adverbials; they are not lexical variants.

Despite the arguments in Andersen (2001), LIKE can affect truth conditions in these contexts (Jucker & Smith 1998; Siegel 2002), a point Andersen acknowledges, even pointing out that its omission can affect the propositional meaning of the utterance (2001:260). In this sense, LIKE is comparable to other truth-conditional approximative adverbs such as roughly and about. Consider, for example, the data (20) and (21).

- (20) a. This is **ABOUT** four years later,
so I guess I’ve been back for **ABOUT** eight years now. (N/fi/m/37)
- b. This is four years later,
so I guess I’ve been back for eight years now.
- (21) a. If there was a lot of traffic, **LIKE** twenty-five minutes maximum.
And usually it takes **LIKE** fifteen minutes to get there. (3/U/f/12)
- b. If there was a lot of traffic, twenty-five minutes maximum.
And usually it takes fifteen minutes to get there.

Neither the amended utterance in (20b) nor the one in (21b) is propositionally equivalent to (20a) and (21a). The use of about in (20) and LIKE in (21) signals that the time spans are approximations and that their literal interpretation is not intended. In contrast, the statements in which these forms are omitted are to be interpreted verbatim.

The ability of LIKE to affect truth conditions in this way is problematic for the pragmatic argument, since one of the primary definitions of discourse features is their inability to interfere with semantics (Hölker 1991; Jucker 1993; Fraser 1996; see also Siegel 2002). As a result, LIKE is sometimes argued to be adverbial before quantified phrases (Schourup 1983; Underhill 1988; Biber et al. 1999). This possibility is of primary importance when circumscribing the variable context of discourse LIKE in nominal and adjectival constructions, syntactic projections that form the basis of two of the following chapters. If LIKE does in fact function as an adverb before quantifiers, then this LIKE is functionally distinct from the discourse uses of this form. Consequently, it must be excluded from the discursive analyses of both noun and adjective phrases.

The question I consider in this chapter is whether the use of LIKE in (16) is akin to that in (19), which arguably derives from grammaticalization processes, or whether it is better accounted for in a model of lexical change. The discussion is organized as follows. In section 2.1, I detail the methodological considerations. In section 2.2, I present the results and interpret the findings in light of previous research on lexical replacement. Finally, in section 2.3, I address further issues raised the analysis, such as the role of synonymy and the implications for future research.

2.1 THE DATA

There are two principal types of quantified phrases: noun phrases (NP), as in (22), and adjective phrases (AP), as in (23). These latter constructions are primarily expressions of age (i.e., X years old(er)), though they also include extent adjectives such as tall, high, and long.

- (22) a. It could have taken you all day to go **LIKE** thirty miles. (N/©/f/76)
 b. My salary was **LIKE** twenty-seven-thousand dollars. (I/1/m/51)
 c. The guy weighed **LIKE** a-hundred pounds. (I*/f/30)¹
- (23) a. He got to high-school when he was **LIKE** twelve. (N/Q/f/72)
 b. They were **LIKE** eighteen years old; they were kids. (N/s/f/52)
 c. He's **LIKE** five feet tall. (I/¶/m/22)
 d. They're **LIKE** sixteen feet long. (I/8/m/32)

Although there are a number of quantifiers available in English (e.g., all, much/many, some, any, and numerals; Biber et al. 1999:275-277), the most common quantifier in discourse is the cardinal number (1999:279). Consequently, the current analysis focuses on numerically quantified NPs and APs. Every instance of a cardinal number used to modify a NP or AP was extracted from the interview materials.

The hypothesis being tested is that **LIKE** functions adverbially in this context, alternating with other approximative adverbs (e.g., about, around, approximately, roughly, etc). As such, only those structures capable of being approximated were included in the analysis; tokens unable to support modification do not form part of the variable context. These

¹ For discussion of the indefinite article as functionally equivalent to unstressed cardinal ONE, see Jespersen (1974:174) and Quirk et al. (1985:128).

include the pronominal use of ‘one’, as in (24a), frozen forms as in (24b,c), comparatives (24d), and reference to a specific occasion or person (24e,f).

- (24) a. I had one [garage] up in O’Connor and one [garage] over on Cosburn. (N/‡/m/85)
 b. But that’s one thing, when you do realize how lucky you are ... (N/fi/m/37)
 c. Then one day in the future he’s like scared. (3/H/m/12)
 d. It’s more than eleven years. (3/G/f/11)
 e. Eventually, one night she took me to the hospital. (I/§/f/21)
 f. There’s one lady that lives in my building who had been in a concentration camp. (I/£/f/83)

A further context in which approximation never occurs is following a direct inquiry about age, as in (25) (N = 16).

- (25) a. SM How old are you?
 CL Ah twenty-four. (N/O/m/24)
 b. DS How old are you now? Thirty-six you said?
 RS Thirty-four. ‘Sixty-nine. (I/œ/m/34)
 c. DW And how old are you now?
 MM Thirty-two. (I/8/m/32)
 d. MH So how old did you turn?
 KG Eleven. (2/g/m/11)

Presumably this derives from the precision of the speaker’s knowledge; other things being equal, one always knows one’s own age, and unless there is a social reason for hedging, approximation would seem odd, even impolite. Importantly, LIKE never occurs in this position either.

However, contrast this with the sequence in (26), where LIKE predictably does not occur in response to the first question, which seeks the speaker’s current age, but it is used in

response to the second question, which depends on the accuracy of the speaker's long-term memory.

- (26) KV How old are you Randy?
 RM I'm eleven years old.
 KV How old were you when you got that?
 RM **LIKE** five. (3/O/m/11)

A total of 3068 numerically quantified phrases allowing approximative modification occurred in the interview materials (2433 NPs and 635 APs). In addition to speaker age and sex, these were coded for whether LIKE occurred or not, as well as for whether they were modified by some other form.

If LIKE and adverbs such as roughly are not lexical variants, as Andersen (2001) implies, then the proportion of adverbs should remain relatively constant across the sample. At the very least, there should be no evidence of weak complementarity in the use of these forms, since such a result would be indicative of dynamic change (Sankoff & Thibault 1981:213; see also Laberge 1980; Sankoff 1982).

Of course, there are a number of alternative ways of signaling approximation in English. For example, it can be expressed through phrasal coordination (Quirk et al. 1985:269), (27a,b), or through the use of coordinating tags such as and something (like that) and or so (Biber et al. 1999:112) as in (27c,d). Among older speakers, some odd as in (27e,f) also performs this function. Other methods include suffixation with -ish, as in (27g), and the use of phrasal tags such as I guess and I think, as in (27h).

- (27) a. This repeats you-know four or five times. (3/B/m/17)
 b. So she'd be six or seven. (I/£/f/83)
 c. I would have been twenty-four or something. (I/œ/m/34)
 d. He was done school, and worked for ten years or so. (I/3/f/24)
 e. There were fifty some odd people died that night from the storm. (N/w/m/84)
 f. Oh that's fifty some odd years ago! (N/©/f/76)
 g. That was when I was you-know twenty-ish. (N/f/f/73)
 h. I took that when I was younger. Like really young, like I think ten. (3/Q/f/16)

2.2 RESULTS

Table 2.1 displays the overall results for numerically quantified contexts in the current data set. The majority of forms are not modified at all. Of those that are, the most frequent form of approximation (excluding LIKE) is adverbial, accounting for 12 percent of the data, followed by coordination at 6 percent. Tags make up the remainder at 3 percent. Included in this latter category are coordinating tags, phrasal tags, and suffixation. Note the frequency of LIKE: it occurs more than one and a half times as often as adverbs do, with a frequency of 19 percent of numerical expressions.

TABLE 2.1
Overall distribution of numerically quantified expressions

Ø		LIKE		adverb		coordination		tag		total
%	N	%	N	%	N	%	N	%	N	
61	1886	19	578	12	354	6	171	3	79	3068

Among the adverbs, about accounts for an overwhelming proportion of its category (93%; N = 330); approximately and around also occur, but their use is rare. In fact, approximately occurs just once. Biber et al. (1999:113) report that about is the most common approximating adverb in all registers and that approximately is used primarily in

academic prose. The prolific rate of about in the Toronto data is therefore typical of spoken materials.

When we consider the use of numerically quantified expressions according to the age of the speakers, a striking change emerges in apparent-time. Figure 2.1 displays the proportions of LIKE, and adverbial, coordinated, and tagged modification in the Toronto corpus.

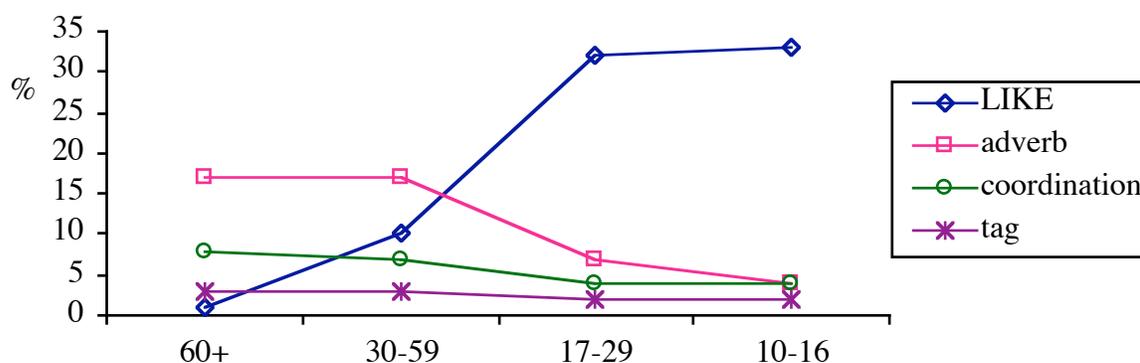


FIGURE 2.1 Distributions of numerical modifiers across apparent-time

These results indicate that the frequency of about and other adverbs of approximation is decreasing rapidly. Where among the oldest speakers in the community they occur at a rate of 17 percent (N = 618), among the youngest speakers, the 10-16 year olds, adverbs account for just 4 percent of the data (N = 509). In contrast, note the use of LIKE. It has risen from a frequency of 1 percent among speakers aged 60 and above, to 32 percent among speakers under the age of 30. This is an extraordinary rate of change.

Crucially, whereas grammaticalization may progress at variable rates (Hopper & Traugott 2003:49), lexical change is characterized by its rapidity (Chambers 2000:193). During the short period of approximately 65 years that separates the youngest and oldest

speakers, the use of adverbial approximators has declined precipitously while LIKE has increased to become the most frequent form occurring with numerically quantified expressions. This suggests that LIKE has replaced the traditional adverbs (about, around, etc.). Indeed, the complementary distribution of these forms in apparent-time is compelling evidence that they are lexical variants (Laberge 1980; Sankoff & Thibault 1981; Sankoff 1982), and therefore, that LIKE is functioning primarily as an adverb in this context. In other words, the rising use of LIKE before numerically quantified expressions represents the replacement of one form, the traditional adverbs, by another.

As Chambers (2000:194) points out, “the replacement of one word by another requires as a precursor the availability in the community of a variant that can replace it.” The results in Figure 2.1 are revealing in this regard. Among the oldest speakers, LIKE is a minority variant. In contrast, the traditional adverbs such as about prevail. The use of these latter forms then remains stable as the frequency of LIKE rises among the middle-aged speakers. In fact, the overall proportions of about, coordination and tags are level across the two oldest age groups; it is only the rate of LIKE that changes and there is robust layering of approximative strategies. This situation, however, is short-lived. Among the 17-29 year olds, LIKE increases rapidly, and it clearly does so at the expense of the traditional adverbs and possibly coordination. What we seem to have, therefore, is a situation in which a new lexical option enters the system and then briefly competes with the older forms for currency before rising to supremacy.

Because the lexicon is relatively accessible to consciousness, speakers are often aware of lexical changes:

Words come to be associated with certain social groups, and their currency waxes or wanes depending on the social status of the groups. When a word declines in frequency, it almost invariably goes through a period when its use becomes increasingly restricted to older people. (Chambers 2000:193-194)

In the case of signaling approximation, the results in Figure 2.1 suggest that once LIKE attained a critical mass of 10 percent in the speech of the 30-59 year olds, it increased rapidly among younger speakers. In fact, a separate configuration of the data (not shown here) reveals that among the 30 year olds, the forms are used in relatively equal proportions (LIKE 12% vs. adverbs 14%; N = 369). Among speakers in their mid to late twenties, a completely different pattern is evident where the use of LIKE has increased exponentially relative to the traditional forms (LIKE 26% vs. adverbs 10%; N = 262). From this point forward, the use of LIKE increases incrementally among younger speakers while the use of the about etc. continues to decline. The traditional adverbs of approximation are becoming increasingly associated with older speakers. This is identical to the findings reported in Chambers (1995), where the use of chesterfield becomes old-fashioned and quaint while couch reigns among younger speakers.

Interestingly, the replacement of the traditional adverbs does not appear to interact with speaker sex: both males and females use LIKE in virtually equal proportions (males: 18%; N = 1364 vs. females: 19%; N = 1704). Cross-tabulation by age reveals no further insights. In fact, among the youngest speakers, the 10-16 year olds, the overall frequency of LIKE is identical: 33 percent (females N = 311; males N = 198).² Thus, despite the potential aware-

² Even within the critical crossover generation, the 30 year olds, there is no significant differentiation based on sex (females: 11%; N = 194 vs. males: 13%; N = 175).

ness in the community regarding this change and the overt stigmatization of discourse uses of LIKE, these factors are not manifested through sexual differentiation of the adverbial function of this lexeme.

Subjecting these data to multivariate analysis, as in Table 2.2, confirms the lack of a sex effect.³ More importantly, two aspects of the regression results support the analysis of lexical replacement argued for here.

TABLE 2.2
Two independent multivariate analyses of the effects of age and sex on the use of LIKE and adverbial approximators

INPUT factors considered	LIKE			Adverbs		
	FW	.12 %	N	FW	.10 %	N
Age						
10-16	.78	33	509	.28	4	509
17-29	.77	32	928	.38	7	928
30-59	.45	10	1013	.64	17	1013
60+	.08	1	618	.64	17	618
range	70			35		
Sex						
male	[.52]	18	1364	[.53]	13	1364
female	[.49]	19	1704	[.48]	10	1704

* Non-significant factors enclosed in square brackets.

First, the age hierarchies are mirror images of each other: LIKE is significantly correlated with speakers under the age of 30 and the traditional adverbials are significantly correlated with those over the age of 30. Second, the corrected means ('input' in Table 2.2), which

³ The disagreement between the distributional results and the factor weights for sex in the first column is further evidence that this factor has no effect on the replacement of about etc. by LIKE as an adverb of approximation. Not only is there no female vs. male trend in terms of overall frequency but there is also some minor fluctuation across apparent-time in the relative use of this form among the sexes.

indicate the propensity for each form to be used apart from the influence of any of the environments (in this case, age and sex), are nearly identical: LIKE .12 vs. the traditional adverbial forms .10. This indicates that despite the ranges in the overall frequency of use of each form across apparent-time, the overall tendency to use either LIKE or the other adverbs is the same. In other words, the form has changed: where one increases, the other decreases. This is the defining characterization of weak complementarity (Sankoff & Thibault 1981:207), thus corroborating the analysis of lexical replacement. The strength of age, however, underscores the speed at which LIKE has risen and the traditional adverbials have been relegated to a minor variant (note the ranges of 70 and 36 respectively).

Besides the age-correlated shift in lexical variants, we can also surmise from Figure 2.1 that the overall rate of modification increases between the oldest and youngest generations: not only does the rate of LIKE in the two youngest groups surpass that of adverbs in the two oldest ones, but the proportion of utterances that include approximations also increases for the younger speakers. This is reflected in the difference in the ranges in Table 2.2: the rate of change toward LIKE has not progressed in direct proportion to the shift away from the traditional adverbs. There are two possible explanations. On the one hand, LIKE may perform extended functions in numerical contexts. That is, some instances of LIKE may be discursive while others are adverbial. If this is the case, it is unclear how to disentangle the two. On the other hand, there may be an age-graded effect at work whereby higher rates of approximation are characteristic of younger speakers. What is needed to evaluate this hypothesis is evidence from diachronic corpora, enabling a comparison of the overall proportion of modification across the community. If at two or more points in time the rate is

higher among speakers under 30 years of age, then this would be evidence of age-grading. If, however, no such trend were apparent, then we would suspect that the current situation is complicated by some other factor. I must leave the issue open, but will conclude by noting that a similar result was found for auxiliary avoir ~ être as well as for allé ~ été in Montreal French. Although correlated with socioeconomic factors rather than with age, Sankoff and Thibault (1981:210, 212) report that the slope of change was consistently steeper for the nonstandard variant (equal in this instance to LIKE) than it was for the standard one. Consequently, the overall rate of use of both variants is lower at one end of the scale than at the other. This is precisely the pattern that we see in Figure 2.1: LIKE rises faster than about declines, resulting in higher proportions of approximation among speakers under 30 than among those above this age. Perhaps regardless of the external index one is using (i.e., socioeconomic status, age, etc.), these inequalities may be inherent aspects of weak complementarity.

Before turning to a discussion of the current findings, I would like to mention the position of LIKE in numerically quantified contexts. An implicit aspect of this analysis has been that when used as an approximative adverb, LIKE occupies the same syntactic slot as other adverbials: it occurs to the immediate left of the quantifier. This too supports the replacement perspective. However, non-adverbial (i.e., discourse) uses of LIKE are never found within these phrases. This was first observed by Andersen (2001:278), and is corroborated here. Of the more than 3000 tokens considered in this analysis, not one contains LIKE in any position other than the adverbial position, exemplified above in (16)

and (21)-(23). This suggests that the hypothetical data in (28) are ruled out by a constraint against the use of LIKE within these frames.

- (28) a. * He's five **LIKE** feet tall.
 b. * He's five feet **LIKE** tall.
 c. * It takes fifteen **LIKE** minutes to get there.
 d. * It could have taken you all day to go thirty **LIKE** miles.

In Chapters 5 and 6 I will explore whether this is an arbitrary constraint on distribution or whether there are independent reasons that can account for this fact.

2.3 DISCUSSION

The distributional results in Figure 2.1 indicate that although LIKE occurs as an adverb of approximation among older speakers in Toronto, it was a minority variant in the spoken vernacular. Indeed, the apparent-time evidence suggests that during the first third of the twentieth century, it was, at best, an incipient lexical alternative, accounting for a mere 1 percent of quantified phrases (N = 612). Support for the hypothesis that LIKE has functioned as an approximative adverb for quite some time is provided by the data in (18) at the beginning of the chapter, which demonstrate this adverbial function among older speakers of both regional and somewhat isolated British varieties (Tagliamonte & Smith 2002:254-256; Tagliamonte, Smith & Lawrence 2005:79). Whatever the mechanism of change in the current materials, LIKE has since accelerated in the system to the point where it is currently the preferred adverb for expressing approximation in numerical contexts among speakers under 30. The crossover pattern in the relative proportions of this form and approximative adverbs such as about in Figure 2.1 indicates a pattern of lexical replacement, one that has

progressed swiftly in the community. Among the oldest speakers in the community, it is possible to find examples in which about and LIKE are used interchangeably, as in (29).

- (29) You-know, it was **LIKE** a hundred and four [degrees]
but it lasted for **ABOUT** two weeks. (N/w/m/84)

However, among the youngest speakers, the traditional adverbs are obsolescent, occurring no more frequently than phrasal coordination, which has been a minority form as far back as can be determined in this corpus.

In summary, the apparent-time results show a path of change in which LIKE first enters numerical contexts simply as another lexical option among many for expressing approximation. In the initial stages it competes with the already existing adverbs and the periphrastic constructions. Fairly quickly, however, it gains currency, as shown by its increase in frequency. For the youngest speakers, LIKE is the dominant form by far, taking over the system. Plainly put, the initial layering of approximative adverbials leads to the eventual ousting of the traditional forms, replaced by the innovative one in the speech of those under 30 years of age.

A question that may be raised at this point is whether LIKE and about etc. are perfect synonyms. In other words, are LIKE and about semantic equivalents? Schourup (1983:30-31) argues that about cannot always be felicitously substituted for LIKE, as in (30), and as such, they are not equivalent.

- (30) a. They've been living in this big three-story house with basements- **LIKE** four floors.
b. My hand'll get into his mouth or **LIKE** just one finger or something like that.

However, these instances do not seem to be examples of LIKE as an approximative adverb. In (30a) it can be paraphrased by “as if (it were)”, and in (30b) it seems to signal a part/whole relationship.

In a vein similar to that of Schourup, Siegel (2002:39-40) points out that while it is “odd” to agree with an approximate, the same restriction does not hold with LIKE. For example, she contends that although (31b) is marginal as a response to (31a), this effect does not hold in (32).

- (31) a. He has **ABOUT** six sisters.
 b. Yes, he has exactly six.
- (32) a. He has **LIKE** six sisters.
 b. Yes, he has exactly six.

For me, the appropriateness is marginal in both cases. Further, despite Siegel’s claim to the contrary, it would be fully acceptable in both (32a) and (32b) if the response were “No, he has exactly six.”

It is therefore my contention that these points are incidental. Moreover, perfect synonymy in language is rare, and individual forms need not overlap entirely in order to be used interchangeably. Thus, Sankoff (1988:153) writes:

While it is indisputable that some difference in connotation may, *upon reflection*, be postulated among so-called synonyms whether in isolation or in context, [...] there is no reason to expect these differences to be pertinent every time one of the variant forms is used.

I assume that in the case of LIKE and approximative adverbs such as about, any such “distinctions in referential value” are neutralized in discourse (Sankoff 1988:153; see also

Labov 1973; Sankoff, Thibault & Bérubé 1978 for lexical variation). This position is corroborated by the weak complementarity of these forms, demonstrating that any differences between them serve “no basic discourse function” (Sankoff & Thibault 1981:210).

The implication for further analyses of discourse LIKE is that numerical contexts must be treated distinctly from other nominal and adjectival constructions in which LIKE unambiguously functions as a discourse particle. Distinguishing adverbial LIKE from discourse LIKE also affects numerous claims regarding the frequency with which LIKE is reported to occur. Various authors report nominal projections to be among the three most frequent slots for this particle (e.g., Schourup 1983; Underhill 1988; Romaine & Lange 1991; Andersen 1997; Levey 2004; Wolgemuth 2003). However, failing to distinguish the adverbial function of LIKE from discourse LIKE has artificially inflated its rate in these contexts overall. Consequently, I exclude quantified phrases from further quantitative analyses of the discourse particle, providing a more defensible notion of its relative frequency in the syntactic frames in which it occurs.

Chapter 3 GRAMMATICALIZATION

One of the murkier aspects of discourse LIKE concerns its history and evolution. If we are to understand the synchronic state of this feature, then a clear understanding of its recent grammatical development is critical. The current status of LIKE as a ubiquitous discourse feature with forward scope (e.g., And there were like [people blocking], Romaine & Lange 1991:244) is espoused as a North American phenomenon which then spread across global Englishes, yet in both its ‘traditional’ (33a) and ‘innovative’ (33b,c) forms (e.g., Andersen 1997 et seq.), LIKE is argued to be the result of grammaticalization, with its pragmatic functions evolving from its role as a preposition and conjunction (Meehan 1991; Romaine & Lange 1991).¹

- (33) a. In an ordinary way **LIKE**. (OED)
b. **LIKE** she’s a space-cadet. (3/T/f/18)
c. It was **LIKE** fake. (3/O/m/11)

This leaves unexplained the relationship between the allegedly British and American uses. Specifically, are these distinct developments or does one follow from the other? I explore here the possibility that the latter has developed from the former. The framework in

¹ The classification of non-traditional uses of LIKE as a North American borrowing seems to be based, in part, on evidence from the development of quotative be like (see, for example, Andersen 2001:221). As discussed in the first chapter, however, the quotative and discourse uses of LIKE are distinct developments, and the origin of one in North America (i.e., the quotative) bears no *a priori* relationship to the origin of the other (i.e., the discourse feature).

which I embed this analysis is grammaticalization, building on recent work by Traugott (1982, 1997 [1995]) and Brinton (1996, 2005, forthcoming).

From this point forward I adopt the restrictive definition espoused by Fraser (1988), Traugott (1997 [1995]) and Brinton (2005, forthcoming) and consider as markers only those forms that occur clause-initially and which serve pragmatically to evaluate the relation of the current utterance to prior discourse. Thus, I refer to DISCOURSE MARKERS as those forms that occur on the left periphery of the sentence as in (33b). Similarly, by the term DISCOURSE PARTICLE I refer to those forms that occur in all other contexts, such as that in (33c). According to Traugott (1997 [1995]), discourse markers form a specialized subset of discourse particles. With this in mind, it is important to note that Traugott (1997 [1995]) and Brinton (1996, 2005, forthcoming) focus specifically on the development of the former type, i.e., discourse markers. As we will see, this has important ramifications for the analyses I present in Chapters 4 through 7.

This section is organized as follows. First I review grammaticalization and the way in which discourse markers are currently seen as fitting into the theory (§3.1). I then provide examples of the development of discourse markers, with evidence spanning the history of English (§3.2). Finally, I address LIKE specifically and propose a path for its diachronic development as a discourse marker in English (§3.3).

3.1 DISCOURSE MARKERS IN A THEORY OF GRAMMATICALIZATION

A long-standing line of inquiry in historical linguistics, dating back to the early nineteenth century (e.g., Bopp 1816 and Humboldt 1825, cited in Traugott & Heine 1991:2) is the

process whereby lexical items are “recruited” into the grammar as functional elements (Traugott 2003:626), i.e., grammaticalization.² However, the term itself dates only to 1912, when it was first used by Meillet to describe “le passage d’un mot autonome au rôle d’élément grammatical” (1948 [1912]:131).

Defined here as the type of change “whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions or grammatical items develop new grammatical functions” (Traugott 2001:1), grammaticalization typically results in grammatical elements such as function words, derivational morphemes and clitics. These gradually develop along an evolutionary cline such as that in (34) (Hopper & Traugott 2003:4-7).³ In Givón’s maxim (1971:413): “Today’s morphology is yesterday’s syntax.”

(34) content item > grammatical word > clitic > inflectional affix (> zero)

These developments are prototypically actualized as the nominal and verbal clines shown in (35) (Hopper & Traugott 2003:106-108), which sketch out the developmental paths along which nouns and verbs may progress as they develop functional roles.

(35) a. relational noun > secondary adposition > primary adposition > agglutinative affix > fusional affix
 b. full verb > (vector verb)⁴ > auxiliary > clitic > affix

² See Lehmann (1982) and Heine (2003) for more thorough discussions of the history of grammaticalization studies.

³ There is no requirement that items undergoing grammaticalization progress to the end of the cline (Brinton forthcoming); see Traugott (2001) for the dangers of viewing grammaticalization as deterministic.

⁴ The term “vector verb” comes from Hook (1974, 1991) who discusses verb complexes in Hindi and other Indo-Aryan languages containing a primary non-finite verb that carries full semantic content and a second ‘vector’ or quasi-auxiliary that is finite. This compound construction may occur with true auxiliaries that carry inflectional information such as person and number (Hopper & Traugott 2003:109).

The steps along these clines are not discrete. Rather, they should be interpreted as marking cluster points along a continuous trajectory (1993:107). For example, the path from noun to adposition is exemplified by the preposition beside, which developed from the noun side in the construction ‘by the side of’ (> beside), and the shift from verb to clitic (35b) is visible in English have, which functions as full lexical verb (We have a cat) but has also developed as an auxiliary (We have owned a cat). Both forms may be cliticized to the subject (We’ve a cat; We’ve owned a cat).

An implicit aspect of the clines in (35), deriving from Meillet’s definition of grammaticalization itself, is the fundamental assumption underlying grammaticalization theory, the HYPOTHESIS OF UNIDIRECTIONALITY. Unidirectionality stipulates that if a given form undergoes a change in its ‘degree of grammaticality’, it will shift from the lexical to the functional end of the scale, but not in the opposite direction (Hopper & Traugott 2003:94,126). From this we can infer that GRAMMATICALIZATION does not refer to a distinct process, but rather encompasses a “subset of cross-linguistically recurring changes” (Traugott 2001:1), i.e., the tendency for some lexical forms to develop functional roles.

To this end, it is worth noting that lexemes equivalent to LIKE have assumed pragmatic functions in a number of languages other than English. For example, Schourup (1983:32) cites data from Sierra Miwok in which the morpheme meaning ‘like’ (nymiŋ) can also be used parenthetically to mean ‘as it were’ (cf. (4) in Chapter 1). This is shown in (36a). Hebrew has a discourse feature akin to English LIKE, as in (36b) from Maschler (2001:296). We can also find equivalents in Bislama (36c) (Meyerhoff & Niedzieski 1998:239), in Canadian French (36d) (Sankoff et al. 1997:198), and in Japanese (36e) (Lauwereyns

2002:244). Finally, (36f) is the Norwegian variant liksom (Hasund 2003:138) and (36g) exemplifies the Swedish equivalent (Kotsinas 1994:93).

- (36) a. mu-uj-NYMIJ
 ‘in the trail, as it were’
- b. Veke’ilu haragláyim sh’xa nitka’ot bifnám KAZE.
 ‘And like your feet get stuck inside like.’
- c. Afta OLSEM hem I jas kambak.
 ‘And like he’d just come home.’
- d. Ah oui on était COMME un des seuls.
 ‘Oh yeah we were like the only ones.’
- e. NANKA atsukunain da yo NANKA attakai tte kanji.
 ‘Like, it’s not hot, y’know, like, it feels like warm.’
- f. Men det er LIKSOM litt harray.
 ‘But they’re like sort of naff.’
- g. Nämen LIKSOM de låter så enkelt.
 ‘No but like it sounds so simple.’

All of this suggests that the development of discourse functions for LIKE in English is not idiosyncratic, but derives from semantic properties that are shared cross-linguistically (see also Meyerhoff 2002:354). It is this universal aspect of LIKE and the changes it may undergo that bolster the grammaticalization argument.

The shift from major (i.e., open) category to minor (i.e., closed) category by a grammaticalizing form has been shown to progress according to a set of recurring principles, including bleaching, generalization, decategorialization, phonological attrition, divergence, and subjectification. Because the diachronic development of discourse markers shares many of these characteristics, a growing body of literature argues for their inclusion of under the umbrella of grammaticalization (e.g., Traugott 1982, 1997 [1995]; Brinton 1996,

forthcoming). As such, it is important to be clear about what exactly is meant by each of these terms. In what follows, I briefly define each; for fuller illustration, see section 3.2.

BLEACHING (Meillet 1948:139; Bybee & Pagliuca 1985:74; Lehmann 1985:307; see also Gabelentz 1891; Givón 1973 and Fleischman 1982) denotes the loss of propositional meaning that is typical of many grammaticalized forms. While this weakening is sometimes seen as a defining characteristic of grammaticalization (e.g., Heine & Reh 1984:15), it is now recognized as a feature of later stages. The early stages, however, are typified by pragmatic strengthening, wherein the significance of invited inferences increases (Traugott 1988; Traugott & König 1991; Hopper & Traugott 2003). In the development of the ‘going to’ future in English, for example, Langacker (1990:23) points out that while the lexical verb go loses its reference to locational orientation and movement, this meaning is replaced by the abstract one of the speaker’s temporal perspective. In short, there is a balance between the “loss of older, typically more concrete meanings, and development of newer, more abstract ones” (Hopper & Traugott 2003:96).

DECATEGORIALIZATION (Heine et al. 1991:213; Hopper 1991:30-31; Hopper & Traugott 2003:103-113) refers to the systematic loss of the morphological and syntactic properties that are diagnostic of membership in a particular lexical category. Thus, decategorialization of a noun, for example, entails the eventual inability to carry inflectional information and/or to be a topic (Hopper & Traugott 1998:103). This is evident in the grammaticalization of *SIDE* discussed above: as a noun, *SIDE* can host plural morphology, while as a preposition it cannot.

PHONOLOGICAL ATTRITION (Bybee & Pagliuca 1985:60; Lehmann 1985:307; Heine et al. 1991:214-215), also known as EROSION, is where a form gradually loses phonological substance. A well-known example is the development of the ‘going to’ future, which becomes progressively reduced relative to its degree of grammaticalization: going to > gointa > goina > gonna (Krug 2000; Poplack & Tagliamonte 2000).

DIVERGENCE (Heine & Reh 1984:57-59; Hopper 1991:24-25) arises when the original lexical item that underwent grammaticalization remains an autonomous element. This results in etymologically related forms with distinct functions, such as French pas ‘not’, which derives from pas ‘step’, and the English indefinite article a(n) from one (Hopper 1991:24). One further example includes LIKE, which, in addition to its discourse roles, continues to function as preposition (2c) and conjunction (2d).

SUBJECTIFICATION (Traugott & König 1991; Traugott 1995; 1997 [1995], 1999; Traugott & Dasher 2002) refers to the development of meanings that encode the speaker’s viewpoint of the speech event. This shift in perspective is evidenced by the development of must as an epistemic marker (i.e., ‘based on the evidence I conclude that’) in English from its earlier function where it encodes deontic meanings of obligation.

GENERALIZATION (Hopper & Traugott 2003:100-103; Heine 2003:579-580) obtains when a form comes to be used in contexts where it could not be used before. Hopper and Traugott (2003:100) cite the development of the English progressive, which was originally constrained to agentive constructions (e.g., They were building a house) and later spread to passive constructions (e.g., The house was being built) followed by stative contexts (e.g., There are statues standing in the park).

In the following chapters, it is this last characteristic of grammaticalization that will be seen to play a critical role in the development of discourse uses of LIKE. On the assumption that the development of discourse markers is characterized by the principles just outlined, a point I illustrate in the following section, subsuming the evolution of these forms in the paradigm of grammaticalization necessarily entails redefining grammatical function to include the pragmatic functions performed by discourse features, including LIKE. This means that in referring to “the grammar” of a language, we are referring not only to the cognitive aspects of phonology, morphology, syntax and semantics, but also to the communicative aspects that draw on the linguistic forms produced by these modules. This includes the pragmatic inferences encoded in words, phrases and the discourse in which their combination results (Traugott 1997 [1995]:5; 2003:626; see also Brinton forthcoming). In taking up this line of argumentation, it is important to stress that discourse markers not only interact with prosodic structure, carrying distinct intonational and stress patterns (Allerton & Cruttenden 1974; Altenberg 1987; Schiffrin 1987; Fraser 1990; Horne et al. 2001), but they also occupy a specific syntactic slot, adjoined to the left periphery of a sentence (Traugott 1997 [1995]:6; see also Kiparsky 1995:140-141).

Brinton (forthcoming) finds evidence for at least three pathways in the development of discourse markers. Due to the intricate nature of these trajectories, they are grossly summarized for purposes of exemplification in (37).

- (37) a. adverb/preposition > conjunction/sentence adverb > discourse marker
 b. matrix clause > matrix clause/parenthetical disjunct > discourse marker
 c. subordinate clause > parenthetical disjunct

Unlike the forms that develop via the nominal and verbal clines in (35), discourse markers evolve from a wide range of linguistic items including nouns, verbs, and adverbs, as well as idioms, phrases and clauses (Fraser 1996; Brinton forthcoming). Consequently, no single cline can account for discourse markers as a class because the developmental pathway derives from the syntactic origin of the individual form (Brinton forthcoming). For example, the cline in (37b) captures the development of forms such as methinks, you know, prithce, I'm afraid, and I say as pragmatic expressions from their original role as matrix clauses requiring 'that' complements (see Brinton forthcoming and references therein). The cline in (37c) accounts for the development of epistemic comment clauses such as I think, I guess, and I suppose, when deletion of the complementizer creates a parenthetical disjunct.

The trajectory of critical import here, however, is (37a), which is the route that has been proposed for LIKE (Romaine & Lange 1991) as well as for a number of other discourse markers including why, so, and now (e.g., Traugott 1982; Aijmer 1988; Blakemore 1988; Schiffrin 1987; Traugott & König 1991). Each of these forms has been argued to develop along a pathway from preposition to conjunction to discourse marker. The trajectory in (37a) also traces the development of markers like indeed and actually in Modern English, anon in Middle English, and witodlice and soðlice in Old English. The difference between these forms and LIKE is that their development is not posited to include a stage as a conjunction; rather, they are believed to have developed directly from adverbial elements. I will elaborate upon the relevance of this distinction in section 3.3 below.

For now, it is important to note that despite the differences in the pathways in (37), they all share the unidirectionality of development in the degree of structural scope. Specifically,

in the grammaticalization of discourse markers we can expect to find a trajectory of the following sort: scope within a proposition > scope over the proposition > scope over discourse (Traugott & Dasher 2002:40; Brinton forthcoming), the last being a defining characteristic of discourse markers. As we will see, it is this aspect of the theory that will have the greatest consequences for the analyses of LIKE to follow.

3.2 EXEMPLIFICATION FROM ENGLISH

In order to illustrate these stages, I now look briefly at two case studies that exhibit the characteristic traits of grammaticalization. In the historical development of these forms, soðlice and indeed, evidence is provided from Old and Modern English. Both of these discourse markers illustrate the path adverb > sentential adverb > discourse marker, and this is the trajectory I will propose for LIKE.

3.2.1 *Soðlice*

Lenker (2000) discusses the evolution of soðlice as a discourse marker. In Old English, soðlice ‘truly’ was a manner adverb that took scope within the predicate, as in (38a). It derives from the adjective soðlic, itself a derivative of the noun soð, preserved in the archaic forms forsooth and soothsayer. Once a manner adverb, soðlice began to function as a sentential adverb, signaling the speaker’s belief that his words are true (38b). Thus, soðlice here has full sentential scope as well as increased subjectivity, since it has come to encode the speaker’s viewpoint of the speech event. The final step in the development of soðlice

from syntax to discourse is the evolution of textual marking functions, where it is argued to draw attention to a particular sequence of discourse, as in (38c).

- (38) a. Ic eam **soðlice** romanisc, and ic on hæftnyd hider gelæd wæs.
 ‘I am truly a Roman, and I was brought hither in captivity.’
 (Lives of Saints, Eustace 344)
- b. Wæs **soðlice** on rightwisnyse weorcum ... swiðe gefrætwod.
 ‘Truly he was greatly adorned ... with works of righteousness.’
 (Lives of Saints, Eustace 4)
- c. **Soðlice** on þam dagum wæs geworden gebod from þam casere augusto.
 ‘Truly, in these days an order was given by the Emperor Augustus.’
 (The Gospels, Luke 2,2)

Thus, this trajectory (noun > adjective > adverb > discourse marker) demonstrates decategorialization, as soðlice shifts from a major word category to increasingly minor ones. Each of these stages represents generalization, as soðlice comes to be used in new contexts. This form also exemplifies increased subjectification, particularly in the development of the sentence adverb. Further, in the progression from adverb to discourse marker, soðlice demonstrates the expected increase in scope, from within the proposition as a manner adverb (38a), to over the proposition as a sentential disjunct (38b), to over discourse as a discourse marker (38c).

3.2.2 *Indeed*

The discussion of the development of indeed is based on Traugott (1997 [1995]). During the Early Middle English period, the noun deed came to be used in the formulaic construction in dede meaning ‘in action’, as in (39a). Traugott (1997 [1995]) shows how this collocation then began to function as an adverb with an evidential meaning of ‘in

actuality’, as in (39b). In this function, in dede has narrow structural scope within the predicate. By the Early Modern English period, in dede occurs clause-initially (39c) used in the evidential sense of (39b), but as a sentence adverb it has scope over the proposition. Finally, this clause-initial context begins to develop meanings of elaboration or clarification (39d), i.e., indeed has developed discourse marking functions.

- (39) a. Al þat þou hauest her bifore i-do, in þohut, in speche,
 All that thou hast here before done, in thought, in speech,
 and **in dede**, in euche oþeres kunnes quede, ich þe forzeue.
 and in action, in each other kind’s evil, I thee forgive.
 ‘I forgive thee for all that you have done heretofore, in thought, in action,
 and in evils of every other kind.’ (c. 1300 Fox and Wolf 34)
- b. but plainly report as it was **in dede** nouȝt sparing for no persone
 but plainly report as it was in fact not sparing for no person
 ‘but plainly report it as it was in actuality, not sparing anyone’
 (1437 Doc. Chancery 168)
- c. somtyme purposely suffrint the more noble children to vainquysse,
 and, as it were, gyuying to them place and soueraintie, thoughe **in dede**
 the inferiour chyldren haue more lernyng. (1531 Governor 21)
- d. any a one that is not well, comes farre and neere in hope to be made
 well: **indeed**, I did heare that it had done much good, and that it hath
 a rare operation to expell or kill diuers maladies. (1630 Penniless Pilgrimage 131)

Like the development of soðlice, the evolution of indeed is characterized by generalization to new linguistic contexts. The fact that deed persists in Modern English as a full noun is an example of divergence. Importantly, however, indeed also provides an example of pragmatic strengthening: its pragmatic functions take on increasing importance as this form develops into a discourse marker and the original semantics of ‘in action’ (> ‘in actuality’) are gradually stripped away.

3.3 DISCOURSE ‘LIKE’

I now turn my focus to LIKE and, building on previous research, explore its diachronic development as a discourse marker. In their seminal article on the origins of discourse (and quotative) (be) like, Romaine and Lange (1991) suggest that it has arisen following the course of preposition to conjunction to discourse marker, as shown in Figure 3.1.

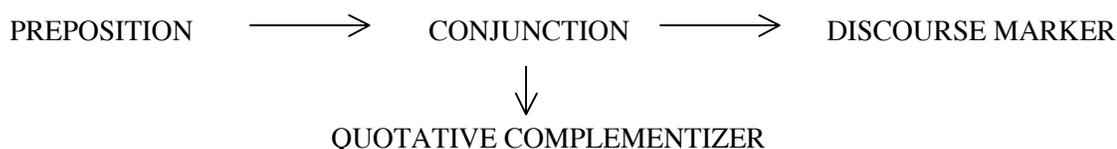


FIGURE 3.1 Grammaticalization of LIKE (Romaine & Lange 1991:261)

As a preposition, LIKE takes a nominal or pronominal complement (e.g., It looks like a snail). The preposition then undergoes recategorialization to a conjunction, enabling it to subcategorize for a sentential complement (e.g., It felt like everything had dropped away.). Romaine and Lange (1991:261) argue that this shift involves treating a subordinate clause as a nominal constituent to which LIKE is extended via analogy. The development of quotative be like follows rather straightforwardly from this point when the clause introduced by LIKE is a quotation. Since English requires all clauses to contain a verb, a dummy ‘be’ is inserted to license LIKE in this context, syntactically embedding it in the quotation frame (*ibidem et passim*); LIKE is thus reanalyzed as a quotative complementizer.

Just as with the quotative form, the discourse marker is shown in Figure 3.1 as deriving directly from LIKE as a conjunction, yet in their description of its evolution, Romaine and Lange (1991:261) state:

Because LIKE can appear as a suffix following an item, as well as precede a clause or sentence, it can be reanalyzed as a discourse marker, which shows syntactic detachability and positional mobility.

This account is problematic. First, the implication of the suffix (e.g., The sculpture looked quite human-like; Romaine & Lange's (24b)) directly contradicts the hypothesis of unidirectionality, an issue pointed out by the authors themselves (see Chapter 1, fn.6), yet their analysis is firmly embedded within a grammaticalization framework. Implicating the suffix is also unmotivated, since the link between a dependent morpheme and 'syntactic detachability' seems contradictory. Second, this account appears to posit two sources for discourse LIKE: the conjunction and the suffix. As far as I am aware, such a development would set a precedent in the grammaticalization literature.⁵ Admittedly, this interpretation may be too strong. Romaine and Lange may simply be suggesting that it is the breadth of scope that LIKE as a morpheme may take — ranging from an individual word at the affix level to an entire clause at the conjunction level — that enables this form to be reanalyzed as a discourse marker, and ultimately, as a discourse particle. If so, then scope is seen as a causal factor in the development of a pragmatic device, rather than as a concomitant of its development more generally.

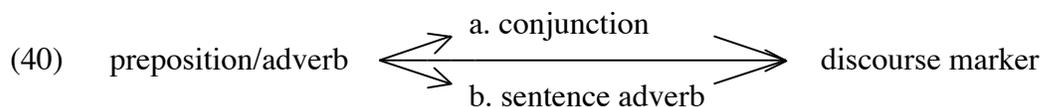
Where Romaine and Lange (1991) focus primarily on the syntactic trajectory of discourse LIKE, Meehan (1991) concentrates on its semantic and pragmatic evolution. Examining its various meanings historically, Meehan (1991) tracks the development of LIKE

⁵ It is true that a single grammatical function may derive from divergent sources, such as the English future, which may be encoded with both be going to from the motion verb go as well as with will, derived from the volitional verb willan (Heine 2003:591). Note that in these instances, we find distinct forms within the functional domain, not a single form as the above suggestion for LIKE entails and further, that these developments occurred at very different points in the history of English.

from an original comparative denoting ‘similar to’ to a focus element with little or no semantic content. Romaine and Lange (1991:262), however, express concern over a linear model of this type and suggest instead that the development of discourse LIKE may involve “a network of related meanings.” Brinton (1996:111) supports this position and argues for the need to allow for “quite varied semantic routes from source to target” in grammaticalization paths (see also Buchstaller 2001).

What remains is a schema for fitting the semantic, pragmatic, and syntactic evidence together. First, as discussed above, the syntactic facts surrounding the development of discourse uses of LIKE require further clarification. It should be noted that since Romaine and Lange’s (1991) groundbreaking analysis, the theory surrounding the development of discourse features has advanced considerably. From this perspective, Brinton (forthcoming) reviews the literature on a number of pragmatic markers, among them LIKE, and states that whether they develop from a conjunction, an adverb, or some other way awaits “fuller diachronic exploration.” This hesitancy stems from diachronic studies of markers such as hwæt and ða (Brinton 1996) where there is no conclusive evidence that they develop from conjunctions, appearing instead to derive from adverbs.

This being said, of the three trajectories proposed in Brinton (forthcoming), the development of LIKE is a variant of (37a), elaborated here as (40):



The direct pathway from adverb to discourse marker discussed in Brinton (1996) is captured in (40) by the middle arrow that progresses directly across the cline from left to right. Path (a) allows for the possibility that discourse uses of LIKE derive from the conjunction, as suggested by Romaine and Lange (1991). Path (b) accounts for the development of forms like soðlice and indeed, discussed in section 3.2 above. These latter forms progress along a trajectory in which their structural scope increases with every shift rightward on the cline: within the proposition as adverbs, over the proposition as parenthetical sentential disjuncts, and finally over the discourse as markers. I would like to pay particular attention to this last path, (40b), because there are reasons to consider whether LIKE has developed via a similar one rather than directly from the conjunction, as in (40a).

Consequently, I return to the ‘traditional’ form of discourse LIKE discussed earlier. This usage, exemplified in (4) and repeated below as (41), situates LIKE in clause-final position where it takes backward scope.

- (41) a. Father grew quite uneasy, **LIKE**, for fear of his Lordship’s taking offence.
(1778 F. Burney Evelina II. xxiii. 222)
- b. Of a sudden **LIKE**. (1801 tr. Gabrielli's Myst. Husb. III. 252)
- c. In an ordinary way **LIKE**. (1826 J. Wilson Noct. Ambr. Wks. 1855 I. 179)
- d. If your honour were more amongst us, there might be more discipline **LIKE**.
(1838 Lytton Alice ii. iii)
- e. “Why **LIKE**, it’s gaily nigh like to four mile **LIKE**.”
(1840-41 De Quincey Style ii. Wks. 1862 X. 224)
- f. Might I be so bold as just to ax, by way of talk **LIKE**, if [etc.].
(1870 E. Peacock Ralf Skirl. I. 112)
- g. He hasn’t passed his examinations **LIKE**... He has that Mr. Karkeek to cover him **LIKE**. (1911 A. Bennett Hilda Lessways i. vi. 49)
- h. “You’re a chauvinist,” Danny said. “Oh, yeah. Is that bad **LIKE**?”
(1961 New Statesman 22 Sept. 382/2)

As noted in the first chapter, LIKE in this construction is parenthetical and can be glossed as ‘as it were’ or ‘so to speak’. Note too that it is adjoined to the right edge of the clause as a sentence adverbial with wide scope over the proposition. This means that the history of LIKE includes a stage identical in form and function to the development of a number of other discourse markers throughout the history of English (cf. soðlice and indeed).

This raises the question: Is this use independent of the ‘innovative’ (North) American one where LIKE precedes the material it modifies, as has generally been assumed (e.g., Romaine & Lange 1991; Andersen 1997 et seq.; Hasund 2003), or are the two forms related? That the ‘innovative’ use occurs historically outside North America (cf. (5) and (6)) and is synchronically productive across varieties of English and among speakers of all ages (cf. (7)) is strongly suggestive of the latter.

LIKE has functioned as a conjunction (2d) since the fourteenth century (Romaine & Lange 1991, fn. 6 271; OED). Despite the clamor raised during the 1950s by the famous advertisement “Winston tastes good like a cigarette should,” Romaine and Lange (1991:244) note that this conjunctive function has likely been established for centuries and was possibly always favoured over the standard ‘as’ in colloquial usage. Note that as a conjunction, the scope of LIKE has broadened from within the predicate as a preposition, to over it, but a core meaning of comparison or similarity is firmly entrenched.

In the century after LIKE developed as a conjunction, the meanings of comparison and similarity began to be extended to include connotations of resemblance and approximation, seen in (42):

- (42) a. The ... seid principal governauncis been of **LIJK** state, condicioun, nature and merit with this present ... principal gouernaunce. (1449 Pecoock Repr. iv.vii 458)
 b. A much more **LIKER** image of God are those good Princes that loue and worshippe him. (1561 T. Holy tr. Castigione's Courtyer. iv. T iij a)
 c. My great Conversion from prodigious Profanese to something **LIKE** a Moral Life. (1666 Bunyan Grace Ab. Section 32)
 d. I have had nothing **LIKE** a bad fall lately. (1791 G. Gambado Ann. Horsem i. 67)

Note that with the extended meanings in (42) LIKE becomes increasingly associated with speaker attitude. In (42d), for example, what constitutes a 'bad fall' is already subjective, since the designation of 'bad' can vary from person to person. This subjectivity is compounded by the speaker professing his opinion that he has had nothing resembling what he considers to be a bad fall.

I would like to suggest that this increasing tendency to be oriented in the speaker's perspective eventually led to LIKE being analyzed as an epistemic stance marker (i.e., a marker of speaker attitude). This is the form that begins to occur clause-finally during the eighteenth century (41) and which must have been transplanted to North America, since it occurs in the speech of older segments of the population synchronically, as shown in (43), albeit rarely.

- (43) a. I had my thing from the school that I passed my entrance **LIKE**. (N/ɜ̃/m/85)
 b. We need to smarten it up a bit **LIKE**. (N/©/f/76)
 c. You'd hit the mud on the bottom **LIKE**. (N/œ/m/62)

In this function, LIKE provides metatextual commentary, signaling to the listener that the proposition only resembles or approximates reported events; it is not meant to be taken literally or verbatim. As a parenthetical sentential adverb, LIKE continues to take scope over the proposition as it did as a conjunction, but its function has shifted to the pragmatic

domain. This is suggestive of pragmatic strengthening, particularly as the notion of similarity has begun to weaken (i.e., it is no longer the central meaning). Positing this stage as an epistemic marker is reminiscent of Andersen's (1997 et seq.) relevance-theoretic analyses where LIKE is argued to mark 'loose talk' (§1.1, Chapter 1).

From its function as a sentential adverb, LIKE developed meanings of elaboration or clarification of discourse intent. In other words, it took on the full discourse marking function (e.g., Fraser 1988:31). In so doing, the scope of LIKE broadens from over the proposition to over the discourse where it "marks relations between sequentially dependent units" (Traugott 1997 [1995]:5).

Note that in its function as a discourse marker, LIKE occurs on the left periphery of the sentence, while as a sentence adverbial it surfaces on the right periphery. While I have no straightforward explanation for this positional shift, it is worth observing that wide-scope adverbials may occur in final position (e.g., Ernst 1984, 2002) while discourse markers tend to occur in initial position (Traugott 1997 [1995]:6). The fronting of the discourse marker may therefore follow from this latter observation. Of course, both sentential adverbs and discourse markers are adjuncts, which in English are not linearized; they may appear on either side of the phrase they modify (Adger 2003:112). Moreover, LIKE is not anomalous: a number of discourse markers have displayed this same type of positional mobility during their development (e.g., look'ee, mind you, I'm sorry, I say, etc.; see examples in Brinton forthcoming).

Regardless, it is clear that LIKE developed a discourse marking function and that in this function it occurs on the left periphery of the sentence. The question is, When did this

occur? Given the belief that uses such as those in (44) have their roots New York's counterculture groups of the 1960s (Andersen 2001:216; OED), the answer to this question seems particularly germane.

- (44) a. **Man, LIKE** the dude really flashed his hole card.
 (1971 Black Scholar Apr.-May 26/1)
 b. **LIKE**, it seems to me that it would be virtually impossible to avoid some contradictions. (1973 Black Panther 17 Nov. 9/4)

As can be seen in (45), LIKE is currently used as a discourse marker by speakers in their seventies and eighties, in both Toronto (45a,b) and across the United Kingdom (45c-e) (from Tagliamonte to appear; see fn. 5, Chapter 1):

- (45) a. **LIKE** our daughter was turning sixteen, and the little girl down the street is sixteen, was given a car for her birthday. (N/d/m/83)
 b. **LIKE**, the week had already gone by so I only got quarantined for probably four days. (N/Q/f/72)
 c. **LIKE**, my neighbours and we got on fine. (AYR/E/f/78)
 d. **LIKE**, you forget that's on at the finish, don't you? (MPT/z/m/78)
 e. **LIKE**, it was a kind-of wee bit of tongue-twister. (CLB/a/f/89)

Assuming that a person's speech changes little over the course of their lifetime (Labov 1963, 1966, 1994), these data suggest that the discourse marking function was already established during the pre-World War II era, at least three decades before the examples in (44). We can extrapolate from this that LIKE began to develop as a discourse marker at some point during either the late nineteenth or early twentieth century. This chronology would allow for it to have been present in the ambient language, enabling the speakers in (45) to acquire it. Crucially, the data in (45c-e) as well as those in (7b,c) also indicate that this function was not restricted to North America. At the same time that we have evidence for LIKE as a

discourse marker in Canada, we have evidence for this use across regional varieties in England, Scotland, and Northern Ireland. In short, not only do these examples pre-date the source that has been hypothesized to this point, they also contradict a strictly North American origin.

Consequently, I would like to suggest that the ‘traditional British’ use of LIKE (Andersen 1997 et seq.), where it occurs in final position and has backward scope, is reflective of an earlier stage in which this form functioned metatextually as a sentential adverb. Its origins in British varieties are not disputed; only its status as a dialectal variant distinct from, yet functionally equivalent to, the ‘innovative American’ use is at issue. This form, where LIKE takes forward scope, reflects a later discourse marking function. In short, these are not separate, variety-specific, developments. Rather, they represent different stages of development in the evolution of a discourse feature. Thus, the trajectory for LIKE parallels that of other discourse markers such soðlice and indeed, where it progresses along the path: conjunction > sentence adverb > discourse marker (i.e., (40b)). The use of the term ‘conjunction’ to refer to LIKE when it takes a clausal complement stems from Quirk et al. (1985:661) and Romaine and Lange (1991:244) and is used here for continuity between the present study and past discussions. It should be noted, however, that this terminology is not definitive, since the OED labels LIKE in its conjunctive function an adverb. Thus, the cline suggested here for the development of LIKE is consistent with those proposed in Traugott (1997 [1995]) or Brinton (forthcoming).

In summary, the path taken by LIKE as it grammaticalizes is similar to other forms (e.g., soðlice, indeed) that followed the route from adverb to sentential adverb to discourse

marker. As LIKE progresses along this cline, it exhibits many of the hallmarks of grammaticalization, including decategorialization (preposition > conjunction > discourse marker), bleaching (weakening of the concrete sense of comparison and/or similarity), subjectification and pragmatic strengthening (development of epistemic meanings and pragmatic functions), and divergence, since LIKE continues to perform as a preposition and conjunction in addition to its role as a discourse marker. Most importantly, LIKE exhibits the archetypal increase in scope characteristic of the development of discourse markers (Traugott & Dasher 2002:40; Brinton forthcoming).

At the same time, there is a critical aspect of LIKE that does not seem to follow from this model of grammaticalization. Specifically, this model is designed to account for the development of discourse MARKERS. The vast majority of occurrences of LIKE, however, consist of its use as a discourse PARTICLE (Andersen 2001:276), as in (46):

- (46) a. I remember there being **LIKE** a solar eclipse. (I/~f/29)
 b. It **LIKE** went **LIKE** seamlessly into it. (N/p/m/20)
 c. She's **LIKE** really smart. (2/m/f/12)

When LIKE functions in this way, it is argued to be syntactically bound to and dependent on the linguistic structure it modifies (Andersen 2001:273). In other words, as a particle, LIKE takes narrow structural scope within a proposition. It does not take wide scope over the discourse. This is the absolute opposite of what is predicted in the development of discourse markers, which, recall, Traugott (1997 [1995]) considers a sub-set of discourse particles. It appears therefore that the syntactic versatility for which LIKE is renowned in fact violates the unidirectional cline outlined by Traugott and Dasher (2002:40) and Brinton (forthcoming).

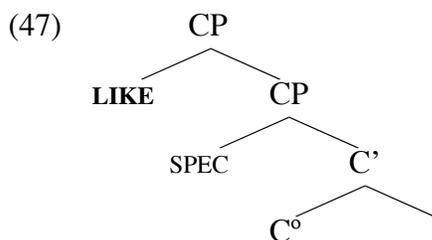
Thus, it seems possible that the development of LIKE is continuing. Over the course of the next few chapters I address this possibility by looking at LIKE in a large body of speech data. I focus in turn on each of the three positions that have been reported to be the most frequent slots for LIKE, clause-initial (Chapter 4), noun phrases (Chapter 5), and verb phrases (Chapter 7), and add adjective phrases (Chapter 6) to the investigation, as these are implicated by the analysis of noun phrases. As we will see, the order in which these chapters are presented reflects the developmental path of LIKE as a discourse-pragmatic feature. Consequently, I return to the issue of development in Chapter 8, where I assess the evidence provided by the current investigation and contextualize the findings. For now, it is sufficient to note that it is only in its function as a discourse particle that LIKE appears to behave contra the predicted pattern of broadening scope. In its discourse marking function, where it occurs clause-initially and brackets sequences of discourse (Fraser 1988, 1990; Schiffrin 1987; Traugott 1997 [1995]), LIKE satisfies the expected tenets of grammaticalization.

Chapter 4

THE CLAUSE-INITIAL CONTEXT

I begin with an analysis of LIKE in the clause-initial context, where it performs textual discourse marking functions. First, the grammaticalization clines proposed by Traugott (1997 [1995]) and Brinton (forthcoming) specifically address the rise of discourse markers. These occur clause-initially. Second, as suggested in the last chapter, it is possible that the grammaticalization of LIKE is ongoing. If this is true, then a clear understanding of its current state of development and use as a discourse marker, the focus of the clines, is necessary in order to contextualize its development and use as a particle. Third, the clause-initial context is claimed to be a highly frequent position for LIKE (e.g., Underhill 1988; Romaine & Lange 1991; Andersen 2001; Levey 2004; Tagliamonte 2005). In Andersen (2001:273), for example, one third of all tokens consist of the discourse marker. These points combine to make the marker an appropriate point of departure for an accountable, and structurally delimited, quantitative analysis of discourse LIKE.

Some clarification of terminology is required. I have described the position of LIKE when it functions as a discourse marker as the left periphery of a sentence, where it takes scope over the discourse. This approach follows from Fraser (1988, 1990), Traugott (1997 [1995]), and Brinton (forthcoming), and is in line with the previous analyses of LIKE listed above. Following Traugott (1997 [1995]), I assume that discourse markers fill the syntactic adjunct slot, adjoining to CP as exemplified in (47).



The left periphery of the clause has been the subject of much recent attention in the literature (e.g., Rizzi 1997, 2001, 2002), and has come to be seen as consisting of distinct functional heads and their projections, much as has happened with the finely articulated DP and VP systems that are argued for in current Minimalist Theory (see Chapters 5 and 7 respectively and references therein). For the purposes of this chapter, the simplified cartography of the CP in (47) will suffice (though see fn.10, this chapter). The crucial point here is that any CP, regardless of level (i.e., matrix and subordinate), has the potential to host LIKE. It is not necessarily the case, however, that every CP is sentence-initial. That is, a CP is not restricted to the top-most projection of syntactic structure. For this reason, I focus on the clause, where “clause-initial” refers to the left edge of both matrix and subordinate CPs but presupposes nothing regarding its position in the full structure of a sentence.

The treatment of discourse markers as syntactic adjuncts (as opposed to, for example, the head of their own functional category) has repercussions for the analyses throughout this dissertation. In the architecture of the Minimalist framework (Chomsky 1995, 2000, 2001), adjuncts, like adjectives, adjoin to the phrasal level, X^{\max} . By defining the variable context for LIKE along structural parameters, for example AP or VP, the adjunction site is implicationally assumed to be the maximal projection heading each structure. As I will

show, this analysis makes the correct predictions regarding the distributional patterns we find in the contexts where LIKE is used.

This chapter is structured as follows. I first outline the methodology, section 4.1, describing the circumscription of the variable context and the coding of the data. In section 4.2, I report my findings, dividing the results into two sections: matrix clauses (§4.2.1) and subordinate clauses (§4.2.2). The rationale behind this division is the following. Clause-initial (i.e., CP adjunction) is not synonymous with sentence-initial. Before collapsing these as the instantiation of a single variable context, we should be sure that in so doing, we are not missing greater generalizations in the data. Finally, in section 4.3 I present a brief summary and discussion of the results.

4.1 METHODOLOGY

4.1.1 Circumscribing the variable context

As just discussed, the focus of this chapter is the clause-initial context, where I define a clause as a CP projection. Because I am applying accountable methodology, CPs were extracted from the interview materials whether they contained LIKE or not (see (11), Chapter 1). However, given the sheer volume of available data, it was not feasible to extract every CP from every speaker. With this in mind, 75 tokens were randomly extracted from each speaker in the sample (Table 1.1), ensuring enough data to allow for an authentic model of variability as well as balanced cells across the corpus.

At both levels of structure, the analysis is limited to declarative CPs. Once the data were compiled, over 99 percent of the tokens were declarative (N = 6260). The remaining clauses consisted of interrogatives (N = 45) and imperatives (N = 10), too little for

meaningful analysis. It should be noted however that both interrogatives and imperatives allow LIKE, as exemplified in (48).¹

- (48) a. **LIKE** how does that work? (N/Δ/f/40)
 b. **LIKE** why don't you go beyond that? (2/d/m/11)
 c. **LIKE** don't worry about it. (2/i/f/19)
 d. **LIKE** don't believe them and stuff. (2/b/f/15)

In circumscribing the context of variation, a number of constructions were excluded. The first is false starts, shown in (49), defined here as those utterances when a speaker fails to complete a thought, changes direction, or hesitates. Only full CPs consisting of a subject, verb and — if transitive — its complement, were included in the analysis.

- (49) a. **LIKE** you- you were- (N/Z/f/60)
 b. **Ø** And this book **LIKE** to!-- (I/7/m/35)

The examples in (50) emphasize the importance of excluding sequences involving false starts: the position of LIKE is ambiguous. Based on the transcriptions alone it cannot be determined whether it constitutes part of the false-started sequence or whether it functions as a discourse marker on the reformulated utterance. Consequently, any tokens of this nature were excluded.

- (50) a. I'll go to **LIKE**- I'd rather go to one game and buy good tickets than go to seven games and have crappy ones. (2/h/m/18)
 b. Well, years ago when **LIKE**, when we were young you-know there was- there wasn't that big bridge, you-know, like, the skyline? (N/©/f/76)

¹ See Chapter 7 (§7.1.1) for further discussion regarding imperatives.

What can be inferred from (49) and (50) is that LIKE as a hesitation device (Andersen 2001:269-271; see also Schourup 1983) forms no part of the current analysis. Because all false start contexts were excluded, the remaining ones can be assumed to constitute the discourse marking position and function.²

A further context that was not included in the analysis are tokens where LIKE does not occur, but had it surfaced, it would function as a conjunction, meaning ‘as (if)’. Examples are given in (51), where the clauses marked with an asterisk were removed from consideration.

- (51) a. LIKE I felt a sensation
 * __ a car might come. (3/G/f/11)
- b. I really feel
 * __ it’s gone. (N/u/f/49)
- c. I like to know
 that they’re watching out for me.
 * __ I watch out for them. (N/v/f/55)
- d. It wasn’t a degree-course.
 * __ It is now. (I/∞/f/75)

Finally, there are three contexts that do not appear to be variable positions for LIKE. The first of these is within an enumeration, as in (52) (for a variationist perspective of enumeration, see Dubois 1995, 1997; Dubois & Sankoff 2001). In these types of sequences, LIKE may occur with the first clause introducing the unit, as in (52a), but it is never used

² Nevertheless, the tendency for LIKE to occur in the context of false starts and reformulated utterances is a recurrent finding (e.g., Andersen 2001; Buchstaller 2001; Wolgemuth 2003) and is deserving of further attention. However, the issues likely intersect with psycholinguistic factors such as processing time (e.g., Fox Tree 1995) and as such, they extend well beyond the scope of the current analysis.

within an enumeration, indicated here by asterisks. Consequently, in these types of sequences only the first clause was retained.

- (52) a. But **LIKE** he's got so many things
that don't fall into the stereotype.
LIKE he's good at ah putting together cars,
* he's a carpenter,
* he's good with tools. (I/8/m/32)
- b. There's too many Bedfords in my life.
Ø I live on Bedford Street.
* I work at the Bedford Academy.
* I went to Bedford Public School. (N/O/m/24)

The second context consists of the first clause in response to a direct question, as in

(53).

- (53) a. Q Do you have any friends that are going to go in there as well or?
A Ø I have a few. (3/H/m/12)
- b. Q What happened to you during the blackout here in the Beach?
A Ø I was actually here working at the rec-centre. (N/fi/m/37)
- c. Q Really? And what else?
A **LIKE** one of my cats meows so much. (3/V/m/11)

As (53c) demonstrates, question-answer sequences do not categorically preclude the use of **LIKE**. However, examples such as this are rare; the discourse marker occurs with just 3 percent of this type of token ($N = 193$).³ This seems to relate to discourse markers more generally, which in these data are also used sporadically in this context: identical to **LIKE**, they occur with 3 percent of first clauses in response to a question. In fact, the only

³ Figures reported in this way refer to the total number of possible contexts, and not to the total occurrences of **LIKE**. Thus, of the 193 interrogative adjacency pairs that were extracted, **LIKE** occurred just 6 times, accounting for 3 percent of the data.

discourse marker that appears to surface in this context in any substantive way is well (see Schiffrin 1987), as in (54), which occurs with 12 percent of answer clauses.⁴

- (54) a. Q: Do you hear- read about those things in East-York?
 A: **WELL**, I think [it happens anywhere]. (N^o/f/80)
- b. Q: And you'd rather represent them than play the sport, you think?
 A: **WELL**, I would like to play the sport. (3/O/m/11)

The third context concerns relative clauses. **LIKE** categorically fails to occur with non-restricted relatives (N = 58), as in (55a), and it is used so infrequently with restricted relatives, as in (55b-c), that it appears to be nearly categorically blocked in these structures (1.6%; N = 180). However, note the position of **LIKE**; I will return to these forms in section 4.2.2.

- (55) a. A lot started in Denmark,
 [which is probably ten years ago]. (I/=/m/67)
- b. And I only have about two friends there
 [who **LIKE** I'm actually good friends with]. (3/F/f/17)
- c. There was this kid
 [who **LIKE**, I really don't like]. (2/u/m/15)

Due to the categorical failure of **LIKE** to surface within lists and in non-restricted relative clauses, as well as its highly restricted use in response to direct questions and in restricted relatives, all such contexts were excluded from the distributional and multivariate analyses.

⁴ Sacks, Schegloff and Jefferson (1974) report that well is frequent in turn-initial positions, and its functions in question/answer adjacency pairs have been documented in numerous works (e.g., Lakoff 1973b; Wootton 1981; Owen 1983; Schiffrin 1987). In Schiffrin (1987:103-105), who suggests that well is a response marker, this form introduces 14 percent of answers, a figure that corresponds to the Toronto results.

In the end, a total of 5737 CPs were retained for analysis. These were coded for a series of internal and external factors. I now turn to a discussion of this part of the methodology.

4.1.2 Coding and analysis

A number of factors, internal and external, were built into the coding scheme. Because this is the first analysis of discourse LIKE both to apply accountable quantificational methods and to subject the data to multivariate analysis, the factors considered throughout this dissertation test various hypotheses and claims that have been gleaned from the existing literature. I make no claim that these are the definitive factors in accounting for LIKE. Rather, they mark a first attempt to model LIKE in the variationist tradition, providing a foundation for future research.

The internal linguistic factors I consider for the clause-initial context are as follows:

- 1) the syntactic type of the clause (matrix, subordinate)
- 2) the status of the clause (bare, conjoined, marked by another discourse feature)
- 3) the position of the clause in the turn (initial, medial, follow-up to an answer)

First, I assume that all finite clauses are headed by a CP projection. If all else is equal and LIKE adjoins to the left periphery of CP as shown in (47), then it should be able to do so regardless of the level at which that projection occurs (i.e., matrix or subordinate). In order to test this hypothesis, the clause type was taken into account. Tokens were coded according to whether they were a main clause, as in (56a), or a subordinate clause. Within the latter category, both embedded clauses that form a complement to the main clause (56b) as well as adjuncts that can occur on either side of the matrix (56c-e) were included.

- (56) a. **LIKE** my uncle's sister married this guy, George-J. (N/‡/m/85)
- b. I just remember
[**LIKE** there were tons and tons of papers and garbage all over the floor]. (I/~ /f/29)
- c. [**LIKE** when it came down to it],
I think I was pretty much like a social reject. (I/ϕ/m/21)
- d. [**LIKE** if you drive up Elgin or Arnold or whatever],
there was just always these big monster homes. (I§/f/21)
- e. **LIKE** one of my cats meows so much,
['cause **LIKE** he's really picky and everything]. (3/V/m/11)

Second, Andersen (2001:284) argues that in clause-initial position, **LIKE** has a tendency to co-occur with other discourse markers and with conjunctions. He writes (2001:285):

We note that it is especially the connectives that tend to collocate with **LIKE** in clause-external contexts. It appears that the most common of these collocations, and like, [(be)cause] like, but like, and I mean like have achieved an almost formulaic status and seem to work as fixed or semi-fixed expressions.

To test this claim, the data were coded for whether the clause was conjoined (57a), introduced by another discourse marker (57b), or was bare (57c), by which I mean that neither conjunctions nor other discourse markers were present.

- (57) a. and **Ø** my other cat always sleeps,
and **LIKE** we almost never see him. (3/V/m/11)
- b. You-know, **LIKE** the people were very, very friendly.
You-know, **Ø** we'd sit out in the park and talk with different people. (N/V/f/60)
- c. **Ø** Nobody said a word.
LIKE my first experience with death was this Italian family. (N/U/f/82)

Third, the data were coded for the position of the clause in the turn. Since markers function to link utterances (Fraser 1988, 1990; Schiffrin 1987; Traugott 1997 [1995]), it is

feasible that the turn-initial position will disfavour LIKE, while turn-medial ones will exhibit higher proportions of use. Vincent and Sankoff (1992:212), for example, report that in Montreal French, pragmatic devices are far more frequent in elaborated genres such as analytic (argumentative, evaluative) and descriptive (narrative) discourse, genres that are characterized by the length of the turn (see also Erman 1987). This length translates into higher ratios of turn-medial clauses, and thus, greater opportunities to elaborate upon previous discourse. Further, while the answer to a direct question nearly categorically blocks LIKE (53), the second clause, where the speaker may clarify or elaborate upon their answer, seemed to contain a large proportion of this discourse marker. To test these hypotheses, clauses were coded for whether they were turn-initial (58a), turn-medial (i.e., any position other than initial) (58b), or a follow-up to a response to a direct question (58c).

- (58) a. Int Sounds like your mom was ahead of her time.
 FD Ø She definitely was, yeah. (N/fi/m/37)
- b. JM Ø There wasn't like an open space between us and downtown Toronto.
 Ø It was all urban.
 LIKE if you interviewed Granny-Wyatt
 Ø you would have a totally different perspective. (N/X/m/46)
- c. Int And you're still in touch?
 CF Yeah. Our parents are actually good friends.
 LIKE we sort-of lost touch for half of high-school
 probably 'cause I went to Upper-Canada-College. (N/O/m/24)

The external factors that I will consider throughout this dissertation are age and sex. As touched on in Chapter 1, previous analyses have tended to focus on younger speakers (e.g., Miller & Weinert 1995; Andersen 1997 et seq.; Siegel 2002; Hasund 2003; Sharifian & Malcolm 2003; Levey 2004; Tagliamonte 2005). In this chapter and the following ones, I

will examine the frequency and patterns of use of LIKE among speakers ranging from 10 to 87 years of age, and I will investigate what effect, if any, sex may have on the use of LIKE. The discussion of sex will focus on issues surrounding frequency. However, in Chapter 8 I will return to the question of sex and discuss it from a functional perspective, i.e., marker vs. particle (cf. Dubois 1992; Erman 1992).

4.2 RESULTS

The sentence-initial position is the canonical slot for discourse markers in English (Traugott 1997 [1995]), and in the majority of cases, the topmost CP is a matrix clause. In the data considered here, for example, 3363 of the 5737 CPs are main level clauses and all but 100 of these are initial.⁵ Consequently, I will focus first on matrix clauses, and will turn subsequently to subordinate ones. This approach reveals the generalization of LIKE across both levels of structure. However, LIKE was established with matrix CPs well before it began appearing with subordinate CPs. Moreover, the subordinate context is by no means straightforward, providing evidence that the discourse marking function of LIKE has begun to spread beyond CP to the left edge of TP. Thus, the clause-initial context is not the monolithic entity that it has previously been assumed to be. Accordingly, I will consider each of these contexts separately.

⁵ A total of 4895 tokens have the potential to occur initially (i.e., they are not obligatory clausal complements of the main proposition). Based on this figure, sentence-initial matrix clauses account for 67 percent of the data (N = 3263). However, this proportion does not reflect actual language use. In the initial extraction phase, too few subordinate contexts occurred (N = 1008) and it was necessary to return to the interview materials and specifically target subordinate structures for extraction. In other words, matrix clauses outnumber subordinate ones by approximately 3:1 in the natural language data.

4.2.1 Matrix clauses

4.2.1.1 Distributional analysis

The overall distribution of LIKE with matrix CPs is given in Table 4.1. When speakers of all ages are taken into account, LIKE occurs with 14 percent of all matrix clauses.

TABLE 4.1
Overall distribution of LIKE on matrix CPs

LIKE		Ø (no LIKE)		total
%	N	%	N	
14	478	86	2885	3363

On its own, this result is difficult to contextualize. There is no point of comparison, since it is unknown how frequent discourse markers are more generally. In these data, the proportion of matrix clauses that are marked by some other discourse form (e.g., well, I mean, so, you see, actually, etc.) is nearly equal to that of LIKE: 13 percent. In other words, LIKE is the single most frequent discourse marker overall, accounting for almost the same amount of the data as all other forms combined.

Does this hold across apparent-time? If we break the sample down according to age, as in Figure 4.1, two trends are noticeable. First, the proportion of discourse markers other than LIKE remains constant across the generations; there is very little fluctuation. The youngest speakers in the sample, the 10 to 12 year olds, use discourse markers at a rate comparable to older segments of the population. This seems to contradict Miller and Weinert's (1995:380) suggestion that discourse strategies are acquired relatively late.

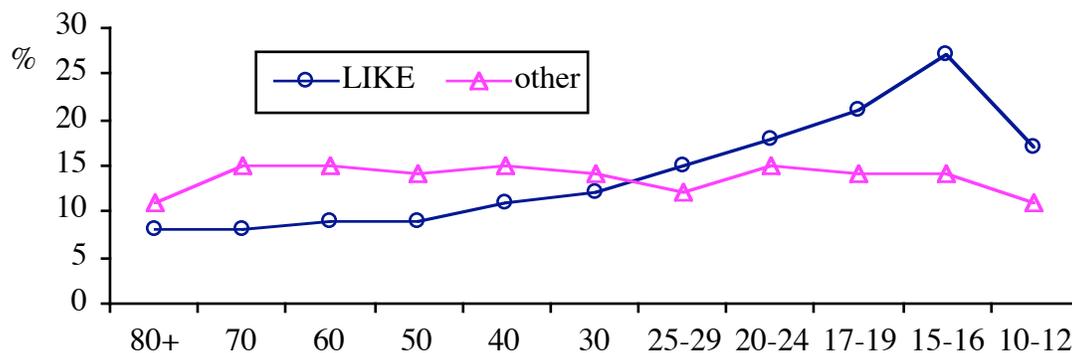


FIGURE 4.1 Distribution of LIKE and other discourse markers with matrix CPs across apparent-time

Second, and more important, is that these results clearly demonstrate the frequency of LIKE increasing in apparent-time. It occurs with 8 percent of matrix CPs among the octogenarians, but where its use peaks among the 15-16 year olds, it marks 27 percent of main clauses. Thus, over a period of approximately 65 years, the proportion of matrix CPs that are marked by LIKE has increased substantially.

A certain amount of caution is necessary. Dubois (1992:185) documents in real time an age-grading effect in the use of extension particles, with high rates associated with youth (i.e., 15-20 years of age). It is therefore possible that the results for LIKE among the younger speakers in Figure 4.1 may reflect in part a higher frequency of use that is characteristic of their age cohort. However, the results for other discourse markers mitigate against this interpretation. As noted, their overall rate of use is level in apparent-time: there is no evidence that the effect noted by Dubois for extension particles extends to discourse markers.

Nevertheless, Figure 4.1 also reveals that the distribution of LIKE relative to other forms has changed. Among the older speakers in the sample, LIKE is less frequent than other discourse markers, but this margin slowly narrows over apparent-time. Among the 30 year olds, the forms have almost leveled, and beginning with speakers in their late twenties, the trend has reversed: LIKE has become the dominant form for marking main clauses. In contrast to the results reported in Figure 2.1 (Chapter 2), this pattern does not indicate that one form (LIKE) is supplanting another (other markers). The trajectory in Chapter 2 revealed the complementarity of LIKE and other approximative adverbs: where LIKE increases, the use of the adverbials decreases (Sankoff & Thibault 1981:207). It is precisely this complementary distribution of forms across apparent-time that indicates lexical replacement. In Figure 4.1, there is no complementarity. The frequency of LIKE increases, but crucially, it does not do so at the expense of other discourse markers. In fact, they remain remarkably stable across the entire community. Thus, we have a situation in which the use of a single form is rising, while the use of others holds constant.

It is important to stress that as discussed in the last chapter, the discourse marking function of LIKE is in evidence among the oldest speakers in the sample, the 80-87 year olds. Further, its overall rate within this cohort is a mere 3 percent lower than the overall use of discourse markers in general. Consequently, the use of LIKE by this generation cannot be set aside; LIKE as a marker is an established and productive form in the grammars of the oldest members of the community.

At the same time, the ongoing increase in the frequency of LIKE, reflected in the monotonic distribution with age (Labov 2001:460), is strongly suggestive that this form is

developing. In instances of ongoing change, it is typical to find evidence of a sex effect. Generally this is manifested as a female preference for incoming forms (Labov 1990, 2001; Chambers 1991). As discussed in Chapter 1 (§1.3.1), however, there is no consensus with regard to LIKE and its correlation with sex.

If we consider the distribution of LIKE according to this external factor in the Toronto data, the marker is fairly evenly divided between the sexes (females 15%; N = 1756 vs. males 14%; N = 1609). Predictably, when viewed in apparent-time, as in Figure 4.2, no overall trend is evident.

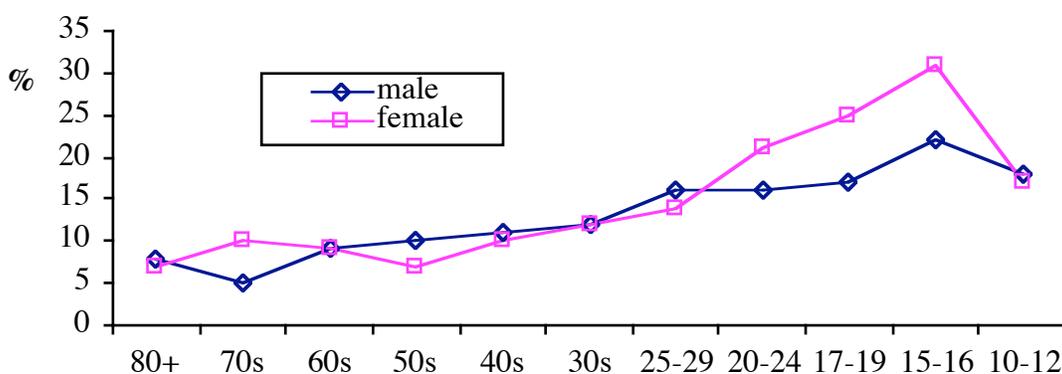


FIGURE 4.2 Cross-tabulation of LIKE with matrix CPs by age and sex

In fact, it is only among the speakers between the ages of 15 and 24 that a distinctive pattern emerges: here females are the leaders in the use of LIKE. This is an intriguing result, since these are also the age groups where LIKE is most frequent overall (see Figure 4.1). While the proportion of LIKE among males increases incrementally from one age cohort to the next younger one among the 15-24 year olds, it is the females in this age bracket who appear to be pushing up the frequency.

This result is reminiscent of recent findings for quotative be like. As the frequency of this form rises, a strong female association develops (Tagliamonte & Hudson 1999; Tagliamonte & D'Arcy 2004, under review). In other words, the more grammaticalized the form is, the greater the sex effect becomes. It is perhaps too early to tell if the same correlation is emerging with regard to the discourse marking function of LIKE, but the results in Figure 4.2 allow for this possibility. However, if such an effect is in fact developing, what would the implications be? Before addressing this question, I think it necessary to examine the use of LIKE in other contexts. This will provide a broader backdrop to address questions of sex, discourse features (both markers and particles), and ongoing change. Consequently, I will return to this topic in Chapter 8.

Labov (2001) has documented in detail that in cases of both stable sociolinguistic variation and ongoing change, adolescents between the ages of 13 and 16 have higher values than other age groups. In Figure 4.1, Labov's adolescent peak is clearly visible. The mechanism behind this peak is social, deriving from the logistic incrementation of change: adolescent girls are the leading recipients in the transmission of linguistic change (Labov 2001:444-445, 463). This pattern is visible in Figure 4.2, where it is the 15-16 year old girls who display by far the highest rate of LIKE.

In light of these general findings, the results for the youngest age group, the 10-12 year olds, require comment. First, the drop in the overall proportion of LIKE that is evident in both Figures 4.1 and 4.2 from its peak among 15-16 year olds is typical of apparent-trajectories for ongoing change (e.g., Labov 2001:170). Second, Figure 4.2 reveals little difference between males and females in this group (18% vs. 17% respectively), yet Figure

4.1 demonstrates that their overall frequency of use is virtually identical to that of the 20-24 year olds, where a female association with LIKE is apparent. How can this be reconciled with the possibility that where frequency of use is highest, a sex effect may develop? Recent findings involving 10-12 year olds indicate that social constraints are developmental, emerging during the high-school years (Tagliamonte & D'Arcy 2004), the same years when the peer group takes on an overarching role in influencing patterns of language use (Eckert 1988). The current results are explicable on the basis of this previous research; little to no differentiation between the sexes is visible among the 10-12 year olds because they have not reached the age where the language of their friends — and therefore the social factors attendant to linguistic form — becomes a primary influence. Notably, the results of the preadolescents mirror those of the speakers in their late twenties, as well as those in their thirties and forties. These adult groups are the peers of the parents of the 10-12 year olds. This too suggests that the youngest speakers are not yet fully adopting the language patterns of their social group, continuing instead to model the language of their parents.

At this point I would like to turn to the internal factors to see how they constrain the discourse marking use of LIKE on main clause CPs. Table 4.2 displays the distribution of LIKE according to the position of the clause in the turn.

TABLE 4.2
Distribution of LIKE on matrix clauses according to position in turn

Initial		Medial		Follow-up		total
%	N	%	N	%	N	
8	279	14	2959	22	125	3363

As hypothesized, LIKE is least frequent in turn-initial position, and most frequent on a matrix clause that follows-up the response to a direct question. The discourse-medial position falls in between. Cross-tabulation reveals that this hierarchy is constant across both age and sex, suggesting that the results in Table 4.2 reflect a bona fide internal pattern in the use of LIKE with matrix CPs; it is constant across the entire community. This is an important observation because although frequencies of use can vary, grammatical constraints are expected to remain stable regardless of “extra-linguistic circumstances” (Poplack & Tagliamonte 2001:92).

Another possible factor affecting the discourse marking function of LIKE concerns the types of elements with which it co-occurs. Discussed earlier, based on the text frequency of certain collocations, Andersen (2001:284-285) posits that LIKE tends to co-occur with other pragmatic markers and conjunctions. He suggests that the most common of these — and like, (be)cause like, but like, and I mean like — “have achieved an almost formulaic status and seem to work as fixed or semi-fixed expressions” (2001:285; see also page 297).⁶ In light of grammaticalization, this is important because it suggests that through frequent use, these collocations are becoming ritualized (see, for example, Haiman 1994; Boyland 1996; Bybee 2003; see also Bybee & Hopper 2001 and papers therein). Thus, whether LIKE in fact occurs in habitual combinations may help shed light on the path by which the discourse marker is developing. Is the routinization of these structures behind its rise in frequency?

Table 4.3 displays the results according to the status of the matrix clause as bare, conjoined, or marked by another discourse marker.

⁶ The collocation (be)cause like, which introduces a subordinate clause, will be discussed in section 4.2.2.

TABLE 4.3
Distribution of LIKE on matrix clauses according to the status of the CP

Marked		Conjoined		Bare		total
%	N	%	N	%	N	
7	448	7	1053	20	1862	3363

These results indicate that the most frequent context for LIKE is with a ‘bare’ matrix CP, one that is not conjoined or marked by another discourse marker. Indeed, the collocations are relatively infrequent in comparison. The tendency predicted by Andersen (2001) is not apparent; the accountable quantitative measure does not corroborate the text frequency results.

Further, based on text frequency, and like is the most common collocation in the COLT materials, occurring 84 times, while but like is far less frequent, occurring only 30 times (Andersen 2001:285). In other words, the ratio is approximately 3:1. Following this same method, the ratio of and like to but like in the Toronto data is also 3:1. However, when we calculate the frequency of each collocation proportionally, and like and but like cease to be differentiated, accounting for 7 percent (N = 751) and 6 percent (N = 297) respectively of the total occurrences of each conjunction. This highlights the importance of accountable methods in the analysis of discourse phenomena. Put plainly, even though the text frequency of and is more than two and a half times greater than that of but, there is no difference in the distribution of LIKE with these conjunctions. This, in concert with the low rate of occurrence of the discourse marker in these contexts, suggests that and like and but like do not in fact represent cases of “incipient fixation” as proposed by Andersen

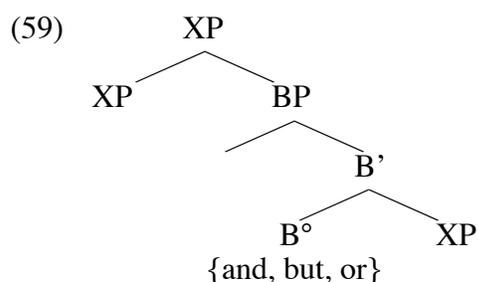
(2001:297). In short, the increasing use of LIKE exposed by Figure 4.1 is not directly attributable to these forms.

What of LIKE and other discourse markers? Table 4.3 displays the overall results for discourse markers as a whole. It does not indicate whether differences exist within the class of markers. Although there are a number of forms that are infrequent (e.g., in fact, actually, you see, I don't know, etc.), there are others that occur with greater regularity. These are forms such as well, I mean, you know, and so. This last form accounts for more than half of the discourse markers in the data (N = 250), not including LIKE.

Of these, the two that occur most often with LIKE are I mean, where LIKE has an overall rate of 13 percent (N = 31), and you know, where the proportion of LIKE is 28 percent (N = 43). In the case of so, LIKE occurs with just 4 percent of tokens. Thus, there is a wide range in the rates at which LIKE collocates with individual discourse markers on the periphery of a clause. The relatively low numbers of I mean and you know tokens make it difficult to generalize from these findings, but there are indications that despite their apparently frequent co-occurrence with LIKE, these chunks are not becoming routinized. Specifically, LIKE does not have a fixed position: LIKE you know occurs nearly as often as you know LIKE does, and LIKE I mean is equally as frequent as is I mean LIKE. If I mean LIKE were developing as a fixed or semi-fixed expression as Andersen (2001:285) suggests, we would not expect to find this high degree of variability in the syntagmatic ordering of the individual forms.

Of course, as discourse markers, both LIKE and I mean are adjuncts, and as such their linear position vis-à-vis each other is not fixed: adjunction can be iterative, but there is no

stipulation that it must follow a given hierarchy (cf. adverb interpolation, which is constrained by the fixed order of functional heads; see Cinque 1999, 2004).⁷ Thus, variability in the order of LIKE and other discourse markers is predictable from structural considerations. In contrast, from a syntactic perspective, the ordering of elements in collocations involving conjunctions and LIKE is predictably invariant: in coordinate structures, the conjunction is situated above the conjoined projection. This is demonstrated, following Munn (1993:13), in (59):



In the case of coordinated clauses, the CP to which LIKE adjoins is the sister of B° (the lower XP in (59)). That LIKE follows the conjunction thus follows directly from the structure. The ordering and + LIKE and but + LIKE does not therefore reflect the fixing of elements in a phrase as a single processing chunk (i.e., a formulaic unit), but reflects instead the hierarchy of projections in the syntax.

I would like to step back for a moment though and consider the results in Table 4.3 from a different perspective. Overall, the presence of another discourse marker on the left periphery of a matrix CP does not appear to foster use of LIKE. Is this apparent co-

⁷ See, for example, Belletti's (2001) proposal to iterate the projections of the complementizer system (e.g., Topic, Focus, etc.). The syntactic adjunct slot proposed by Traugott (1997 [1995]) in (47) corresponds to Topic in Kiparsky (1995).

occurrence restriction characteristic of LIKE, or is it typical of discourse markers more generally? In fact, discourse markers do not typically collocate together. In these materials, excluding LIKE, there are only 15 instances of co-occurring markers, an overall distribution of just 3 percent.⁸ Thus, there seems to be a general constraint against stacking discourse markers, and the results for LIKE can be seen to follow from this.⁹ At the same time, the frequency of LIKE with other markers is more than double the rate of collocation for markers apart from LIKE: 7 percent. Thus, in its greater ability to collocate with these forms, LIKE is distinguished from other discourse markers, which nearly categorically fail to co-occur. In this regard, the quantitative results corroborate those of Andersen's (2001) text frequency account.

In sum, the distributional analysis has revealed that the frequency of LIKE appears to be increasing in apparent-time, but that this change does not seem to be affected by speaker sex. A caveat to this generalization is suggested by the results for the 15-24 years olds, where a female correlation may be developing. The overall results also suggest that language internal factors are operative: LIKE is least frequent in turn-initial positions and most frequent with the second clause of a turn in response to a direct question. It is also used with greater regularity on matrix clauses that are not conjoined or marked by some other discourse form, suggesting that the status of the CP is an important factor in the use of LIKE. What is the role of each of these factors when they are considered simultaneously?

⁸ Interestingly, most of these involve so (e.g., so anyway, so you know, so I mean, so yeah), though I mean and you know co-occur as well.

⁹ The source of this effect is likely pragmatic but it awaits fuller exploration.

Are the results revealed by the proportions significant? Are some factors stronger than others?

4.2.1.2 Multivariate analysis

Table 4.4 displays the results of a multivariate analysis of both the external (sex, age) and internal (position, status) factors on the probability that LIKE will occur with matrix CPs. Due to the gradual slope of change revealed in Figure 4.1 and low distribution of LIKE overall in matrix CP context across the sample (14%; N = 3363), the speakers have been reorganized in order to allow for a better model of variability. These groupings categorize the speakers according to four life stages: adolescence (10-16), young adulthood (17-24), middle age (30-59), and golden years (over 60). Further, the appearance of conjunctions and other discourse markers on the periphery of the matrix CP has been collapsed as a single factor here because the proportional results in Table 4.3 indicate that the frequency of LIKE is identical in each of these contexts.

A number of observations can be made. First, the results of the multivariate analysis corroborate those of the proportional analysis. Each of the hierarchies suggested by the overall distributions is reflected here. Concerning status, bare CPs are strongly favoured, while those marked by other discourse forms and conjunctions are strongly disfavoured. With regard to position in the turn, the follow-up position is the strongest position, followed by other turn-medial clauses; turn-initial matrix CPs are disfavoured.

TABLE 4.4
Contribution of external and internal factors on the probability of LIKE on matrix CPs

INPUT factors considered	FW	.12 %	N
Status			
bare	.64	20	1862
conjoined/marked	.32	7	1501
range	32		
Position			
follow-up	.59	22	125
medial	.51	14	2959
initial	.34	8	279
range	25		
Age			
10-16	.65	22	585
17-29	.58	18	1056
30-59	.43	11	957
60+	.36	8	765
range	29		
Sex			
female	.52	15	1756
male	.47	14	1607
range	5		
TOTAL N			3363

Second, sex is significant. In light of the lack of a visible trend in Figure 4.2 and the negligible proportional difference between the sexes, this result is of considerable interest. Closer examination of the interaction of age and sex reveals that the fine distinctions between the age groups in Figure 4.2 obscure a greater generalization, which is that, with the exception of the middle aged speakers, the frequency of LIKE is higher overall among females than it is among males. This is shown in Figure 4.3, and it explains why this social correlation is selected by the step-wise procedure. Thus, despite the marginality of the sex effect across the sample, its consistency suggests that it is a fundamental factor in the use of

LIKE as a discourse marker in matrix CP contexts. As we investigate the use of LIKE in other contexts throughout this dissertation, the significance of this result will become clear.

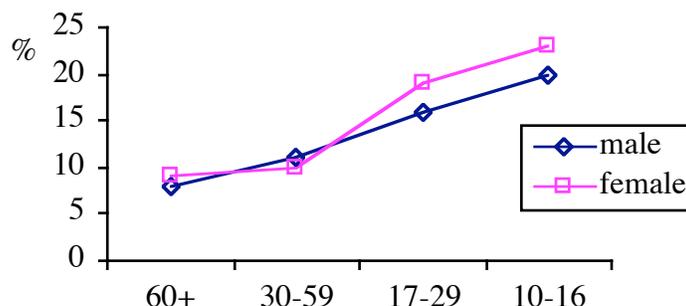


FIGURE 4.3 The use of LIKE in four age groups according to sex: matrix CPs

Third, the ranges in Table 4.4 — which indicate the magnitude of effects — indicate that each of the three remaining factors, status, position, and age, exerts a strong constraining effect on the use of LIKE. However, the strongest of the three is the status of the clause (note the range of 32). Thus, not only is LIKE significantly disfavoured with conjunctions and other discourse markers, but of the factors considered here, this co-occurrence constraint is the most important one conditioning the use of this marker. Based on Andersen (2001), who argues for the lexicalization of conjunction/discourse marker + LIKE, this result is unexpected. If these collocations were in fact one of the mechanisms behind the development of LIKE, then such a strong constraint against them is not predicted.

Fourth, the results for age support the finding that LIKE is slowly gaining momentum over apparent-time. The factor weight for one group is always slightly higher than that of the next older group, demonstrating the characteristic monotonic association between age and use that is characteristic of ongoing change (Labov 2001:460). Crucially, Figure 4.4 reveals that both the co-occurrence constraint (status) and the positional constraint operate

uniformly across the sample. In other words, as the frequency of LIKE increases, the direction of effects seen in Table 4.4 holds constant. Poplack and Tagliamonte (2001:92-93) contend that the conditioning of variability is immutable in the face of extra-linguistic factors, operating distinct from fluctuations in overall rates of occurrence. As such, they stipulate that it is the hierarchy of constraints, and not proportional results or even significance, which determines the variable grammar. The results in Figure 4.4 thus demonstrate that a single grammar exists for the discourse marking function of LIKE with matrix CPs in the Toronto speech community, and that it operates regardless of speaker age.

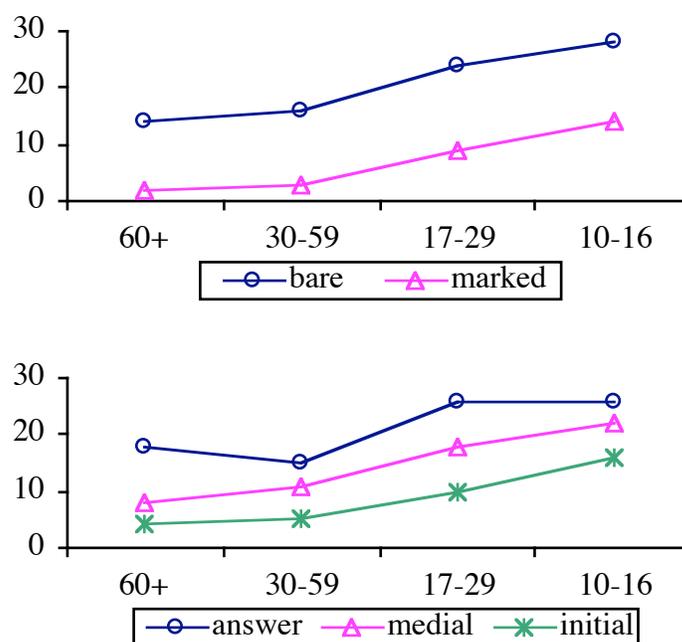


FIGURE 4.4 Cross-tabulation of LIKE on matrix CPs: internal factors and age

In addition to unveiling the shared configuration of effects, Figure 4.4 also provides a dramatic display of the Constant Rate Effect, CRE (Kroch 1989a, 2001). The generalization underlying the CRE is that during change, disfavoured contexts acquire new forms no later

than do favouring ones (Kroch 1989a:238). As a consequence, the constraints governing variability remain constant across time. Although this model was intended to capture the parallelism across contexts of syntactic change deriving from parametric variation (i.e., ‘grammar competition’; see Kroch 1989a, 1989b), the results in Figure 4.4 reveal its applicability in the rise of a pragmatic device (see also Cukor-Avila 2002 for evidence of the CRE with quotative be like). Indeed, the patterns in Figure 4.4 provide a striking display of regular and systematic change across the entire Toronto speech community.

4.2.1.3 Summary

The analysis of matrix CPs has shown that LIKE functions as a discourse marker in the speech of all age groups, and the apparent-time evidence, both distributional and multivariate, indicates that use of this feature continues to increase. Further, the lack of interaction among constraints demonstrates stable conditions on variability across the community, constraints that were already in place among the oldest generation of LIKE users: the 80-87 year olds. Thus, at least in so far as the discourse marking function of LIKE is concerned, it appears to represent bona fide linguistic change. Further, this analysis has shown that a strong co-occurrence restriction is in place which significantly disfavours the collocation of LIKE with either other discourse markers or conjunctions.

The question now is, what of subordinate clauses? It was hypothesized above that because both matrix and subordinate clauses project a CP layer, both provide possible adjunction sites for LIKE. I turn to this possibility now.

4.2.2 Subordinate clauses

A total of 2374 subordinate clauses were extracted from the interview materials and retained for analysis. As discussed earlier, these comprise a number of different types of clauses, including embedded complement clauses (56b) and adjuncts which provide background or context to the main proposition (e.g., because, if, when, etc.) and which are positionally free to either precede or follow the matrix clause (56c-e). For reasons that will soon be clear, a third type of clause will be considered here; these are then clauses, consisting of both apodoses of conditional if clauses and temporally ordered adverbial clauses. These are matrix-level structures, yet the slot occupied by LIKE forces them to be considered separately.

Indeed, the subordinate context is complex, distinguished on the one hand by structures in which LIKE precedes the clause, adjoining, as in matrix clauses, to the left periphery of CP, as in (60), and on the other hand by structures in which LIKE occurs to the left of the subject, as in (61). In other words, there are two distinct patterns.

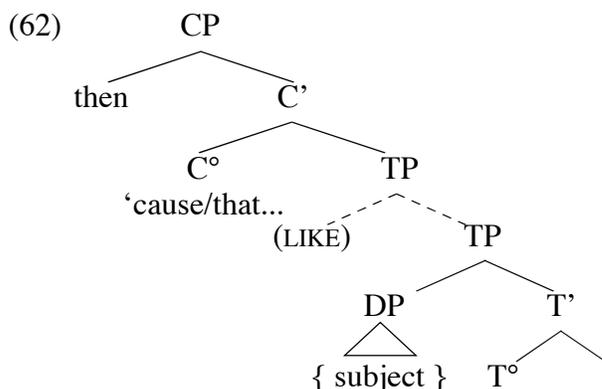
- (60) a. **LIKE** if you go up here to Saint-Francesco's it is the last bastion. (N/D/m/52)
 b. **LIKE** when I first heard that, I was still teaching. (N/T/m/64)
 c. There's still some, **LIKE** if you go up on Kipling. (I/™/m/40)
 d. So I get it all done **LIKE** when I get home. (3/F/f/17)
- (61) a. It's weird because **LIKE** you didn't really fit in the Black group. (I/≠/f/21)
 b. I think that **LIKE** there's been a desire instilled in me. (I/§/f/21)
 c. If they're like cut or something, then **LIKE** they have a teacher come in. (2/e/f/11)

In (60a,b), LIKE occurs on the left edge of the sentence. In both examples, the initial clause is subordinate to the main proposition. This is not proof that both matrix and subordinate CPs may host discourse markers, since LIKE continues to target the topmost CP

in the structure. That is, it is sentence-initial. In contrast, the examples in (60c,d) do provide evidence of this: here the subordinate clause is the lower clause. Thus, regardless of the level of the CP, LIKE may adjoin to the left edge of this projection, occupying the slot discussed in Traugott (1997 [1995]).

However, a different adjunction site is evidenced in (61). In these constructions, LIKE is not on the left periphery of the entire clause. Instead, in (61a,b) it intervenes between the head of CP (because, that) and the subject of the clause. This is particularly noteworthy in the case of because, which is analogous to if and when (also after, while, until, and whenever) both structurally and semantically: all are hosted in C°, all head clauses that may occur in a range of positions, and all modify the main proposition. In the instance of then clauses, as in (61c), the situation is different only in that I assume then to be situated in SPEC CP rather than C°. It should be noted that in all constructions like those exemplified in (61), the position of LIKE is fixed: when used in causal clauses, embedded clauses, and clauses headed by then, it categorically occurs to the immediate left of the subject (N = 1284).

The question is, what is this position in structural terms? I assume here that it is the left edge of TP. First, as an adjunct, LIKE targets maximal projections. Second, following Pollock (1989), I assume that the subject is hosted in the SPEC TP, as in (62):



Since LIKE surfaces between this position and the higher CP projection, and because I assume that it targets the phrasal level, I posit that in constructions like those in (61), LIKE adjoins to the left periphery of TP, analogous to its position on the left edge of CP in (47).¹⁰

Further support for the TP analysis comes from instances when the complementizer is absent, as in (63).

- (63) a. We decided [\emptyset we needed to be more centrally located]. (N/A/f/56)
 b. I'm just saying [**LIKE** it would be so sick to live there.] (3/S/m/18)
 c. I thought [**LIKE** University-of-Toronto is big.] (3/K/f/18)

¹⁰ The proposal that LIKE adjoins to TP in (61) stems from the structure in (47), which I adopt for ease of exposition. Nothing hinges on this cartography. Indeed, the analysis presented here is rather straightforwardly translatable into the 'split Comp' framework proposed by Rizzi (1997, 2001, 2002), where four separate projections are argued for (ForceP, InterrogativeP, FocusP, FiniteP), with possibly recursive TopPs interspersed. The position of LIKE in relation to the types of elements that may occur on the left periphery of the clause falls out quite naturally from this architecture: because and that are hosted in Force, the highest projection, while if and when are hosted lower down, in Int. Thus, if we assume that LIKE targets some position between Force and Int, then the two patterns observed in (60) and (61) are explicable on structural grounds. The missing link to this analysis is the position of then. Rizzi (pc, June 10 2005) posits that then seems to be situated quite high (i.e., above FinP and FocP), and by analogy with French puis, which may function as a coordination, suggests that English then may be closely connected with coordinating conjunctions. This would situate it very near the top of the Comp domain. The crucial point for the current analysis is that as long as then is hosted above IntP, then the syntagmatic position of LIKE is also accounted for structurally. This remains, however, an open issue.

Following Boskovic (1994), I assume that these types of structures do not project a CP. Evidence for this comes from both coordination (64) and preposing (65), since both (64c) and (65d) are ungrammatical.

- (64) a. Chris believes [_βthat he's nice] and [_βthat he's handsome].
 b. Chris believes [_□he's nice] and [_□he's handsome].
 c. *Chris believes [_□he's nice] and [_βthat he's handsome].
- (65) a. Pat thinks [money can buy happiness].
 b. Money, Pat thinks, [can buy happiness].
 c. Pat thinks [that money can buy happiness].
 d. *Money, Pat thinks, [that can buy happiness].

In spite of lacking a CP, the data in (63) demonstrate these types of complement clauses are nonetheless able to host LIKE. If LIKE targets the X^{\max} level of structure, then it must adjoin in this instance to TP since CP is not available.

Consequently, the subordinate context is divided into what I will refer to CP SUBORDINATES and TP SUBORDINATES. CP subordinates consist of constructions like those in (60), where LIKE occurs on the left edge of CP. TP subordinates consist of constructions like those in (61) and (63), where LIKE adjoins to the left edge of TP. In other words, subordinate clauses provide us with two distinct functional projections for the analysis of LIKE: CP and TP. Consequently, I will consider each context separately. Moreover, the discussion is restricted to distributional factors; no multivariate evidence is presented because it is unclear how the factors examined for the CP matrix context relate to subordinate contexts. For example, subordinate clauses are often lower level structures. Therefore, the position of the clause in the turn will, more often than not, be medial. Further, these clauses are rarely conjoined, rendering the status of the clause moot.

4.2.2.1 CP subordinate clauses

The overall distribution of LIKE with CP subordinates is 14 percent (N = 1090). This is identical to the result for matrix CPs. However, when we observe the frequency of LIKE in the CP subordinate context over apparent-time, as in Figure 4.5, the CP contexts differ in one critical respect. Unlike matrix CPs, which provide a variable context for all age groups, LIKE first appears with subordinate CPs among the 60 year olds. Among speakers in their seventies and eighties, LIKE never occurs (N = 133).¹¹ This suggests that the use of LIKE to mark subordinate clauses developed later and did not coincide with the emergence of the discourse marker in matrix clause CPs.

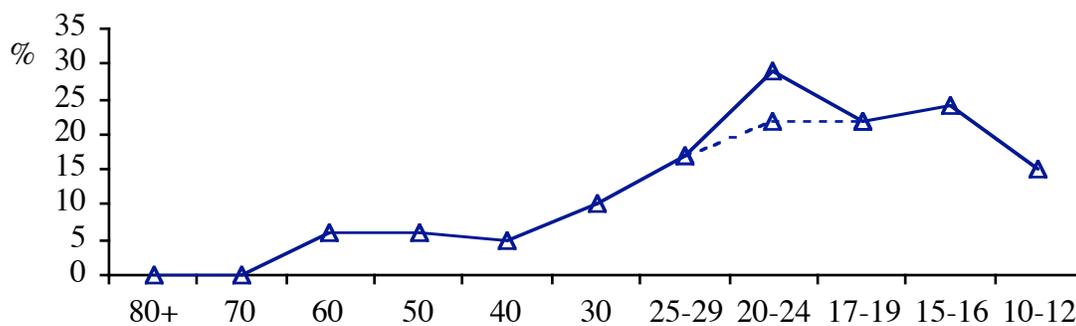


FIGURE 4.5 Distribution of LIKE on CP subordinates across apparent-time
(dotted line represents results when the anomalous 22 year old is omitted)

This path is not predicted by structural factors, since both levels of structure project a CP. Moreover, among the two oldest age groups nearly three-quarters of CP subordinates occur in sentence-initial position. If LIKE simply targets the left edge of a sentence, there is no syntactic explanation for its failure to mark those whose highest clause happens to be

¹¹ Data sampling is not responsible for the categorical lack of LIKE within the two oldest age groups; there are simply no tokens of LIKE with subordinate clauses for these speakers.

subordinate rather than main. The explanation may be semantic. That is, it is possible that in the initial stages of grammaticalization, discourse markers adjoin strictly to the main proposition, linking it to prior discourse, and only later spread to subordinate clauses through analogy. This is an interesting question, but it is beyond the scope of the current research; I leave the issue open.

There is something else going on here. The apparent-time results reveal that, just as we saw in the CP matrix context in Figure 4.1, the frequency of LIKE is increasing. Similarly, the adolescent peak is again in evidence, though this time it is overshadowed by the one that appears among 20-24 year olds. Examination of the individuals who comprise this age group reveals that the high rate of LIKE derives from a single individual, a 22 year-old male who uses this particle at an astonishing rate of 60 percent (N = 15). Once his data are removed, the anomalous 20-24 year old peak levels out.

None of these observations, however, help to account for the identical distributions of LIKE in both the matrix and subordinate CP contexts overall. The adjunction of LIKE to subordinate clauses was clearly a later development, yet among those for whom subordinate clauses are a variable context, the rate of occurrence matches that of the older context. Put plainly, the use of LIKE in subordinate contexts starts later, but rises faster. How can we account for this? Or, stated another way, do we need to?

Let us return to the CRE (Kroch 1989a, 2001). This model does not predict that the frequency of increase will hold at a constant rate for the duration of a change. Indeed, such a trajectory would directly contradict the S-shape curve of linguistic change (Weinreich, Labov & Herzog 1968; Bailey 1973). As summarized by Pintzuk (2003:514), “what is

‘constant’ in the Constant Rate Effect is that the change is the same across linguistic contexts, so that the frequency of the new form changes in the same way in all contexts.” In other words, all the contexts change together. How does this correspond to the use of LIKE in matrix CPs on the one hand, and subordinate CPs on the other? It does not. First, it is not a matter of one, matrix CPs, being favoured over the other, subordinate CPs. Second, this relates to the nature of the change. The spread of LIKE to CP subordinates does not reflect a change in an underlying parameter. Not only are discourse markers argued here to be adjuncts, but further, were this a matter of two distinct syntactic options (e.g., V to I movement in the rise of ‘do’ support), LIKE would appear in both contexts simultaneously (Kroch 1989a:238). The apparent-time results in Figures 4.1 and 4.7 clearly demonstrate that this was not the case. Rather, the emergence of LIKE in the CP subordinate context among 60 year olds reflects the generalization of a new form, LIKE, to a context where it could not previously occur (Hopper & Traugott 2003:100-103; Heine 2003:579-580). We are dealing with grammaticalization, not a competition between distinct grammars. Thus, the faster rate of increase by LIKE in the CP subordinate context may simply derive from analogical processes. As discussed above, there is no structural cause for a distinction between CPs based on their level within the syntax. Once whatever semantic or pragmatic motivations that restricted LIKE to matrix clauses were overcome, it is entirely plausible that subordinate CP projections were simply assimilated into the variable context, leaving no reason to differentiate between the two.

Examination of LIKE according to the social correlate of speaker sex may be helpful in this matter. For example, if the same language external pattern is evident for the CP matrix

and subordinate contexts, then this could be interpreted as corroboration that the levels of structure have been assimilated as a single context, CP. The first piece of evidence is the overall distribution across the sexes: just as with matrix CPs, LIKE is more frequent among females (17%; N = 493) than it is among males (15%; N = 464). The second piece of evidence is provided by the distribution of LIKE across apparent-time. Recall from section 4.2.1.2 that when the sample is re-grouped into four categories, a female association with LIKE is evident in every cohort except the middle-aged speakers. The results in Figure 4.6 are virtually identical to those in Figure 4.3.

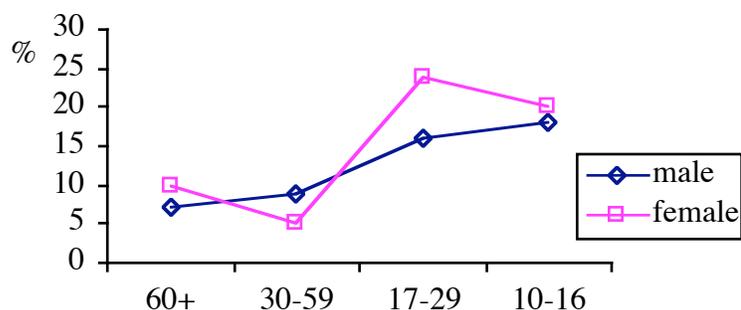


FIGURE 4.6 The use of LIKE in four age groups according to sex: subordinate CPs

Thus, in both the matrix and subordinate contexts, it is the females who are leading the move toward the use of LIKE on the periphery of CP.

4.2.2.2 TP subordinate clauses

We have seen that the CP subordinate context represents a later stage in the development of the discourse marking function of LIKE. Such an expansion is representative of generalization, whereby LIKE spreads to a context where it did not formerly appear. Initially restricted to matrix CPs, it comes to be used on subordinate CPs as well. It is this type of

extension that is typically attributed to a rise in the text frequency of a grammaticalizing form (Hopper & Traugott 2003; Bybee 2003). The generalization of LIKE to the left edge of TP, exemplified in (61) and (63) above, provides another example of this type of change, demonstrating how this marker gradually comes to be used in a broader range of contexts as it develops. Unlike the generalization from matrix to subordinate CP, however, the spread of LIKE to TP marks a broadening of the functional category it selects for adjunction. The former is shift across the same functional projection, CP > CP, differentiated only by the status of the clause as containing the main proposition or not. The generalization of LIKE to TP marks the first time here that we see this marker appear in a position other than on the left periphery of CP, the adjunct slot discussed by Traugott (1997 [1995]). The implication is that adjunction to TP developed after the establishment of the CP context, and that the ability of these two maximal projections to host LIKE did not evolve contemporaneously. This implication is corroborated by the apparent-time evidence.

Figure 4.7 charts the distribution of LIKE in the TP context across apparent-time (N = 1284). These results graphically demonstrate how LIKE is slowly spreading in the grammar. Among the four oldest groups in the sample, the 50 to 87 year olds, this form never occurs in the TP subordinate context (N = 396). It is not until the 40 year olds that LIKE is first attested on the left periphery of TP (1%; N = 122). Figure 4.7 also reveals that once LIKE spreads to this position, the TP is established as a possible adjunction site: from the time it appears in this context, the frequency of LIKE continues to increase. The same result is also evident in Figures 4.1 and 4.7. This pattern, where a form enters a domain and then

continues to rise in frequency is not the behaviour of a random, *ad hoc* linguistic feature. Rather, it is indicative of systematic, rule-governed change.

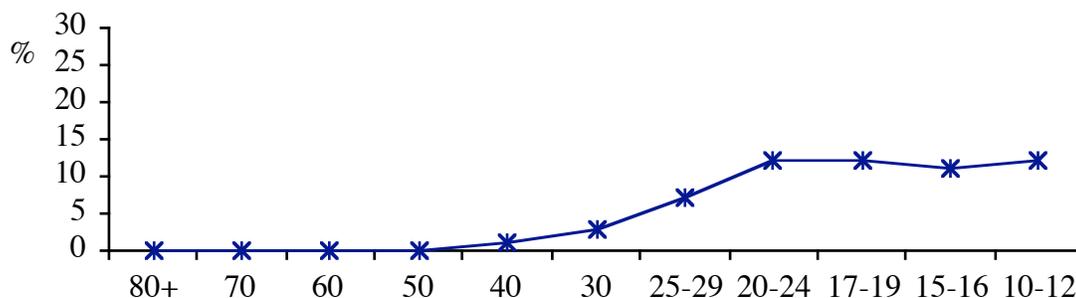


FIGURE 4.7 Distribution of LIKE on TP subordinates across apparent-time

It is this point that brings us back to relative clauses. As discussed earlier, LIKE nearly categorically fails to be used in this context: it never occurs with non-restricted relatives (N = 58) and has a rate of less than 2 percent on restricted relatives (N = 180) (see (55b,c)). When LIKE does appear in this context, it surfaces between the relative pronoun and the subject. Thus, these types of clauses are analogous to the ones in (61), where, despite the availability of a CP, LIKE adjoins to TP. The apparent-time results clearly demonstrate that the generalization of LIKE to this functional projection is a recent development. This, combined with the low accumulation of relative tokens, may have obscured our view of how this context actually behaves with respect to receptivity to LIKE.

An interesting aspect of the TP context is that it includes causal clauses, where because is hosted in C°. These are the final collocation pattern discussed by Andersen (2001:284-286). Table 4.5 displays the results for the individual structures that make up the TP subordinate context. I have included only the data for those age groups where LIKE is attested (i.e., the 10-49 year olds; Figure 4.7).

TABLE 4.5
Comparison of results for LIKE within the TP subordinate context

	%	total
complement	9	461
‘because’	7	294
‘then’	5	133
total	8	888

The overall distribution of LIKE is 8 percent (N = 888). Represented by 294 tokens, because clauses have an overall rate of LIKE of 7 percent. In light of the context as a whole, this suggests that because LIKE is fairly robust, but does it indicate that this collocation is becoming routinized? One way to disentangle this type of developmental path is to consider the distribution of because LIKE across the sample. Beginning with the 25-29 year olds, the rate of this collocation is level, ranging between 8 and 10 percent. Crucially, this stabilization occurs while the frequency of LIKE is still increasing (cf. Figure 4.7). This suggests that, akin to and/but LIKE and I mean LIKE, because LIKE does not represent incipient fixation (Andersen 2001:297), nor is this collocation responsible for increases in the use of LIKE overall.

4.3 DISCUSSION

As discussed in Chapter 3, the fact that LIKE is currently used productively with matrix CPs by octogenarians indicates that the discourse marking function of this form was already firmly embedded in the English of Toronto in the late 1930s. Based on this, it was hypothesized that this function began to develop near the turn of the twentieth century. The results in Figure 4.1 support this analysis.

In contrast, the apparent-time results indicate that the use of LIKE in subordinate contexts arose much later. It is first attested with subordinate CPs among speakers in their sixties, dating the generalization beyond matrix CPs around 1960. Interestingly, this is the same period when its roots have been ascribed to certain American counterculture groups (see Andersen 2001:216, and references therein). Here, then, we may have an explanation for this attribution. As an incoming change, the discourse marking use of LIKE had been increasing slowly. This is both typical (e.g., Weinreich, Labov & Herzog 1968; Bailey 1973) and predicted by the incrementation model in Labov (2001). However, the acceleration of a change “logically begins when the incipient change is attached to or is associated with a particular style or social group” (Labov 2001:462). Moreover, in order for a change to progress, such association is necessary (*ibidem*). Consequently, I would like to suggest that the rise of LIKE, as least in its discourse marking function, derives from its association with these counterculture groups. In this view, these groups are not the source of the “innovative American pattern,” as indeed, I have argued they cannot be, an argument that is bolstered by the apparent-time results. Instead, they represent the social vehicle for the further development of LIKE.

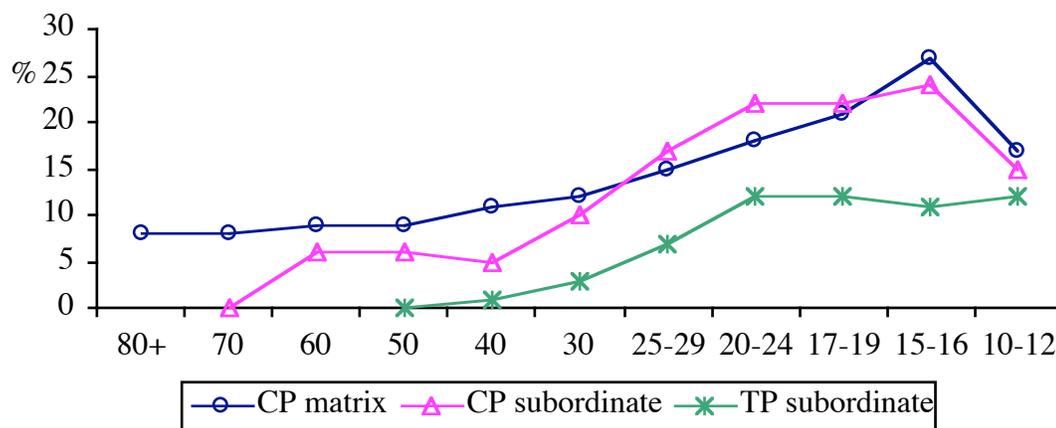


FIGURE 4.8 Distribution of LIKE in the clause-initial context across apparent-time

Displayed in Figure 4.8, with the 60 year olds the generalization of LIKE to subordinate CPs increases the overall text frequency of this marker in discourse. Thus, beginning in the 1960s, the time when these speakers would have been adolescents, the number of possible contexts for this marker has risen, and consequently, its use rises in tandem. It is possible that at this point, the frequency of LIKE passes some threshold, increasing in saliency so that this formerly incipient feature of the vernacular comes to be associated with the counterculture groups of the time. This is an arbitrary event (Labov 2001:462). But, it is also an event that can be contextualized in the broader social climate of the time. Consider Labov's NONCONFORMITY PRINCIPLE (2001:516):

Ongoing linguistic changes are emblematic of nonconformity to established social norms of appropriate behavior [...].

As a feature of discourse, LIKE is first interpreted along the formal/informal dimension. This polarity is then extended to the social dimension of “good” and “bad” (Labov 2001:512). In the United States in particular, the 1960s were a time of extreme turmoil, and many of the

counterculture groups of the time enacted the overt manifestation of nonconformity to accepted social norms. Thus, it is possible that, in the case of LIKE, the Nonconformity Principle can be interpreted literally: its use within these groups came to be associated with widespread nonconformity, allowing for the acceleration of the marker.

The most recent step in this evolutionary path is the generalization of LIKE from CP to TP. The apparent-time results suggest that this occurred no more than 25 years ago. In fact, the only speaker in the 40-49 year old cell to use LIKE in this context was 40 years of age at the time of the interview. Thus, the rise of TP as a possible adjunction site for LIKE may be even more recent than the apparent-time results first implied.

The question at this point is, does the use of LIKE in the TP subordinate context constitute the discourse marking function? Two pieces of evidence bear on this discussion. One, the TP context remains a clause-initial position in the sense that LIKE occurs to the left of the subject and before the full proposition. Consequently, the definition of “clause-initial” must be broadened to include the left edge of TP in addition to that of CP. Two, even in this position, LIKE continues to take wide scope over the discourse, linking the previous utterance, in this case a clause, with the current utterance. It is precisely this function that defines a discourse marker (Fraser 1988, 1990; Schriffrin 1987; Traugott 1997 [1995]).

Taken as a whole, the clause-initial context (CP and TP) provides important evidence regarding the nature of LIKE. The consistent finding that the frequency of this form is increasing in apparent-time is strongly suggestive that the discourse marking function of LIKE represents generational change. However, it is the spread from one functional

projection, CP, to another, TP, that provides key evidence regarding the type of change represented by LIKE. As outlined in Chapter 3, generalization is a hallmark of grammaticalization (Hopper & Traugott 2003; Bybee 2003; Heine 2003), obtaining when a form comes to be used in contexts where it could not be used before. Figure 4.8 above, which displays the results for each of the three syntactic positions that make up the clause-initial context, graphically demonstrates this type of trajectory. Beginning in apparent-time with the 80 year olds, LIKE is already established on matrix CPs. It then spreads to subordinate CPs, followed by its appearance on the periphery of TP approximately 20 to 30 years later.

Viewed from this perspective, the layering of LIKE across syntactic structures is plain. It is precisely this broadening of the structural contexts in which a construction may occur that is behind the rise in frequency of a grammaticalizing form (Hopper & Traugott 2003; Bybee 2003). Thus, not only has this chapter revealed that the discourse marking use of LIKE is continuing to develop, following a path of generational change, it also represents grammaticalization, characterized by generalization and a concomitant rise in its frequency of use.

Chapter 5 **THE DETERMINER PHRASE**

A key element in the development of discourse markers is a progressive broadening of scope (Traugott 1997 [1995]; Traugott & Dasher 2002; Brinton forthcoming). Thus, as a form evolves as a pragmatic device, its modificational domain eventually becomes global, where it links sequences of discourse (Fraser 1988, 1990). In Chapter 3, we saw that the attainment of this increase in scope does not mark the end of the development of LIKE. Although it is already established as a wide-scope discourse marker among the oldest speakers in the sample, the 80 year olds, there is evidence for ongoing development across apparent-time. The most important aspect of this trajectory is the generalization of LIKE beyond matrix CPs to subordinate CPs, followed by the evolution of the TP, a distinct functional projection, as an adjunction site for the marker. In each position, LIKE continues to evaluate pragmatically the relation of the current utterance to prior discourse, and thus does not conflict with the pathways proposed by Traugott (1997 [1995]) and Brinton (forthcoming).

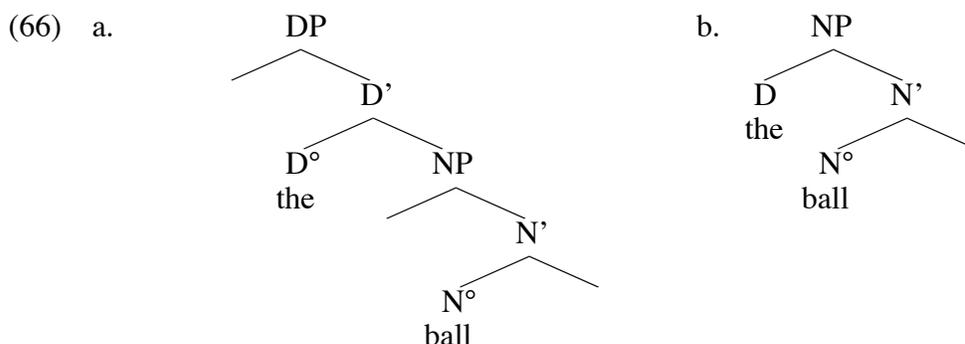
Once we begin to talk about the use of LIKE outside the rigidly circumscribed notion of a marker, that is, when it occurs within a clause, as in (65), we immediately encounter a problem.

- (65) a. I remember there being **LIKE** a solar eclipse. (I~/f/29)
b. It **LIKE** went **LIKE** seamlessly into it. (N/p/m/20)
c. She's **LIKE** really smart. (2/m/f/12)

Its domain is not the global level of discourse. It is the local level of the proposition. As a particle, LIKE modifies the element to its right, performing a number of pragmatic functions such as exemplification, clarification, metalinguistic focus, hesitation marker, evincive, etc. (Schourup 1983; Andersen 1997 et seq.). This multifunctionality is one of the primary characteristics of LIKE, but it is also one of its most problematic aspects because each function operates within the narrow scope of the following element, such as the noun phrase, the verb phrase, the adjective phrase, etc. (e.g., Underhill 1988; Romaine & Lange 1991; Andersen 1997 et seq.). We therefore have two types of scope: wide scope as a marker and narrow scope as a particle. Unless we wish to consider LIKE as a marker to be a development distinct from LIKE as a particle, then we must find a way to reconcile the evolution of the particle with the model proposed by Traugott (1997 [1995]) and Brinton (forthcoming).

In this chapter, I present the first piece of the puzzle, the determiner phrase. Reference to this context, both in the literature and in the current work, has been the noun phrase (e.g., “before NP” Underhill 1988:243; “before a noun phrase” Andersen 1997:43; “NP preceding/entering” Andersen 2001:277; “pre-NP” Wolgemuth 2003:44; “before a noun phrase” Tagliamonte 2005). I abandon that label here and refer to this context instead as the determiner phrase, DP. There are two reasons for this. First, one of the primary objectives of this investigation is to offer an account of the syntactic distribution of LIKE. This entails embedding the analysis in contemporary theoretical syntax. Since Abney (1987), the determiner has been considered the head of its own functional projection, one that takes an NP as its complement (see, for example, Bernstein (2001) and references therein).

Consequently, we have the structure in (66a), and not one along the lines of (66b), where the determiner fills the specifier slot of an NP.



Second, we will see that DP (i.e., what has traditionally been referred to as “NP” in analyses of LIKE) forms a variable context for LIKE distinct from the noun phrase (i.e., the projection which takes ‘ball’ as its head in (66a)). From this point forward, by referring to the DP context I intend the full structure in (66a). When I discuss the NP context, I am making reference to the lower lexical projection in that structure.

This distinction has ramifications for Andersen’s (2001:284) Principle of Lexical Attraction. Recall from Chapter 1 that although LIKE “has a great capacity to enter verb phrases and prepositional phrases, [...] it only rarely enters noun phrases and adjective phrases” (2001:277). One of the issues I will begin to address in this chapter is that these tendencies do not reflect any special properties on the part of the phrases themselves. Instead, the patterns observed by Andersen derive from the developmental path of LIKE and the projections it targets.

5.1 METHODOLOGY

Both Andersen (2001) and Wolgemuth (2003) present detailed accounts of the types of determiner phrases where LIKE is either rare or categorically blocked. In this section, I summarize their findings, and elaborate further considerations in circumscribing the variable context (§5.1.1). I then outline how the data were configured in the current analysis (§5.1.2). As with the CP context, I make no claim that these are the definitive constraints on LIKE. Rather, they have been gleaned from the literature and provide a stepping off point for further analysis. What they reveal, once again, is that LIKE is not random but is conditioned by language internal factors and further, that these have been active constraints on this particle from the time it was established in the DP context.

5.1.1 The variable context

As discussed earlier, the notion of scope is intimately entwined with the positions in which LIKE occurs, as whatever its pragmatic function, LIKE communicates information about the material in its domain. In circumscribing the variable context of LIKE with DPs, it is therefore necessary to restrict the analysis to those tokens where this projection is in the immediate scope of the particle, and omit those that incidentally fall within its domain, but do not comprise it in its entirety. Consequently, subject DPs are excluded. Because they occur clause-initially, a context in which LIKE is assumed to function as a wide scope marker, they do not fall within the restrictive definition of particles. As such, the variable context was defined as DPs that occur in positions other than the subject, either as

arguments of the lexical verb (i.e., direct object), as in (67), or embedded in other projections (PP, AdvP, etc.), as in (68) and (69).

- (67) a. I haven't seen **LIKE** a huge difference. (N/÷/f/45)
 b. You're like giving them **LIKE** the death stare. (N/c/f/36)
 c. I literally was **LIKE** this side of the fence. (N/m/m/26)
- (68) a. We stayed at **LIKE** a motel. (N/©/f/76)
 b. He was born in **LIKE** a slave camp. (N/o/m/25)
 c. They're doing all the calculations on **LIKE** a piece of cardboard. (I/φ/m/21)
 d. You're right by **LIKE** the coast. (3/S/m/18)
- (69) a. They just look at it towards **LIKE** the violence or mischief. (2/c/m/16)
 b. They went through **LIKE** all their old law stuff. (3/A/m/15)

Ultimately however it was necessary to further restrict this circumscription. The examples in (69) represent the sole instances of LIKE modifying a DP that forms the complement of an adverbial phrase (2%; N = 120). These tokens were excluded from the distributional and multivariate analyses because of the low occurrence of LIKE; this infrequency is an issue to which I return in section 5.2.1 below.

Once the target contexts for investigation have been defined, there are a number of constructions that are not variable. The first was noted by Wolgemuth (2003:46): personal pronouns. As exemplified in (70), LIKE categorically fails to occur with these forms, both in Wolgemuth's 1995 Ottawa data and in the current materials as well (0%; N = 252).

- (70) a. I had a crush on him. (N/i/f/53)
 b. We'd kick it around. (N/X/m/46)
 c. He built this second storey on himself. (N/E/m/36)

A second categorical context is idiomatic or fixed expressions, as in (71).

- (71) a. Someone had a higher mark, which I suppose is [the story of my life]. (I/~f/29)
 b. But [at the same time], there's a lot of science involved. (N/I/m/24)
 c. He just [kicked the crap] out of me. (I/J/m/22)
 d. It wasn't [my cup of tea]. (2/h/m/18)
 e. The storm was coming in and they're [batting down the hatches]. (I/1/m/51)
 f. You think that they're **LIKE** [friends for life]. (2/h/m/18)

Although LIKE can precede formulaic or routinized sequences (71f), it never occurs within these strings, DP or otherwise (0%; N = 30). This finding corroborates the observations in Wolgemuth (2003:47-48). It also corroborates Andersen's (2001:277-278) point regarding proper and compound nouns, since the elements that comprise these constructions are also fixed. As in the COLT data, LIKE does not intervene between these types of nominal expressions. This is exemplified in (72). Consequently, proper and compound nouns were treated as a single context.¹

- (72) a. Ryerson started after the Second-World-War. (N/I/m/53)
 b. That was also um the time of the Kent State shootings. (N/i/f/53)
 c. I bought a bus ticket to England. (I/TM/m/40)

An additional construction that was treated as providing a single context is the genitive. There are, of course, two types of possessive structures: -s genitives, as in (73), and 'of' genitives, as in (74).

¹ The inability of LIKE to occur within fixed expressions differentiates its use from lexical insertion processes, which, for example, permit profanities to intervene not only within idiomatic strings (e.g., the story of my stinking life; kicked the bloody crap out of me), but also between morphemes (e.g., un-frigging-believable).

- (73) a. They like hung out with **LIKE** their brother's friend or something. (3/Q/f/16)
 b. Last time I saw them was at **LIKE** my friend's bat-mitzvah. (2/m/f/12)
 c. I used that idea of Lucia's **LIKE** struggle and everything. (I/&/f/21)
- (74) a. The kids would go off to **LIKE** the corner of the school-yard. (I/3/f/24)
 b. You see **LIKE** the daily life of these people. (I;/m/22)
 c. Do it at the height of **LIKE** the summer. (N/I/m/24)

In both kinds of genitives, LIKE tends to precede the entire string, rather than the individual DPs that comprise it. As (73c) and (74c) demonstrate, this is not a categorical pattern, but the rates of LIKE before the second nominal expression are extremely low: 2 percent regardless of type (–s genitive N = 43; periphrastic N = 147). As such, all possessives were considered to provide one, not two, contexts for analysis. However, I will return to the topic of genitive constructions in section 5.2.1, as structural and developmental issues are implicated in the low frequencies of LIKE “within” possessive constructions.

At this point, I diverge from previous analyses in three critical aspects (e.g., Andersen 1997 et seq.; Wolgemuth 2003; Tagliamonte 2005). First, I exclude structures like the bracketed sequences in (75) in which neither the specifier nor the head of DP is overtly filled but the NP is modified by an attributive adjective. Second, I do not exclude conjoined DPs. Third, I exclude all sequences that allow for approximation.

- (75) a. There'd be [yellow dinosaurs] with lights on them. (I/œ/m/34)
 b. They die through [different things]. (3/B/m/16)

Regarding (75): standard methodological practice is to focus on constructions in which the locus of syntactic variation can be unambiguously determined along structural parameters (see §1.4.2, Chapter 1). According to this maxim, constructions such as (75)

were excluded because the slot filled by LIKE cannot be deciphered. Moreover, it cannot be tracked accountably. Consider, for example, (76).

(76) I'm listening to **LIKE** younger people. (I/%/f/28)

Does LIKE modify the DP [_{DP} [_{NP} [_{AP} younger] people]], the NP [_{NP} [_{AP} younger] people]], or the AP [_{AP} younger]? When a DP contains no overt material, the position of LIKE cannot be determined from the relative order of the elements in the phrase. In contrast, in (77a), where the specifier position is lexically filled, and in (77b), where the head is filled, we can assume that LIKE scopes over the full DP.

(77) a. Like he got into **LIKE** some serious drugs. (N/o/m/25)
 b. I remember there being **LIKE** a solar eclipse. (I/~f/29)

Despite the fact that DP contains overt material, the data in (78) are less straightforward, since two phrasal projections may provide an adjunction site for LIKE: NP and AP.

(78) a. They were all **LIKE** good friends for a really, really long time. (N/≤/m/26)
 b. Like everyone's in this **LIKE** torrid race to grow up. (N/E.m/36)

There is reason to suspect, however, that this position is in fact NP, but I set this aside until section 5.2. The crucial point here is that LIKE cannot be situated on DP in these structures, since it follows the determiner. Thus, (78) is contrasted with (76), because DP can be ruled out on structural grounds in the former, but not in the latter.

Returning then to sequences such as those in (76), one possible means of distinguishing between a DP or an NP/AP interpretation could be prosodic: if the stress falls on the adjective, for example, we may decide that LIKE modifies the NP/AP. Similarly, rising intonation that peaks over the noun could suggest that LIKE modifies the full DP.² However, this raises the issue of accountability. Even if the position of LIKE in every instance similar to (76) could be unambiguously determined using prosodic (or other) cues (and this is doubtful), the forms where LIKE does not occur must still be accounted for. The problem is that there is no objective means for determining which level of structure LIKE would target in each case if it were to be used. Consequently, these problematic cases (i.e., (75) and (76)) were simply removed from the analysis (N = 168). Tokens such as those in (77) and (78), where DP contains overt material, were retained.

However, this raises another important point. Recall from above that I do not treat the DP as a monolithic context, but distinguish instead between DPs and NPs. In contrast, regardless of the syntagmatic position of LIKE in a DP, Andersen (2001:278) considers all tokens as comprising a single context. Thus, (78) would be treated as a DP marked by LIKE. In the current analysis, such tokens were treated as unmarked DPs. In other words, LIKE “within” a DP projection is distinguished from LIKE occurring on the periphery of DP.

The decision to include conjoined structures was based on two factors. Although Wolgemuth (2003:47) reports that LIKE never modifies the second of conjoined DPs, this finding is not corroborated in the Toronto data. As exemplified in (79), a full 12 percent of these phrases are preceded by the particle (N = 130).

² I am not claiming that this is the correct way to analyze these types of suprasegmental factors with regard to pragmatic features. They are offered here simply as a means of elucidating my point.

- (79) a. I could pay that much and get a diploma or **LIKE** a degree. (N/I/m/24)
 b. There's like a guy with like a bell and **LIKE** a hood. (2/h/m/18)

Further, I assume that LIKE targets maximal projections. In the case of DP adjunction, the inclusion of conjoined DPs as a variable context is predicted by structural factors. By the same token, however, LIKE should not occur before a noun that is conjoined below this level (80):

- (80) a. They just look at it towards like the violence or mischief. (2/c/m/16)
 b. Now he's living with my nanny and poppa. (2/e/f/11)

Though the Ns are low, this prediction is met by the data; LIKE categorically fails to appear in this context (0%; N = 16). Thus, conjoined DPs were included and only conjoined NPs were excluded.

Finally, because LIKE functions as an adverb in the context of quantified expressions (Chapter 2), DPs allowing approximative modification, as in (81), were excluded.

- (81) a. It could have taken you all day to go **LIKE** thirty miles. (N/©/f/76)
 b. My salary was **LIKE** twenty-seven-thousand dollars. (I/1/m/51)
 c. The guy weighed **LIKE** a-hundred pounds. (I/*/f/30)

Following these methods, a total of 4408 tokens were retained for analysis. As with the CP context, these were then coded for a series of internal factors, gleaned from the literature, as well as for speaker age and sex. I now turn to a discussion of this aspect of the analysis.

5.1.2 Coding and analysis

The coding procedure for sex and age was discussed in Chapter 4, and the motivation for these factors was outlined in Chapter 1; I do not repeat those points here. This section will focus on the internal linguistic factors that I consider in the analysis of LIKE in the DP context. These are: the syntactic status of the DP, the definiteness of the DP, and modification of the DP.

Syntactic status

There are two potentially competing findings in the literature. On the one hand, there is the well-known result that DPs provide one of three preferential positions for LIKE (e.g., Underhill 1988; Andersen 1997 et seq.; Wolgemuth 2003; Levey 2004). On the other hand, there is the more recent finding that the type of phrase significantly constrains the probability of “LIKE-insertion” (Andersen 2001:277). These two results intersect where ‘occurring within a phrase’ is tantamount to modifying the DP that that phrase takes as its complement. For example, PPs and AdvPs both subcategorize for a DP (PPs do so categorically and AdvPs optionally).

As we saw above, however, LIKE almost never modifies a DP within an AdvP (2%; N = 120). This seems to corroborate Andersen (2001:277), who reports that LIKE never enters AdvPs in his data.³ It also suggests, however, that not all DPs take LIKE at the same rate. In order to test this hypothesis, the data were coded according to whether the DP is the argument of the VP, (67), or the complement of a PP, (68).

³ However, Andersen’s (2001:277) sole example consists of the predicative construction “She tried to act **LIKE** really friendly” and does not entail a DP.

Definiteness

A number of pragmatic functions have been ascribed to LIKE (e.g., exemplification, clarification, metalinguistic focus, hesitation marker, evincive, etc.). I do not consider these here because, following Poplack and Tagliamonte (2000:321), “such nuances tend to reside in speaker intent and hearer inference, both of which are inaccessible to the analyst” (see also Dines 1980; Stubbe & Holmes 1995; Vincent 1986; Dubois 1992; Dubois & Sankoff 2001). Rather than devise an arbitrary classification system, I focus on factors that can be established through objective means. With this in mind, many researchers have argued that LIKE may also mark non-contrastive focus (e.g., Underhill 1988; Meehan 1991; Romaine & Lange 1991). This can be defined as introducing new as opposed to given or presupposed information (Halliday 1967; Chomsky 1971; Rochemont & Cullicover 1990).

Working from this assumption, Dailey-O’Cain (2000) presents an analysis of LIKE in focus constructions. Although the crux of her work is the social and attitudinal aspects of this particle and not its function, the crucial point is that focus constructions contain a non-negligible proportion of LIKE: among those who make the most propitious use of this particle, males between the ages of 14 and 29, its overall distribution is 14 percent (2000:66). In short, Dailey-O’Cain’s findings support the suggestion that LIKE performs in the textual component of discourse, marking information structure (see Chapter 1, §1.3).

In a statement that appears to contradict the focus analysis, Andersen (2001:247) claims: “rhematic status does not seem to be a sufficient condition for LIKE-qualification.” In support of this position, he points out that LIKE does not introduce new information “indiscriminately” because it does not mark information that is familiar, even if it is new to

the discourse (e.g., names of friends). This is not an argument against focus. What Andersen questions is the newness of the information. While it is true that focus constructions introduce new material, the relevant contrast is not between new and old (i.e., previously established in the discourse), but between new and known, since ‘given’ or ‘presupposed’ in the above definition presumes that the information is in some way familiar to the interlocutors.

Although we cannot disambiguate between what is familiar and what is not, there is a criterion to distinguish between new and known: definiteness. The English articles encode this information directly, as the indefinite article indicates new content while the definite article denotes known material.

Consequently, tokens were coded for whether they were definite, as in (82), or indefinite, as in (83). Included in the former category were the definite article (82a), definite demonstrative pronouns (82b), and personal pronouns (82c).

- (82) a. Maybe there is a catch to **LIKE** the homework. (3/K/f/18)
 b. So usually like I just look for **LIKE** those types of signs. (3/J/m/18)
 c. It was **LIKE** my very first day. (N/f/f/32)

Included in the latter category were the indefinite article (83a) and the indefinite use of this (Prince 1981), exemplified in (83b).

- (83) a. He lives like right on Queen Street in **LIKE** an apartment. (I/J/m/22)
 b. I think the drummer is so cute. He’s **LIKE** this big Irish guy. (N/r/f/22)

Note that instead of testing focus, which is compromised by issues of consistency deriving from subjectivity, this configuration of the data tests the tangible and objective

measure of definiteness. Abstracted from the literature is the hypothesis that LIKE will be less frequent in definite contexts, as these will often, though not always, correspond to known information (cf. (82c)).

Modification

The purpose of Chapter 2 was to demonstrate that when LIKE occurs in the context of quantified expressions, it functions adverbially, carrying the full force of other approximative adverbs such as about. I suggested that this lexical function is likely to have inflated the number of DPs that are modified by LIKE in previous investigations (see, for example, Andersen 1997:44, 2001:277-278; Wolgemuth 2003:68; Tagliamonte 2005). Recall from Figure 2.1, for example, that among speakers between the ages of 10 and 29, a full 32 percent of measurable DP and APs are approximated using LIKE (N = 1437). As Andersen (1997:43) points out, however, a number of constructions are modified by LIKE where this lexeme does not express inexactness (see also Schourup 1983:31). Examples are given in (84).

- (84) a. He saw **LIKE** all the kids walking around. (N/t/f/36)
 b. There tends to be **LIKE** quite a dramatic split in there. (I/5/f/24)
 c. He's got **LIKE** this whole band with him. (N/I/m/24)
 d. I mean you meet **LIKE** a lot more people. (2/u/m15)

Since these types of tokens cannot be interpreted as approximative, they must be a discourse use, i.e., the particle. The question is, once the adverbial tokens discussed in Chapter 2 are omitted from the analysis, do quantified and degree-modified DPs continue to contain high proportions of LIKE? In order to test for this, tokens were coded for whether

they were modified either by a quantificational or degree adverb (i.e., overt material in SPEC DP), as in (84a,b), or periphrastically, as in (84c,d), or whether the DP is unmodified.

In sum, I have suggested three internal factors that may condition the use of LIKE with DPs and configured the data so as to test for their effects. These are status (argument, complement), definiteness (definite, indefinite), and modification (modified vs. unmodified). I now turn to the results, focusing first on the external factors of sex and age.

5.2 RESULTS

5.2.1 Distributional analysis

Table 5.1 reports the overall distribution of LIKE in the DP context according to age. Note the 80 year olds: in over 360 tokens, LIKE occurred just once, not enough to register in the distributional analysis (.003%).

TABLE 5.1
Distribution of LIKE in the DP context across apparent-time

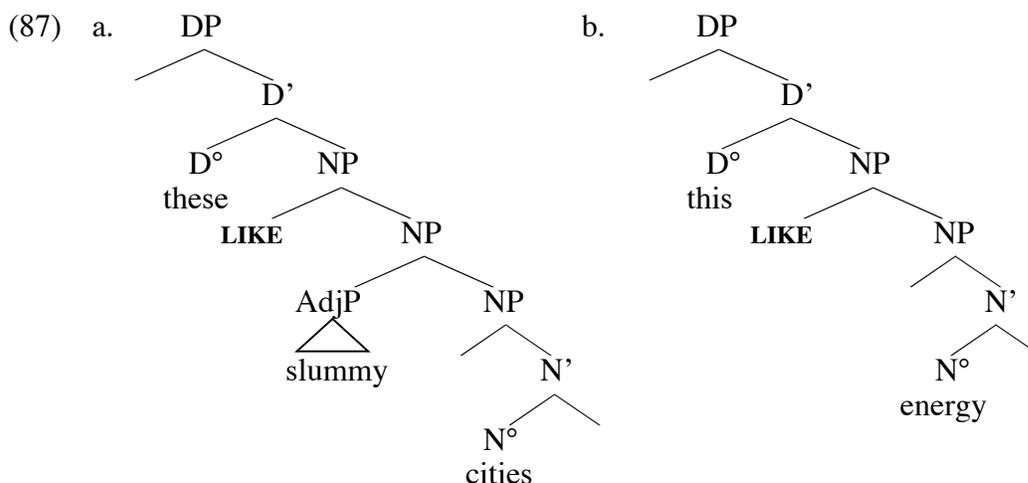
	%	N	total
10-12	16	424	} 4047 OD = 10%
15-16	18	368	
17-19	17	467	
20-24	16	466	
25-29	9	488	
30-39	7	463	
40-49	5	337	
50-59	3	349	
60-69	2	372	
70-79	1	313	
80+	0	361	
total		4408	

In contrast, recall that in Chapter 4, the use of the discourse marker (i.e., where LIKE occurs in clause-initial position) was already well established among this same group of speakers. This suggests that the DP constitutes a later stage in the development of LIKE. In light of the developmental path proposed by both Traugott (1997 [1995]) and Brinton (forthcoming) for discourse markers, the possibility that the particle, which surfaces in clause-internal positions, may have evolved after the marker has important ramifications for theories concerning the grammaticalization of discourse features. This is a point to which I return in Chapter 8, once evidence from adjective and verb phrases is available.

However, before leaving this topic, I would like to address the tendency for LIKE to also appear within a DP, adjoined to either AP or NP. These types of tokens are not included in Table 5.1 since they comprise a context distinct from DP. Andersen (2001:276-277) reports that LIKE is rarely situated within a DP; of the 220 occurrences of LIKE that he considers to be nominal, just 27 enter the phrase. Of the more than 4000 DPs extracted from the Toronto corpus, only 47 contain LIKE in a position other than on the periphery of DP: 26 precede an adjective (and follow a determiner or quantifier/degree adverb), as in (85), and 21 precede a noun (and follow a determiner or quantifier/degree adverb), as in (86).

- (85) a. We had to go through these LIKE slummy cities and stuff. (I/7/m/35)
 b. She had this LIKE old-fashioned tricycle. (3/V/m/11)
- (86) a. They have this LIKE energy you-know? (I/&/f/21)
 b. Like we were supposed to memorize some LIKE parts. (3/V/m/11)

Although LIKE in (85) could modify either the NP or the AP, I will assume here that it is adjoined to NP.⁴ Thus, we have the structures in (87).



Crucially, examples such as (85) and (86) do not occur among speakers of all ages. Indeed, they are first attested among speakers in their thirties, where they are so infrequent that they account for just .004 percent of the data (N = 463). Among speakers in their late twenties, the NP position accounts for 1 percent of the data, a proportion that rises to 2 percent among 20-25 year olds, and 3 percent among 17-19 year olds. These facts are summarized in Table 5.2.

⁴ As I discuss next, constructions like (85) and (86) appear in the data simultaneously. Further, in Chapter 6 I will show that the AP position was established well before (85) is attested (see also (76), a construction which first occurs in the data among speakers in their sixties). These two facts suggest that LIKE in (85) represents NP adjunction, not AP adjunction.

TABLE 5.2
Distribution of LIKE in the NP context across apparent-time

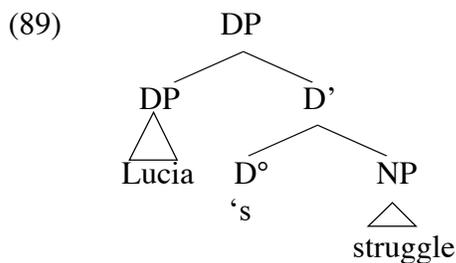
	%	N	total
10-12	2	424	} 2213 OD = 2%
15-16	2	368	
17-19	3	467	
20-24	2	466	
25-29	1	488	
30-39	0	463	
40-49	0	337	
50-59	0	349	
60-69	0	372	
70-79	0	313	
80+	0	361	
total		4408	

Thus, the distributional evidence not only suggests that the DP context is a later development than the CP context, but it also strongly suggests that the NP context, a position which comprises the ability for LIKE to “enter” a DP, was established approximately half a century after the DP. It is this that accounts for the lower occurrence of LIKE following a determiner, since the internal position developed later than the ‘outer’ DP adjunction site.

The late emergence of the NP as an adjunction site for LIKE explains the tendency for the particle to appear at the front of an –s genitive rather than within one. Recall the data from (73), which I repeat here as (88).

- (88) a. They like hung-out with LIKE their brother’s friend or something. (3/Q/f/16)
 b. Last time I saw them was at LIKE my friend’s bat-mitzvah. (2/m/f/12)
 c. I used that idea of Lucia’s LIKE struggle and everything. (I/&/f/21)

For this type of possessive, I assume the structure in (89):



Based on this structure, it is clear that until the NP developed as an adjunction site for LIKE, the only position in which we would predict LIKE to appear is that in (88a,b). The example in (88c), which can only represent NP adjunction, is the only one of its kind in the data. Note the age of the speaker: she is 21. Thus, this example fits the proposed chronology for the emergence of the NP position.

Returning then to the discussion of the DP, observe that in Table 5.1, once LIKE begins to modify DPs, its rate of use continues to rise. This type of trajectory, where frequencies gradually increase among younger speakers, is indicative of ongoing, generational change (Labov 1994:84). In other words, even though there is a large jump in the proportion of LIKE among speakers under the age of 25, it is clear that the use of LIKE with DPs is not a passing fad that will disappear with age. Rather, the apparent-time results suggest that these younger groups have inherited the context from older members of the community who themselves use LIKE in this position.

At this point it is fitting to turn to the issue of speaker sex. Recall that in the analysis of the CP context in Chapter 4, the only correlation between this factor and LIKE appeared where the frequency is highest, i.e., among speakers aged 15-24 (Figure 4.2). Within these groups, the females emerge as the leaders. Is the same result apparent in the DP context as well? As Figure 5.1 reveals, it is not.

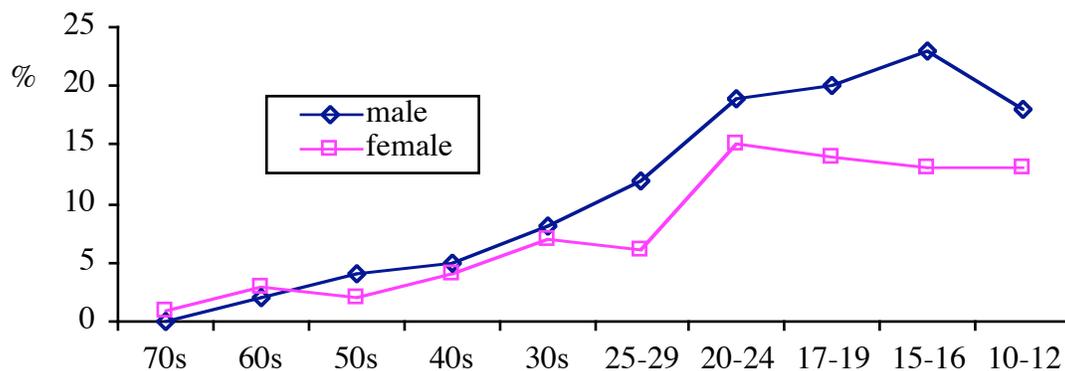


FIGURE 5.1 Cross-tabulation of LIKE in the DP context by age and sex

Figure 5.1 differs from Figure 4.2 in two critical respects: 1) there is a clear male preference and 2) this correlation is manifest nearly from the outset, not simply where the rate of use is highest. We therefore have evidence for two distinct patterns: with the discourse marker, females eventually emerge as the leaders, but with the particle, at least for the DP context, males are at the forefront virtually from the time LIKE first begins to modify these structures.

Until we see what the AP (Chapter 6) and VP (Chapter 7) add to this picture, there is little point in interpreting these patterns yet. I will simply point out that previous research has suggested that the primary difference between males and females in the use of discourse features is qualitative, not quantitative (e.g., Dubois 1992; Erman 1992). Thus, a possibility suggested by the comparison of the CP and DP results is that females may come to favour the marker, while males favour the particle. I leave the issue open for now, but return to it in Chapter 8 (§8.2).

We have seen the effects of sex and age on the use of LIKE in the DP context, but what of the internal constraints? The first possible condition is the status of DP itself. Since Andersen (2001:277) reports that the type of phrase significantly affects the likelihood of

whether or not the particle can appear between constituents, it is possible that not all DPs are equally receptive to modification by LIKE. Recall, for example, that LIKE rarely occurs with DPs embedded in AdvPs (2%; N = 120). The distributional results in Table 5.3 further support this hypothesis.

TABLE 5.3
Distribution of LIKE according to the status of the DP

	%	N
Complement	5	1649
Argument	13	2398
total		4047

LIKE is nearly three times more frequent with DPs that are part of argument structure (i.e., direct objects) than with those that form the complement of some other projection, represented here by PPs. Cross-tabulation by age, as in Figure 5.2, reveals that this effect is consistent across the community.

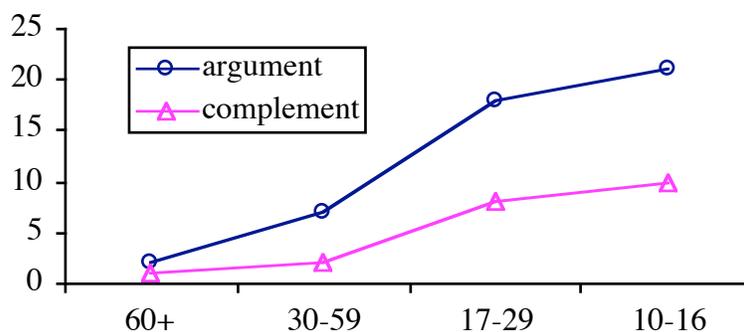


FIGURE 5.2 Cross-tabulation of status by age

Why should this be? Structurally, all DPs are identical. Consequently, the syntax cannot be directly responsible for the difference between complement DPs and argument DPs. However, I have suggested that differences in the likelihood that LIKE will ‘enter’ a

given phrase derive from the projection to which it adjoins. In the sense in which Andersen framed his result, my claim remains true, which I will demonstrate further in Chapters 6 and 7. Nonetheless, the results in Table 5.3 and Figure 5.2 reveal that some other factor is also at work, at least with respect to DPs.

One possible explanation might come from frequency effects (e.g., Bybee 2003). Table 5.4 reproduces the information from Table 5.3, but I have added in the results for DP complements of AdvPs, which were excluded from the analysis because of the extremely low rate at which LIKE appears in this context. Note the correspondence between the frequency of a given construction and the frequency of LIKE: the more common the context, the greater the proportion of LIKE becomes.

TABLE 5.4
Distribution of LIKE according to the status of the DP II

	%	N
Complement — AdvP	2	120
Complement — PP	5	1649
Argument	13	2398

The effects of frequency have been well articulated in the lexical diffusion of sound change: high frequency words typically undergo change faster than low frequency words (e.g., Fidelholtz 1975; Hooper 1976; Wang 1977; Phillips 1980; Labov 1994; Bybee 2001, 2002; etc.). More germane, however, is the effect of frequency on certain grammaticalization processes. Bybee (2003:610-613), for example, uses the development of auxiliary can during the Middle English period to demonstrate how a high type frequency (e.g., verbs of communication) corresponds to the increased generality of the schema can +

infinitive. This in turn corresponds to a higher degree of grammaticalization for constructions such as can say and can tell.

In the case of LIKE, we have a situation where arguments represent the class of DPs with the greatest type frequency, and it is in these constructions that we find the highest proportion of the particle. This suggests that the consistency of the distinction between complements and arguments across apparent-time reflects stages in the evolution of LIKE: the greater rate of LIKE in arguments derives from their more grammaticalized status as adjunction sites for the particle. Stated another way, arguments, as high frequency constructions, undergo change faster than phrasal complements, which are lower frequency constructions.

An interesting consequence of this finding is that the low rate of LIKE within periphrastic genitives that was discussed in section 5.1 (2%; N = 147) and which ultimately resulted in these structures being treated as a single variable context, may derive from this broader effect of frequency on the development of the particle. Specifically, in this type of possessive, (N), the second DP is hosted within a PP (e.g., the height [of LIKE the summer]). Thus, it is possible that LIKE is infrequent here simply because of the projection in which the DP is situated.

Now, what of the other internal constraints? How do they condition the use of LIKE? Consider Figures 5.3 and 5.4, which are analogous to Figure 5.2.

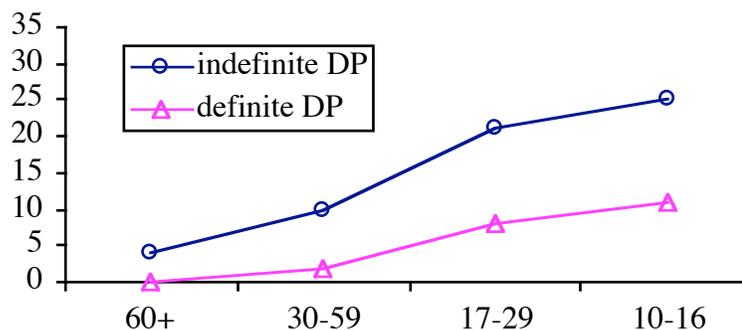


FIGURE 5.3 Cross-tabulation of definiteness by age

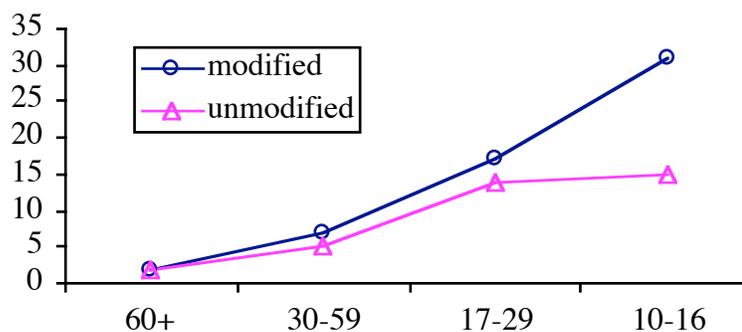


FIGURE 5.4 Cross-tabulation of modification by age

Figure 5.3 reveals that LIKE is more frequent with indefinite DPs than definite ones. Further, just as with the status of the DP in Figure 5.2, this effect is apparent from the oldest speakers who use the particle in the DP context (i.e., 60 and 70 year olds). Figure 5.4 tracks the rate of LIKE by modification of the DP across apparent-time. Among speakers in their sixties and seventies, LIKE is equally as frequent regardless of whether the DP is modified or not, but within the next generation the modified context begins to break away.

These distributional results therefore suggest that LIKE is not only conditioned by linguistic factors, but that these factors have been operational from the inception of the DP context as a possible adjunction site for this particle. Since it has been argued that it is the

operation of internal factors that determines the variable grammar (Poplack & Tagliamonte 2001:93-94), these results reveal that despite the association of LIKE with younger speakers, it is uniformly constrained across the entire community. In other words, speakers of all ages share the same grammar for LIKE. The importance of this result cannot be overstated. Not only does it support an analysis of ongoing change, one which is “percolating successively through the generations” (Tagliamonte & D’Arcy 2004:503), it also unmistakably demonstrates that younger speakers have not laid claim to LIKE *ex nihilo*. Although their frequency of use is higher, they use LIKE in the DP context following the parameters already established in the discourse of the older members of the population.

5.2.2 Multivariate analysis

To provide support to the distributional analysis, I turn to the results of three independent multivariate analyses that demonstrate the consistency of these effects across the Toronto English sample. These results are given in Table 5.5.⁵

The over-arching observation to be made about these results is the uniformity with which the constraints operate across the sample. Although not all factors are significant in each age group, the direction of effects within factor groups is consistent from the oldest speakers to the youngest: males favour over females, and modified, argument, and indefinite DPs favour over unmodified, complement, and definite DPs. In other words, Table 5.5

⁵ The oldest age group, the 60 and 70 year olds, has not been included because LIKE occurs too infrequently to allow for statistical analysis (2%; N = 685). In a separate run in which all the age groups were run simultaneously (not shown here), the following factor weights were assigned to each cohort: 60+: .17; 30-59: .39; 17-29: .65; 10-16: .71.

reveals that for 10 year olds and 59 year olds alike, the conditions affecting the use of LIKE are identical.

TABLE 5.5
Contribution of external and internal factors on the probability of LIKE in the DP context

INPUT	30-59 .04			17-29 .12			10-16 .15		
Definiteness⁶									
indefinite	.70	10	346	.62	21	472	.64	25	245
definite	.36	2	506	.41	8	603	.41	11	393
range	24			21			23		
Status									
argument	.58	7	701	.57	18	861	.56	21	469
complement	.37	2	448	.39	8	560	.42	10	323
range	21			18			14		
Modification									
modified	[.56]	7	113	[.52]	17	136	.70	31	72
unmodified	[.49]	5	1036	[.50]	14	1285	.48	15	720
range							22		
Sex									
male	[.53]	6	577	.55	16	698	.58	20	417
female	[.47]	5	572	.45	11	723	.44	13	375
range				10			14		
TOTAL N			1149			1421			792

* Non-significant factors enclosed in square brackets.

Notice, however, the shifting strengths of the factors across apparent-time. Definiteness is the strongest constraint on LIKE, consistently ranked first in every age group. It is therefore a stable factor affecting the use of LIKE in the DP context. The same is not true for any of the remaining three factors considered in Table 5.5. The sex effect, not significant among 30-59 year olds, is selected among 17-29 year olds and gains in strength among the

⁶ Since the coding for definiteness relies on the form in D°, the head of DP, tokens in which the determiner slot was empty have been removed from the test for this factor (see Tagliamonte & D'Arcy 2004:fn.12, 512).

youngest speakers, where it is as strong as the status of the DP (both have a range of 14). Indeed, this last factor can be seen to lose strength. Within the two oldest age groups in Table 5.5, it is the second ranked constraint. In the youngest age group, it has fallen to third, behind modification of the DP and definiteness. However, it is the factor group of modification that displays the most dramatic shift of all the constraints on LIKE. Not significant for speakers between the ages of 17 and 59, it leaps into effect among the 10-16 year olds, becoming the second ranked factor. In fact, it operates at nearly equal strength to definiteness, the quintessential constraint on LIKE in the DP.

What does all this mean? First, let us consider the results for speaker sex. The increasing strength of this social factor indicates that as the frequency of LIKE rises, the difference between males and females becomes more marked. This type of trajectory is typical as a linguistic change progresses. Consider phonological change. The study of a number of phonological variables has revealed that the sex effect is greater in older changes than it is in more recent innovations (e.g., Labov 1990:242). The same result was reported by Tagliamonte and D'Arcy (2004) for a grammatical feature, quotative be like. Thus, there seems to be a recurrent correlation between the degree of diffusion and the degree to which the sexes will be differentiated. The results in Table 5.5 clearly establish the operation of this effect for a discourse variable. This result was perhaps predictable, yet what is unusual about LIKE is that it is males who are at the forefront, not females. However, as discussed above, I set this aspect of LIKE aside until Chapter 8 when we have corroborating evidence from the AP and VP contexts.

Turning to the internal constraints, the results for definiteness are particularly striking since they reveal that this factor has been operative since the DP was established as a variable context for LIKE. That it is also the strongest factor in each age group (note the ranges of 24, 21, and 24 respectively), suggests that it is a fundamental constraint on the use of LIKE in the DP contexts. Thus, even though indefinite constructions are not synonymous with focus constructions, the consistent and significant favouring effect of indefinites corroborates earlier work (e.g., Underhill 1988; Meehan 1991; also Dailey-O’Cain 2000) positing that, in addition to the many interpersonal functions that have been proposed, most notably in Schourup (1983) and Andersen (1997 et seq.), LIKE can mark new information.

Within the youngest group, however, the contribution of definiteness is nearly equal to that of modification (23 and 22 respectively). The difference in the ranking of this latter factor is quite dramatic: it is marginal and non-significant in the two older groups, but then rises to become the second ranked constraint among pre- and adolescent speakers. In particular, DPs that are quantified or intensified by a degree adverbial come to be a strongly favoured context for LIKE at .70. In contrast, DPs that are unmodified remain relatively neutral, just slightly disfavoured at .48.

We therefore have evidence for two distinct patterns. On the one hand, LIKE occurs slightly more often with modified DPs than with unmodified DPs, but the difference is minor. This is evident in both the distributional analysis and the multivariate analysis. This pattern, which is visible across the majority of the sample (i.e., the 17-59 year olds; see Figure 5.4), supports the suggestion that the inclusion of approximative contexts exaggerated the numbers of DP marked by LIKE in previous analyses, since when these are

excluded, quantified DPs cease to be substantially differentiated from unmodified DPs. On the other hand, LIKE is significantly more frequent with modified DPs among 10-16 year olds. There are two possibilities for this sudden shift within the youngest group: 1) it could be an age-related phenomenon or 2) it could signal the rise of a formerly incipient constraint. Let us consider each of these in turn.

First, consider the implications if the frequent use of LIKE with modified DPs is age-graded. In this view, younger speakers generalize LIKE to all types of quantification, regardless of function. If this is the case, then future research should reveal a recurrent pattern whereby preadolescents and adolescents use high proportions of LIKE with modified DPs, but soon after, the levels lower and there is a retrenchment toward adult norms (Chambers 2002). In fact, this pattern is evident in Figure 5.5, where the frequency of LIKE nearly levels across DPs among the 17-19 year olds. From this point upward in the age scale, modified DPs are not significantly differentiated from those that are unmodified (cf. Figure 5.4 and Table 5.5). In short, the late adolescent model reflects the adult pattern.

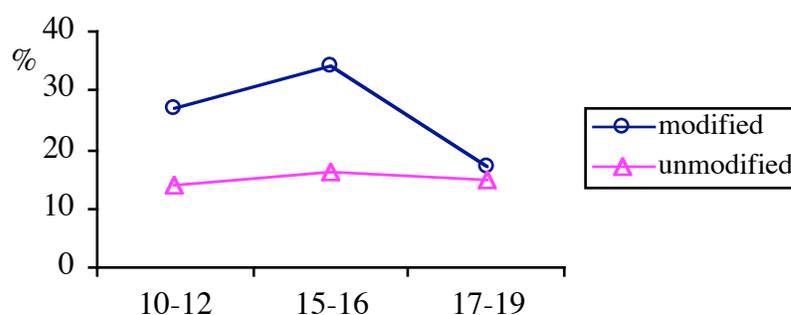


FIGURE 5.5 Distribution of LIKE according to modification of the DP

There are two further pieces of evidence that suggest age-related behaviour. First, the proportion of LIKE with unmodified DPs is constant across the three age groups represented

in Figure 5.5. It is only the frequency of LIKE with modified DPs that changes. Second, and more important, is that among 10-16 year olds, the rates of use are virtually identical across functions: adverb 33 percent (N = 509) vs. particle 31 percent (N = 72). This suggests that 10-16 year olds do not distinguish between the two kinds of quantified contexts (i.e., those that can be approximated and those that cannot).

Second, consider the implications if the significant effect of quantified DPs among the 10-16 year olds reflects the crystallization of a constraint that was only incipient in the speech of older age groups in the Toronto sample. If so, then modification should continue to play a significant role in conditioning LIKE as the speakers currently between the ages of 10 and 16 enter adulthood. However, this premise is suspect; younger speakers participate in change (e.g., Labov 1994, 2001; Eckert 1988, 2000; D'Arcy 2005) but preadolescents are not usually the innovators. Indeed, 10-12 year olds typically reflect the patterns of their parents, not those of their peers (Eckert 1988). Crucially, Figure 5.5 reveals that the large proportional difference between modified and unmodified DPs that is evident among 10-16 year olds in Figure 5.4 is not due to the adolescents alone. The preadolescents also make a clear distinction between the two types of DPs. This type of pattern is reminiscent of Macauley's (1977) findings for glottalization in Glasgow, where the use of glottal stops recedes dramatically during adolescence. The only difference here is that the frequent use of LIKE in modified contexts perpetuates into adolescence, instead of dropping off after pre-adolescence. Nonetheless, the similarities in the patterns between Glasgow and Toronto suggest that the high rates of LIKE in modified contexts among 10-16 year olds marks a

maturational stage (Chambers 2003:188) and not the sudden emergence of a linguistic constraint.

Finally, we have the syntactic status of the DP. In light of my suggestion that this factor reflects varying degrees of grammaticalization, the weakening of its effect that is revealed by the regression analyses is particularly intriguing. It falls from competing closely with definiteness in the two older groups to being tied with sex as the third ranked constraint in the DP context among 10-16 year olds. Thus, it is possible that as the development of LIKE continues, this factor will continue to decline in strength. We cannot project how frequent LIKE will become, nor can we presume to understand the full range of mechanisms driving its evolution. However, as LIKE comes to be used in a wider range of contexts, here different complement structures, these contexts will eventually become entrenched as LIKE-modifiable structures as well. If it is indeed this type of developmental trajectory that is responsible for the differentiation between argument and complement, then a leveling of the distinction along grammatical lines should lead to a leveling of the constraint as well. Time will reveal whether this hypothesis has substance or not.

5.3 DISCUSSION

The analyses presented in this chapter have shown that the frequency of LIKE in the DP context is increasing in apparent-time, and that this particle generalized to NPs much later than when it first began appearing on the periphery of DPs. I have suggested that it is this latter expansion that accounts for the lower rates of LIKE within a DP. In other words, LIKE

is more likely to precede a full DP than it is to enter one because the internal position is a more recent development.

Further, the analysis has revealed that unlike the CP context, in the DP context it is males who are at the forefront of LIKE use. There are therefore qualitative differences between the use of LIKE in each of these contexts: as a marker, LIKE is associated with females, but as a particle, at least in the DP context, it is associated with males.

Finally, the use of LIKE is not indiscriminate. From its inception with DPs, it has been significantly favoured in indefinite contexts and in argument positions.

In sum, the analysis of the DP context has shown that LIKE is subject to both social and linguistic factors, and that these constraints have been operational from the establishment of the DP as an adjunction site for this particle. Consequently, these results indicate that those who use LIKE the most, adolescents, are not the linguistic mavericks intimidated by popular stereotypes, but instead, share a variable grammar for this feature with much of the speech community, all age groups under 80. By implication, just as with the CP context, this chapter has shown that LIKE does not appear randomly within linguistic structure.

However, this chapter has raised a number of questions:

- 1) Why does sex exert a different effect on LIKE as a marker than it does on LIKE as a particle in the DP context?
- 2) Why should the status of the DP as an argument or a complement be relevant?
- 3) Will this distinction eventually level?

None of these can be answered definitely. I will consider issues involving sexual differentiation in Chapter 8. Regarding the status of the DP, I have suggested, following the work of Bybee (2001, 2002, 2003) and the lexical diffusion model (e.g., Fidelholtz 1975;

Hooper 1976; Wang 1977; Phillips 1980; Labov 1994) that this result derives from frequency effects, with the most frequent type of structure, arguments, representing a more developed context in the evolution of LIKE. This analysis is supported by the weakening strength of this factor in apparent-time, which may indicate that the entrenchment of complement structures as a context for LIKE is gaining ground.

In the next chapter I turn to adjective structures. These do not belong to the “Group of Three” — those contexts that are recurrently reported to contain the highest proportions of LIKE. However, the investigation of DP has raised the question of what their role in the development of LIKE may be. Consequently, before presenting an analysis of the VP, the third alleged preferential host for this particle, I turn my attention to predicate adjectives because these, in contrast to the attributive contexts discussed earlier ((75) and (76)), provide a clear and unambiguous context for analysis.

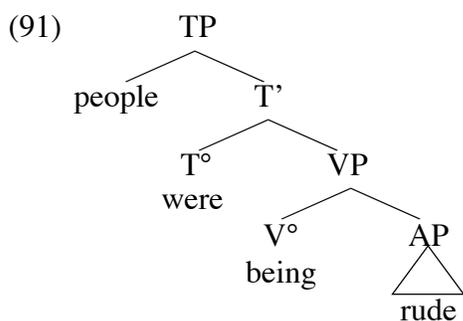
Chapter 6 **PREDICATE ADJECTIVES**

We have seen that the ability of LIKE to appear “within” a determiner phrase is a later development in the history of this particle. Thus, while LIKE adjoins to DP in the grammar of speakers in their seventies, unambiguous evidence for its occurrence on the left edge of NP first appears only among speakers in their thirties. However, the DP context provides a second possible adjunction site, the AP. Consequently, in this chapter, I turn my attention to predicate adjectives, one of the slots reported by Underhill (1988:243). His examples of LIKE with adjectives, of which there are only two, are strictly predicative. Subsequent investigations have not differentiated between the two types of adjectives, but have corroborated the AP as a position for LIKE while simultaneously pointing out its relative infrequency in this context (e.g., Andersen 1997 et seq.; Wolgemuth 2003; Levey 2004; Tagliamonte 2005). Since Underhill (1988) is the only study to have specifically addressed predicate adjectives, why distinguish between predicative and attributive constructions here? The answer is two-fold.

First, attributive APs are embedded within a DP, adjoined to NP (see (87a) in the previous chapter). As a consequence, if neither the specifier nor the head of the DP contains overt material, the position of LIKE cannot be disambiguated on structural grounds. Specifically, is it adjoined to DP, NP, or AP? In contrast, predicative constructions provide a clearly delineated variable context. Consider (90).

- (90) a. She's **LIKE** dumb or something.
 Like I love her but she's **LIKE** dumb. (3/T/f/18)
- b. All these people were being **LIKE** rude to my supervisor, so I was like there making the waffles and all these people were like being rude. (3/P/f/16)

Following Andersen (1997 et seq.) and others (e.g., Schourup 1983; Underhill 1988; Romaine & Lange 1991; Siegel 2002, etc.), I assume that when LIKE occurs in clause-internal positions (i.e., as a discourse particle), it functions pragmatically to modify the element that follows it. Further, I assume that this effect can be captured syntactically, with LIKE adjoining to the left periphery of the modified element. In predicate adjective constructions, henceforth referred to as PredAPs, I take the AP to form the argument of the lexical verb, as in (91).¹



Crucially, no projection intervenes between the VP and the AP in (91). Consequently, if LIKE appears between the lexical verb and the adjective, as in (90), it must adjoin to either the AP or within the adjectival projection at some point above the head, A°. I have argued in

¹ Two comments are required regarding (91). First, as in previous chapters, I collapse nodes unrelated to the immediate discussion. Second, I represent nonfinite be as occurring within a VP. This is for purposes of exemplification; I take no position regarding whether copular constructions involve a VP or a Predication Phrase (e.g., Bowers 1993, 2002). This representation is maintained throughout the dissertation (see especially Chapter 7). Nothing hinges on be being hosted in V; the critical point in the following chapter is that the nonfinite copula occurs in some projection below the light verb.

Chapters 4 and 5 that as an adjunct, LIKE targets maximal projections, and this is a position that I take up here as well. Consequently, in PredAP contexts, LIKE is assumed to adjoin to AP. In section 6.2, I will discuss structural evidence in favour of this analysis.

Second, when we consider strictly those structures in which LIKE intervenes between overt material in DP and an attributive adjective, as in (92), its overall frequency is relatively low: 4 percent (N = 672). Further, these 672 tokens represent just 13 percent of the 5140 DP tokens that were initially extracted. One could argue therefore that the omission of attributive structures was strategic: the investment of extracting sufficient quantities of this construction would net low returns.

- (92) a. They were all LIKE good friends for a really, really long time. (N/≤/m/26)
 b. Like everyone's in this LIKE torrid race to grow up. (N/E.m/36)

However, the infrequency of (92) coupled with the point at which these types of structures begin to appear with LIKE suggests a more compelling reason for excluding them: LIKE may be adjoined to NP, not AP. Regardless of whether an adjective is present or not, LIKE is first unambiguously attested “within” a DP among 30 year olds in the Toronto data. In other words, constructions such as a LIKE kid and a LIKE little kid appear simultaneously. Consequently, I have suggested that LIKE is hosted on NP in both types of constructions, since this is the adjunction site in a LIKE kid. I have also suggested that the low frequency of structures such as (92) reflects the recent evolution of the NP as a variable context for LIKE.

To summarize, I limit the analysis to PredAP constructions because when DP is lexically empty, the position of LIKE cannot be structurally disambiguated in attributive

constructions. In contrast, when DP is lexically filled, LIKE is both infrequent and assumed to adjoin to NP, not AP.

6.1 THE DATA

Based on the above discussion, the variable context is restricted to PredAP constructions. Unlike nouns, verbs, and clauses, which are fundamental elements of grammar and accordingly, occur with extreme frequency in discourse (e.g., Altenberg 1990), the use of predicate adjectives is susceptible to intraspeaker variation deriving from both social and stylistic factors. For example, in an hour-long interview, one 74 year-old male produces just 15 predicate adjectives and another man, aged 67, produces only 17. In contrast, for many of the speakers whose interviews were equally as long, the number of occurrences was much higher. This may reflect sex differences, but as the use of adjectives is greater in orientation and evaluation sequences of narrative structure (in the sense of Labov & Waletzky 1967) than elsewhere, it also may derive from the nature of the discourse. In order to minimize the effects of these individual patterns, a goal of 50 PredAPs per speaker was set. This resulted in a total of 4329 tokens.

From these, one construction was singled out and removed. This is the expression I'm not sure, as in (93).

- (93) a. I might have a grade twelve course here or something. I'm not sure. (3/B/m/17)
 b. I think a few people died. I'm not sure. (N/æ/m/51)

There were just 31 of these extracted from the interview materials, but they were excluded from the analysis because I'm not sure functions as a fixed pragmatic expression, similar to I

think and I guess (e.g., Brinton 2005, forthcoming). That is, it signals metalinguistic commentary on the previous statement. Further, LIKE occurred with none of these tokens. Since LIKE does intercede elements of idiomatic and lexicalized expressions (see Chapter 5, §5.1.1; also Andersen 2001:277-278; Wolgemuth 2003:47-48), its failure to occur in this string supports the decision to exclude it.

The remaining tokens, totaling nearly 4300, were then coded for speaker age and sex, as well as for the lexical verb and the position of LIKE within the PredAP. I will address these points in the results section, to which I now turn.

6.2 RESULTS

Table 6.1, which tracks the use of LIKE across apparent-time, reveals that this particle first appears with predicate adjectives among speakers in their fifties. Consequently, I set aside the data from the oldest speakers in the sample, the 60-87 year olds, and consider only the results for those between the ages of 10 and 59 (N = 3455).

An interesting trajectory is evident in these data. When LIKE first enters the PredAP context, its frequency is fairly constant: from no occurrences among the 60 year olds, it rises to 2 percent among the 50 year olds before stabilizing at 3 percent in the three successive age groups. Thus, during a period spanning approximately a quarter of a century, the proportion of PredAPs modified by LIKE remains level in apparent-time. At this point, its rate of use suddenly doubles, increasing to 6 percent among speakers in their early twenties. The frequency of LIKE then continues to rise until it peaks among the 15-16 year olds at 14

percent (N = 392). In short, the proportion of LIKE increases across apparent-time, gradually at first, but then a clear association with speakers under the age of 25 emerges.

TABLE 6.1
Distribution of LIKE with predicate adjectives across apparent-time

	%	N	total
10-12	9	474	} 3455 OD = 6%
15-16	14	392	
17-19	9	494	
20-24	6	471	
25-29	3	454	
30-39	3	490	
40-49	3	358	
50-59	2	322	
60-69	0	223	
70-79	0	271	
80+	0	349	
total		4298	

What is the mechanism behind this sudden and dramatic shift in the rate of LIKE-modification that occurs among speakers in their twenties? Andersen (2001:281) reports that copular be has “a special ‘triggering effect’” on LIKE and suggests that the combination of be + LIKE is almost formulaic. Interestingly, the vast majority of PredAP contexts occur with be (94%; N = 3455). The remaining data comprise arguments of verbs such as look, sound, and get (i.e., ‘become’), as in (94).

- (94) a. He had all this cream on it and it looked LIKE freaky. (2/e/f/11)
 b. It sounds LIKE mysterious. (2/h/m/18)
 c. If you make fun of him he’ll get LIKE really mad. (3/H/m/12)

Is the copula triggering the use of LIKE and driving up its frequency? To get a view of the collocation patterns of LIKE with verbs that subcategorize for an AP, we can track the distribution of each combination across apparent-time as in Figure 6.1.

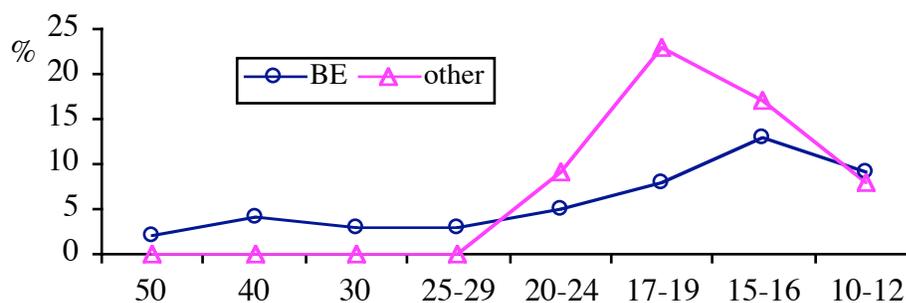


FIGURE 6.1 Distribution of LIKE in PredAP contexts according to verb

These results illustrate that when LIKE first enters the PredAP context, it co-occurs only with be. This seems to support Andersen's analysis. Beginning with the 20-24 year olds, however, LIKE starts to spread to other verbs, collapsed here as 'other'. This is the same group of speakers where fundamental increases in the proportion of LIKE become apparent. In other words, it is precisely at that point where LIKE generalizes beyond the copula that its frequency begins to rise quickly. This strongly suggests that one of the primary mechanisms behind the increase in the proportion of PredAPs modified by LIKE is its diffusion across a broader range of predicative constructions. Ito and Tagliamonte (2003:271,276) report a similar finding in the development of really as an intensifier: it first spreads across a wider selection of adjectives before increasing in frequency. Here, in Figure 6.1, any individual effects of expansion and frequency are difficult to disentangle, as the rate of LIKE jumps simultaneously with its spread beyond be. Nonetheless, these results graphically demonstrate why increases in overall rates of use are such important concomitants of

grammaticalization (Hopper 1991; Hopper & Traugott 2003; Bybee 2003; Bybee, Perkins and Pagliuca 1994), as they signal advancing development.

Thus, we can extrapolate from Figure 6.1 that, for the PredAP context, although be plays an important role in the initial stages of LIKE, it takes on a lesser role in the later stages. Among the first generations to use LIKE with PredAP constructions, the 25-50 year olds, the only variable context is following be. I would like to suggest that this is not accidental. Arguments of be are by far the most frequent token type, accounting for 94 percent of PredAP contexts overall. The early restriction to be is therefore analogous to the tendency for LIKE to modify argument DPs over complement DPs that we saw in Chapter 5. In the case of PredAP constructions, as the development of LIKE proceeds, the particle eventually generalizes to the arguments of other verbs. From this point forward, LIKE continues to occur with be and its frequency in this context continues to rise, but it is its use in other contexts that propels the overall frequency upward (see also Ito & Tagliamonte 2003).

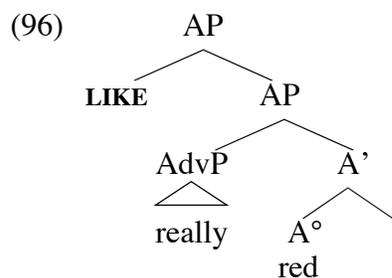
In contrast to Andersen (2001:281), who suggests that the raw frequency of be + LIKE collocations indicates a specialized role for the copula in the development of LIKE, the PredAP context provides no evidence that be continues to prompt the use of this particle. As the most frequent context overall, the copula provides the entry point for LIKE with predicate adjectives. However, once LIKE spreads beyond be, this verb takes on a lesser role as it is in the newer contexts that use of the particle predominates.

The question I would like to address at this point is whether there is support from the full range of PredAP constructions that LIKE adjoins to the left edge of the phrase. Since the

relative order of elements in a phrase can help disentangle structural factors (see also Pintzuk & Kroch 1989; Pintzuk 1993 et seq.; Santorini 1993; Taylor 1994), the most convincing evidence comes from APs that contain overt material in the specifier position, as in (95).

- (95) a. Sean would be **LIKE** really quiet and stuff. (I/8/m/32)
 b. I'd be **LIKE** pretty mad. (3/H/m/12)
 c. She's **LIKE** all surprised. (2/j/f/18)
 d. Everything is **LIKE** so complicated. (N/D/m/50)
 e. I was **LIKE** very scared. (I/3/f/24)

The data demonstrate that when the adjective is intensified (i.e., with intensives or downtoners, Quirk et al. 1985), **LIKE** occurs to the immediate left of the adverb. Given this syntagmatic order, we can conclude that **LIKE** adjoins to AP, as shown in (96):²



Of the 212 instances of **LIKE** that were extracted in the PredAP context, all but 13 are consistent with the analysis in (96). Some of the exceptions are given in (97).

- (97) a. They're very **LIKE** um, you-know, pacifist. (I/*/f/30)
 b. I get really, **LIKE**, flabbergasted. (I/3/f/24)
 c. My whole mouth was getting incredibly, **LIKE**, dry. (3/J/m/18)
 d. She was so, **LIKE**, touched. (2/a/f/16)

² An alternative analysis is that **LIKE** may modify the adverb, in which case I assume that it adjoins to AdvP. Regardless, **LIKE** targets the X^{max} level of structure (see fn.2, Chapter 8).

I have no explanation for these types of sequences at this time. It is interesting to note, however, that (97a) clearly contains a hesitation, and that LIKE in (97b-d) is surrounded by pauses. One possibility, therefore, is that LIKE in these exceptional cases is motivated by processing considerations rather than pragmatic factors (see, for example, Siegel 2002; cf. Miller & Weinert 1995). Although this issue is beyond the scope of this analysis, the data in (97) suggest that LIKE bears examination from a psycholinguistic perspective.

At this time I would like to turn to the role of sex in conditioning LIKE in PredAP constructions. Similar to what we saw in Chapter 5 for the DP context, Table 6.2 reveals that overall, LIKE is slightly more frequent among males than it is among females.

TABLE 6.2
Overall distribution of LIKE in PredAPs according to sex

males		females	
%	N	%	N
8	1641	5	1814

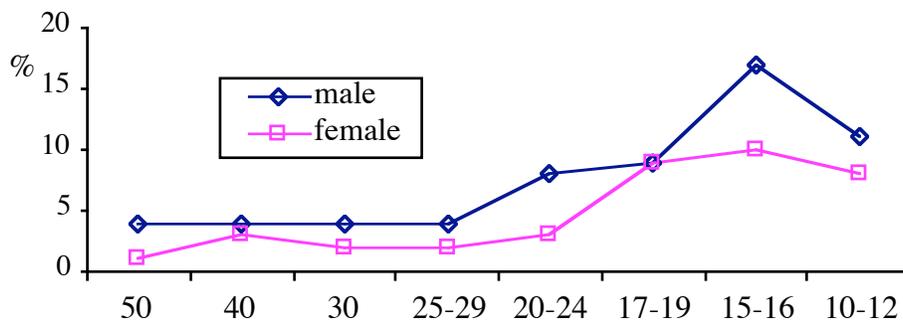


FIGURE 6.2 Cross-tabulation of sex by age in the PredAP context

When sex is cross-tabulated by age, as in Figure 6.2, the distributional pattern is reflected in all but one age group, the 17-19 year olds. In other words, the correlation with males is in

place from the outset: even among the first group to use LIKE in the PredAP context, the 50 year olds, the men are more advanced than the women. Further, note the 15-16 year olds. This is the age where the use of LIKE peaks (Table 6.1), and it is clearly the males who are responsible for the high rate within this cohort.

If we consider these results in light of the two previous chapters, the PredAP reflects the pattern observed in the DP context. In the CP context, where LIKE functions as a discourse marker, a strong female association is emerging. This effect is not apparent across the entire community, but is only visible among younger speakers where the frequency of LIKE is highest. In the DP and PredAP contexts, on the other hand, we find a consistent male lead, one that is evident from the establishment of LIKE in each context. I will not address this issue further here, but will return to it in Chapter 8, once we also have the results from the next chapter, the verbal domain, to consider.

To summarize, the results for the PredAP context to this point, the frequency of LIKE is increasing in apparent-time. Among speakers in their early twenties, however, the proportion of PredAPs modified by LIKE increases dramatically, an effect which I have shown to derive from the generalization of LIKE beyond copular constructions. Further, the analysis of speaker sex reveals a fundamental correlation with males. The question at this point is, how many of these results reflect significant constraints on the use and development of LIKE? Table 6.3 displays the results of a multivariate analysis of LIKE in the PredAP context: each factor considered here — age, sex, and verb — is selected by the stepwise procedure.

As in previous chapters, the speakers have been regrouped into four age categories. The older group, speakers above the age of 60, is included in the table strictly for purposes of exemplification. When configured in this way, the regular increase in the overall frequency of LIKE across apparent-time is plain: 0 > 3 > 6 > 11. Indeed, with a range of 22, the multivariate analysis reveals that age is the strongest of the factors considered here.

With ranges of 13, the two remaining factors are equally strong: sex and verb type. Although the difference in the overall frequency of LIKE is not particularly robust across the sexes, the consistency of the effect across the sample suggested that the male association reflects a true social constraint. This is confirmed here by the selection of sex as a significant factor in the use of LIKE.

TABLE 6.3
Contribution of external and internal factors on the probability of LIKE in the PredAP context

INPUT factors considered	FW	.06 %	N
Verb			
other	.62	14	115
copula	.49	9	1716
range	13		
Age			
10-16	.63	11	866
17-29	.50	6	1419
30-59	.41	3	1170
60+	—	K/O [0%]	843
range	22		
Sex			
male	.57	8	1641
female	.44	5	1814
range	13		
TOTAL N			3455

The identity of the lexical verb was the sole internal factor to be tested. For this factor, only the results for the age groups where LIKE has diffused beyond be, the 10-24 year olds, were included. The data for the speakers aged 25-50 were omitted, since there is no variability with regard to the verb within these groups: LIKE is restricted to the copula and has not expanded to a broader range of PredAP constructions.

The results are unequivocal. Not only is LIKE favoured in the context of verbs other than the copula, this effect is significant. In comparison, the copula exerts little effect, hovering near .50. Thus, while the collocation of be + LIKE was a “driving force” (Andersen 2001:281) behind the grammaticalization of this discourse particle during the initial stages of its development in the PredAP context, it no longer fills this role. In fact, further evidence of the weakening role of be in predicate adjective constructions comes from the interface between external and internal factors. Consider Figure 6.3, which plots the frequency of LIKE according the sex and age of the speaker and the lexical identity of the verb.

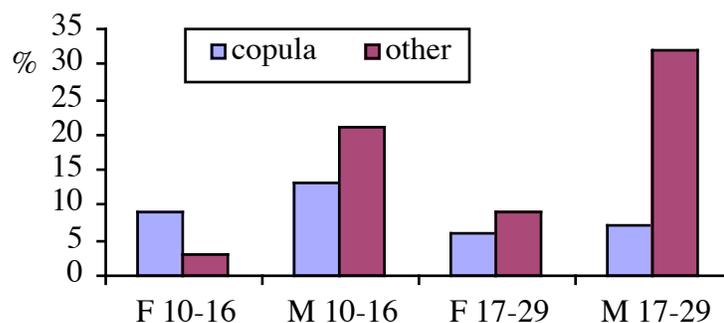


FIGURE 6.3 Cross-tabulation of sex by verb in the PredAP context

From the moment LIKE enters the PredAP context, males are in the lead, consistently using this particle more frequently than their female peers, regardless of age. Here in Figure 6.3 we can see that they are also responsible for the expansion of LIKE beyond be. In both age groups, the difference in the rate of LIKE according to the lexical verb is greater among males than it is among females. Indeed, the younger females reflect the older pattern, that where LIKE is more frequent with be. Thus, they lag behind the males, who are clearly at the forefront of shifting towards the use of LIKE in non-copular contexts. In other words, males are the innovators.

6.3 DISCUSSION

Even though adjectives do not belong to the “Group of Three,” the contexts that have consistently been reported to constitute the most frequent slots for LIKE in terms of token frequency (Underhill 1988; Andersen 1997 et seq.; Wolgemuth 2003; Levey 2004), I elected to examine the AP because the analysis of the DP in Chapter 4 implicated these structures in

the ongoing development of LIKE. However, for purposes of exposition I concentrated on predicative structures rather than confounding the analysis with attributive contexts.

Nevertheless, these other contexts cannot simply be ignored. The apparent-time view of PredAPs in Table 6.1 indicates that it was established as a variable context for LIKE among speakers currently in their fifties. This is not the same group of speakers who were the first to use LIKE in contexts such as (98a), nor are they the same group to use it first in contexts such as (98b).

- (98) a. They are **LIKE** red balls.
 b. They are the **LIKE** red balls.

In fact, if we trace a time-line for LIKE in the various contexts in which, in terms of the linear order of syntactic elements, it precedes an adjective, as in Figure 6.4, we observe a perplexing pattern where LIKE enters each construction type in what appears to be discrete stages.



FIGURE 6.4 Time-line for LIKE in adjectival contexts

Structures like (98a) were excluded from both the analysis of the DP as well as that of the AP because the position of LIKE cannot be structurally disambiguated. That this context first appears within a decade of LIKE in PredAP contexts further supports these decisions,

since the chronology allows for AP adjunction but the syntax cannot rule out DP adjunction. Consequently, the AP context can be considered as having developed at some point between the 50 and 60 year olds.

Structures like (98b) were excluded because their first attestation in the Toronto data appears concurrently with structures in which LIKE unambiguously adjoins to NP. That the AP was already well established as an adjunction site for LIKE by the time (98b) occurs supports the analysis that LIKE adjoins to NP in this type of construction. However, we cannot rule out that LIKE never adjoins to AP in (98b). Given this possibility, the question is, why do these constructions not appear until later? The AP is an older layer than the NP in the development of LIKE. If this particle simply adjoins to the phrase, there is no structural reason why it could not appear in (98b) from the same point at which it appears in (98a), assuming some proportion of these represent AP adjunction, or in predicative constructions. This is an intriguing issue, one I will leave open for now.

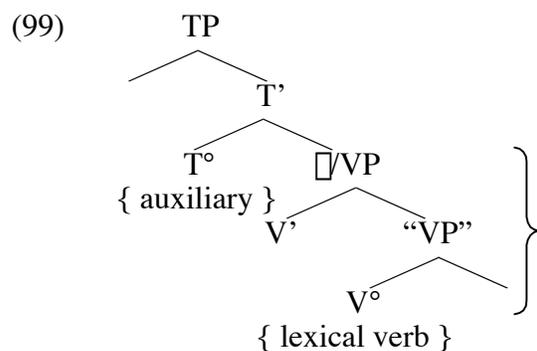
In summary, the results presented in this chapter indicate that the overall proportion of LIKE in PredAP constructions is increasing in apparent-time, and further, that an important concomitant of this rise in frequency is the spread of LIKE beyond copular be to other verbs that subcategorize for an AP. This finding corroborates previous research demonstrating the role of generalization in grammaticalization processes (e.g., Bybee 2003; Ito & Tagliamonte 2003). It also reveals that although the copula was instrumental in the initial stages of the generalization of LIKE to PredAPs, probably because it represents the most frequent token type, this role has been usurped among younger speakers where LIKE is significantly favoured with verbs other than be. Finally, this developmental pathway is tightly entwined

with social factors, and with speaker sex in particular. Males not only use LIKE significantly more often than females do, but it is they who are advancing the spread of this particle to new constructions within the PredAP context.

Chapter 7
THE VERBAL DOMAIN

The preverbal position is often cited as one of the three most frequent slots for discourse LIKE. When the text frequency of this particle is calculated, the VP has been ranked first (Underhill 1988), second, falling behind noun phrases (Andersen 1997 et seq.), or third, where both noun phrases and the clause-initial position outrank it (Wolgemuth 2003; Levey 2004; Tagliamonte 2005). The most exhaustive analysis is that of Andersen (2001), and as such, it is used as a reference point for most of the work I present here. Due to the perceived status of the verbal domain as a preferential location for LIKE, this context forms the focus of the current chapter, in which I challenge the received wisdom in the literature.

As I will outline below, a number of observations have been made regarding the position of LIKE when it co-occurs with verbs. Hereafter, when I speak of the VP context, I use ‘VP’ to refer broadly to the structures of [□P VP] and any projections such as aspectual categories that may intervene between them and the head of TP. This is shown in (99).



This structural approach to the analysis of LIKE reveals that its position in the VP falls out from syntactic factors. In other words, the slot filled by this particle is consistent across the full range of predicate structures and is by no means random. In addition to these structural considerations, I am able to demonstrate the effects of certain semantic constraints on the use of LIKE.

7.1 METHODOLOGY

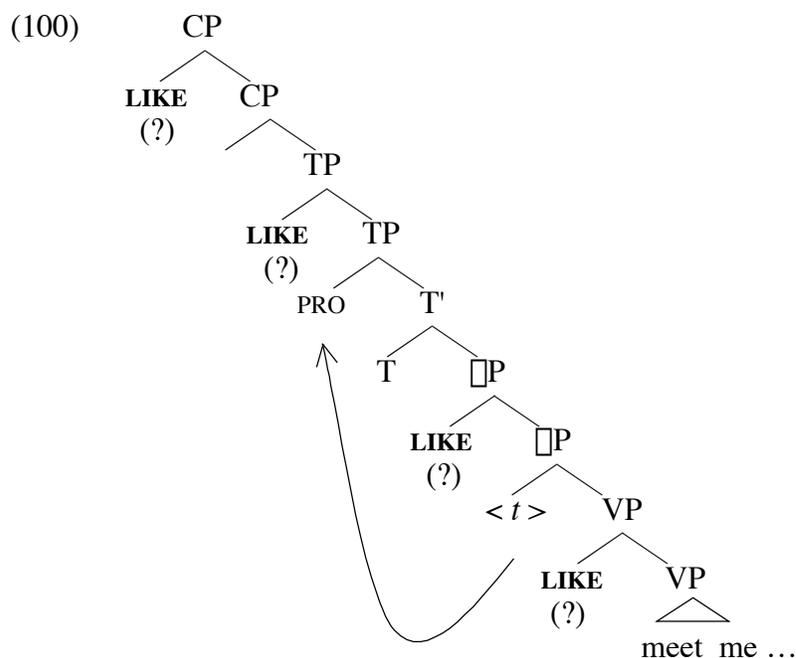
Previous research reporting the high number of LIKE tokens in the verbal domain has been based on the raw number of occurrences of this particle. Given the frequency of verbs in language in general (e.g., Altenberg 1990:185), these results are difficult to put in context. What do they tell us about LIKE? Is there a reason why it may be preferentially located before verbs? Further, are all verbs equal in terms of receptivity to LIKE? In order to address questions such as these, I will apply accountable methodology as in previous chapters to determine whether there is any systematicity to where LIKE occurs as opposed to where it does not. In this section, I detail the methodological considerations that form the basis of the analysis.

7.1.1 The variable context

The context for this analysis is restricted to the VP of declarative sentences. There are three reasons for this. First, recall from Chapter 4 (the CP) that interrogatives and imperatives are rare in the Toronto English materials. Second, of the interrogatives that were initially extracted in the analysis of the CP context, none contained LIKE within the clause. That is,

although LIKE occurs on the left periphery of interrogative CPs, it never occurs in any other position in these types of clauses. Whether this is due to low Ns ($N = 45$) or some other factor is unclear. Nonetheless, due to their apparently categorical status, they are not included in the analysis presented here.

The third cause for the declarative focus concerns the position of LIKE in imperatives. Due to a number of factors, we cannot disambiguate whether LIKE adjoins to CP, TP, or \square /VP in sentences such as ‘LIKE, meet me behind the portables’. For example, the lack of a complementizer means that we cannot tell if LIKE is to the right or left of CP. Because the subject is null, we cannot tell if LIKE is to the right or the left of TP. Further, because there is no auxiliary, whether LIKE is to the right or the left of \square P cannot be determined. These facts are demonstrated in (100).¹



¹ Note that as in previous chapters, I collapse nodes when their internal structure is incidental to my point.

Consequently, affirmative imperatives are removed from consideration on structural grounds (see §1.4.2, Chapter 1). Negative imperatives must also be excluded, but for a different reason. Consider the data in (101).

- (101) a. **LIKE** don't worry about it. (2/i/f/19)
 b. **LIKE** don't believe them and stuff. (2/b/f/15)

Following Pollock (1989), I assume that auxiliaries, modals and infinitival 'to' are hosted outside the light verb complex in the head of TP. This includes the dummy auxiliary 'do' that is inserted in negative constructions. Regardless of where these inflectional elements are base-generated, at PF they are higher in the syntax than the lexical verb, which is in V°. In imperatives, though, 'do' raises to C° (e.g., Han 2000). Since **LIKE** is before this element in (101), it will be assumed that in imperatives, it is situated outside the VP context, adjoined to the higher CP projection.²

In order to determine the ways in which **LIKE** is distributed within declarative VPs, it was necessary to have access to a wide range of phrases. Consequently, 75 tokens were extracted per speaker, whether the VP contained **LIKE** or not. This method revealed that not all verbal constructions constitute part of the variable context for **LIKE**. In other words, there are certain VPs in these data in which **LIKE** never occurs.

² There were very few negative imperatives extracted from the Toronto materials (N = 4). As such, the evidence in (101) probably does not represent an invariant pattern, and there is no reason to postulate that **LIKE** could not intervene between don't and the lexical verb in these constructions. As this chapter will establish, all that is required structurally for **LIKE** is the presence of a light verb.

The finite inflected copula, as in (102), is never preceded by LIKE (0%; N = 482).³

- (102) a. They Ø are just on my bed. (2/g/m/11)
 cf. *They LIKE are just on my bed.
- b. He Ø was so happy to take a bath. (2/b/f/15)
 cf. *He LIKE was so happy to take a bath.
- c. There Ø is a light at the end of the tunnel. (3/K/f/18)
 cf. *There LIKE is a light at the end of the tunnel.

In contrast, LIKE does occur in nonfinite copular constructions, as in (103), although its overall rate is low (3%; N = 86).

- (103) a. So then it was cool ‘cause you get to LIKE be smart. (2/u/m/15)
 b. So it went from LIKE being like that, to like that. (2/h/m/18)

Both the copula contexts were ultimately excluded from the distributional and multivariate analyses, but I return to them in section 7.2.1 where I offer an explanation for both patterns.

A further context that does not seem to readily admit LIKE is the passive voice, shown in (104). Here, the overall distribution of the particle is 4 percent (N = 67). Following Guy (1988:132), these tokens were excluded from the analysis.

- (104) a. I was LIKE threatened to get beat up. (ITM/m/40)
 b. All the Chinese school was LIKE packed in the room. (2/d/m/11)

Finally, LIKE occurs extremely infrequently with the perfective aspect (2%; N = 135), (105), and these too were excluded.

³ Unless stated otherwise, I report distributions and total Ns based on the results for those age groups who use LIKE in this context. Data from those whose VPs do not contain LIKE (see §7.2.1, Table 7.2) are not included in these calculations.

- (105) a. I've **LIKE** lived here like my whole life. (3/S/m/18)
 b. And they had **LIKE** scraped her. (I/7/m/35)

Why should perfectives distribute in this way? One possible explanation rests in their status as characteristic of more formal styles, and with written registers in particular (Elsness 1997). In this view, constructions representing more formal registers, where discourse features are highly stigmatized (e.g., Quirk 1972; Evan-Zohar 1982; Östman 1982; Erman 1987; Schiffrin 1987), may be resistant to LIKE.

Circumscribing the variable context in this way resulted in a total of 5483 VPs (all speakers), which were subsequently coded for speaker age and sex, as well as for the type of surface string and the verb type. I will now outline in detail these last two aspects of the methodology.

7.1.2 Coding and analysis

Surface string

Despite the popular belief that LIKE can occur grammatically anywhere (e.g., Siegel 2002:64), constraints on its distributional properties in VPs have been fairly well documented. For example, Underhill (1988:243) reported that when LIKE is used in the verbal domain, it surfaces in the following three positions:

- 1) between an auxiliary and the lexical verb
- 2) between infinitival 'to' and the lexical verb
- 3) otherwise, to the left of the lexical verb

Andersen (2001:280-281), working with a much larger database, reports identical patterns.

To test these claims quantitatively, the data were coded for the type of surface string. These include a wide range of structures, including finite verbs, in which the head of TP is empty (i.e., not lexically-specified), as in (106a), participles (106b), auxiliary constructions (106c), infinitival ‘to’ constructions (106d), as well modal and semi-modal constructions (106e,f). Due to the position of LIKE in this last context, the semi-modals were recoded with infinitival constructions.

- (106) a. I’m not sure if my eight year old LIKE understands that. (N/X/m/46)
 b. I’ve seen a lot of people just LIKE walking down through Yorkville. (N/m/m/27)
 c. I was LIKE playing in bands like all the way through high-school. (I/J/m/22)
 d. They still manage to LIKE wave at you. (I/&/f/21)
 e. For instance, uh she would LIKE call twenty-four-seven. (3/C/m/15)
 f. So I had to LIKE pull it off my ankle. (2/e/f/11)

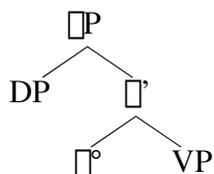
Verb type

The effect of verb type was tested because in many languages, including English, verbs can be differentiated on both semantic and syntactic grounds (e.g., Perlmutter 1978; Hoekstra 1984; Burzio 1986; Hale & Keyser 1993, 2002; Chomsky 1995). This distinction divides the monadic, intransitive verbs into two classes, unergatives and unaccusatives, and groups unergatives together with transitives. Prototypical unergative verbs are dance, laugh, and run; prototypical unaccusatives are break, arrive, and fall.

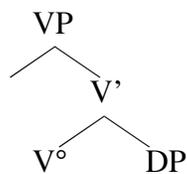
Semantically, the subject of an unergative clause is usually an agent, while that of an unaccusative is typically a patient or a theme. These roles can be summarized as follows. Agents cause some action to occur through their own volition (e.g., she danced). Patients and themes undergo the action or event denoted by the predicate (e.g., she fell).

Syntactically, the subject of unergative and transitive verbs is argued to be merged external to the VP, as the specifier of the light verb, as in (107a). The argument of an unaccusative, however, is merged internally, as the complement of VP, as in (107b).⁴

(107) a. unergative/transitive



b. unaccusative



The structure in (107b) does not indicate a \bar{P} layer. There are two schools of thought regarding the status of the light verb in unaccusative structures. One view is that unaccusatives lack this projection entirely. This has been argued for by both Hale and Keyser (1993, 2002) and Chomsky (1995). The opposite view posits a \bar{P} for all verbs (e.g., Levin & Rappaport Hovav 1995; Marantz 1997; Harley & Noyer 1998; Alexiadou 1999; Massam 2001). I assume the second line of argumentation here.

Since verb type has both syntactic and semantic correlates, a finding that LIKE is distributed differently across these categories would suggest that LIKE is constrained grammatically. The prediction would be that the use of LIKE will differ across the monadic verbs, differentiating unergatives and unaccusatives, and that unergatives will pattern together with transitives. Consequently, verbs were coded according to whether they were unergative (108a), unaccusative (108b), or transitive (108c).

⁴ This DP then raises to the subject position to be assigned Case.

- (108) a. And she's **LIKE** praying, praying, praying. (I/≠/f/32)
 b. So I was **LIKE** buzzing from that. (N/I/m/24)
 c. The kid was **LIKE** pouring water on him. (I/~f/29)

Of course, there are verbs whose status as either unergative ~ transitive or unaccusative ~ transitive is variable (e.g., eat, play, break, burn, etc.). These were coded following their categorization in Levin (1993), relying on the context in which the verb occurred. For example, play in (109a) was treated as a transitive verb, while play in (109b) was treated as unergative.

- (109) a. We were **LIKE** playing this weird game. (2/e/f/11)
 b. I was **LIKE** playing outside. (I/§/f/21)

TABLE 7.1
 Criteria used to determine the class of a monadic verb

UNERGATIVE	<ul style="list-style-type: none"> • Cognate object (Pesetsky 1982) • Derivational suffix -er (Burzio 1986) 	<p>dream a dream; sing a song *arrive an arrival; *vanish a vanish</p> <p>jumper, singer, teacher, bleeder *ariser, *goer, *faller, *dier</p>
UNACCUSATIVE	<ul style="list-style-type: none"> • Causative alternation (Levin & Rappaport Hovav 1995) • Locative inversion (Coopmans 1989; Hoekstra & Mulder 1990) • Resultative construction (Levin & Rappaport Hovav 1995) • Past participles > adjectives (Levin & Rappaport 1986) • Derivational suffixes -ee, -able, -re, -un (Horn 1980) 	<p>The toy broke / the child broke the toy. *The toy fell / the child fell the toy.</p> <p>Out of the market came my mother. *In the car talked my parents.</p> <p>The river froze solid. *We cried dry. (cf. cried ourselves dry)</p> <p>frozen lake; withered leaves *jumped child; *slept woman</p> <p>escapee, shrinkable, unfreeze, reerupt *sneeze; *jumpable; *unlaugh; *resmile</p>

Finally, the class of a monadic verb as either unergative or unaccusative was determined according to the criteria outlined in Table 7.1. As is typical of linguistic phenomena, these diagnostics do not apply across the board (Levin 1993). They simply present guidelines to distinguish among intransitives. The theta role of the subject was also considered, though the syntactic diagnostics took precedence over the semantic one.

I turn now to the results.

7.2 RESULTS

7.2.1 Distributional analysis

Table 7.2 charts the distribution of LIKE in the VP context according to age. These results suggest that the expansion of LIKE to the VP context has occurred relatively recently, since just as we saw in the last chapter, LIKE does not begin to appear until the 50 year olds. These results also show that once established in the VP, its frequency increases in apparent-time. However, whereas this increase is incremental among 25-59 year olds, a different pattern is evident among the younger speakers. Beginning with the 20-24 year olds, the frequency of LIKE jumps first to 10 percent, and then to 15 percent among the 17-19 year olds. Within the two youngest age groups, its use remains steady, though it is still high relative to those in their late twenties and beyond. Does this sharp stratification represent age-grading or some other factor? This is a question I return to in the following chapter.

TABLE 7.2
Distribution of LIKE in the VP context over apparent-time

	%	N	total
10-12	10	687	} 4389 OD = 8%
15-16	10	555	
17-19	15	598	
20-24	10	648	
25-29	4	534	
30-39	3	545	
40-49	2	407	
50-59	1	416	
60-69	0	394	
70-79	0	325	
80+	0	374	
total		5483	

In view of this time-line, when Underhill (1988) writes that VPs contain the largest number of LIKE tokens, he must have been reporting on data from speakers currently in their thirties. The results in Table 7.2, however, indicate that speakers in this age group, in Toronto at least, rarely use LIKE before verbs.

Is the appearance of the VP as a common location for LIKE in type/token analyses a direct result of the high frequency with which verbs occur? This is what I had suggested in Chapter 1 (§1.4.2), and the apparent-time results confirm this. Consider Figure 7.1, which displays the frequency of LIKE across the Toronto English sample in both PredAP constructions, the topic of the last chapter, and in the VP.

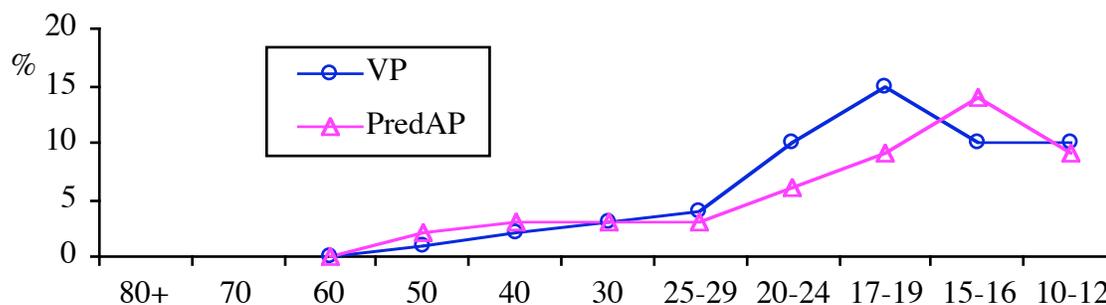


FIGURE 7.1 Comparison of the distribution of LIKE in the PredAP and VP contexts

These contexts provide an ideal opportunity to test previous claims regarding the frequency of LIKE, because the VP is believed to be a favoured position for this particle, while APs in general are not (e.g., Andersen 2001; Levey 2004; Wolgemuth 2003; etc.). The results in Figure 7.1 demonstrate that the distribution of LIKE in the VP context parallels its distribution in PredAP constructions. There is very little divergence. In fact, among Underhill's (1988) age group, the 30 year olds, the proportions of LIKE in these two contexts are identical. Moreover, when the overall distribution is calculated for the 10-59 year olds, the results are again nearly identical: LIKE accounts for 8 percent of the data in the VP (N = 4389) vs. 6 percent in PredAPs (N = 3455). Thus, a comparison of the proportional results for these two contexts refutes the claim that the VP presents a preferential location for LIKE (e.g., Underhill 1988; Romaine & Lange 1991; Andersen 2001).

In contrast, consider the results from Andersen (2001:276). Adjectives appear to present an infrequent context since they account for only 8 percent of the total occurrences of LIKE, compared to verbal contexts, which account for 28 percent of tokens in the COLT materials (N = 505). However, adjectives, and predicate adjectives in particular, occur much less often than verbs do, and the sheer rate at which the latter occur provide increased

opportunities to use the particle. This does not mean that VPs are preferred over PredAPs, but rather, that their raw numbers are simply greater.⁵ In sum, even though the type frequency of LIKE + verb is higher than that of LIKE + adjective, comparison of each context proportionally reveals striking evidence that little difference exists between the two.

Establishing that the VP is not a preferential location for LIKE is not the end of the story. When it is used, where exactly does it appear? As reported in Underhill (1988:64) and Andersen (2001:280-281), the syntagmatic position of LIKE is highly fixed. It cannot take “any” position within the VP, but categorically occurs to the immediate left of the lexical verb. Thus, when functional morphemes such as modals, auxiliaries, or infinitival ‘to’ are present, LIKE intervenes between these and the element hosted in V°. Each of these is exemplified in (110):

⁵ Nevertheless, the results in Figure 7.1 indicate that among speakers between the ages of 17 and 24, LIKE does occur more frequently within the VP than it does within PredAP contexts. It is unclear whether this reflects data fluctuation or some other factor. Regardless, this does not impinge on my argument here. Andersen’s (1997 et seq.) data are from speakers between the ages of 13 and 17, speakers who are currently aged 24-28 years old. Within this group of speakers in the Toronto data, the two contexts are virtually level. Among the 10-16 year olds, ages that overlap those of the COLT corpus at the time the data were collected, the VP does not contain a higher proportion of LIKE overall. In fact, within the adolescents of this cohort, the 15-16 year olds, LIKE is more frequent in PredAP constructions than it is within the VP.

(110) Position of LIKE in the VP:

between modal and the lexical verb

- a. I'd **LIKE** wake up and feel good. (2/u/m/15)
- b. They'll **LIKE** stack them for me. (2/b/f/15)

between auxiliary and the lexical verb

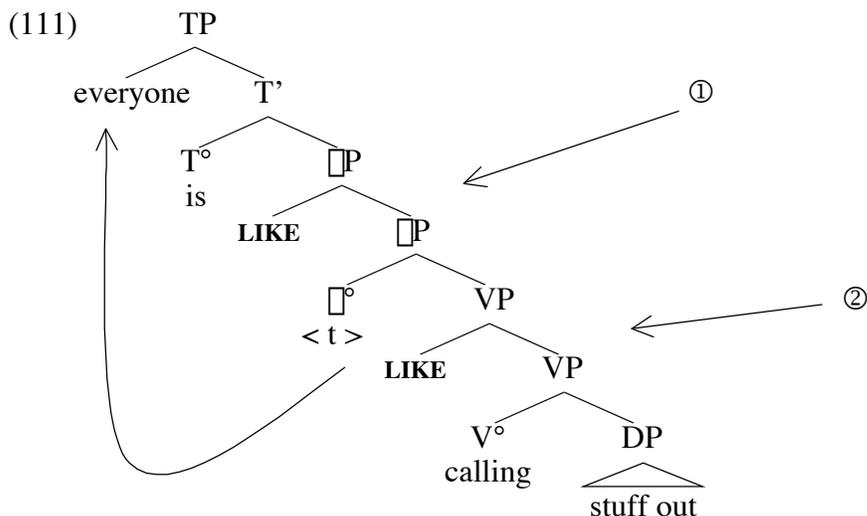
- c. I've **LIKE** grown into that. (3/H/m/12)
- d. Everyone is **LIKE** calling stuff out. (1;/m/22)

between infinitival 'to' and the lexical verb

- e. They like to **LIKE** intervene a lot. (3/J/m/18)
- f. It's a lot more easier to **LIKE** work out like my own schedule. (2/m/f/12)
- g. So I had to **LIKE** pull it off my ankle. (2/e/f/11)

Is this position accidental, or can it be explained on structural grounds? Andersen (2001:276-277) has argued that the type of phrase significantly conditions the probability that LIKE will occur within a phrase ($p < .001$). In Chapters 4, 5 and 6, I have suggested that it is the projection to which LIKE adjoins that determines where this particle surfaces. I continue this line of argumentation here.

The structure in (111) is that of the sentence in (110d); the CP is omitted.



Assuming that LIKE adjoins to the X^{\max} level, as I have argued throughout this dissertation, there are two possible positions where it could adjoin in the VP context: $\square P$ or VP.⁶ This is shown in (111) by the arrows labeled 1 and 2. I have not found any structural diagnostics for a distinction between these two sites, yet there is reason to consider that LIKE targets the light verb (position 1) and not the lexical verb (position 2).

As a discourse feature, LIKE first appeared on the left periphery of CP, a functional category. As seen in Chapter 5, it next generalized to DP, another functional category. The expansion of this particle to lexical categories, namely NPs and APs, were later developments. Further, the “DP to NP” shift seen in Chapter 5 suggests that within a constituent, LIKE must first be able to adjoin to the functional projection before the ‘deeper’ lexical category is a viable site. Consequently, I assume that when LIKE appears in the VP context, it is situated on the left edge of $\square P$. I will return to this point below.

The crucial point for now is that, using Andersen’s (2001) terminology, LIKE occurs “within” the VP because of the structure it targets, $\square P$. The lexical verb is situated below this level, in V° . Modals, auxiliaries, and infinitival ‘to’, on the other hand, appear in the head of TP (Pollock 1989), above both the lexical verb and the light verb.

This positioning supports the hypothesis that LIKE targets maximal projections. If LIKE could adjoin to some other level (e.g., X'), then we should find evidence of it intervening between the subject DP and T° , but we do not. In the VP context, LIKE never precedes

⁶ In assuming just two possible adjunction sites, I collapse $\square P$ and any other functional projections (e.g., Agr, Voice, etc.) that may alternatively be the locus of adjunction.

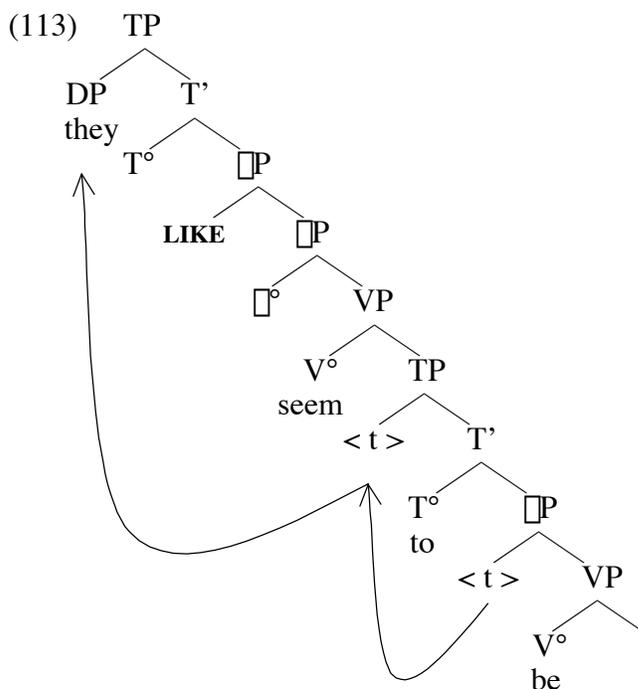
material hosted in T° (see also Andersen 2001:280). In sum, we can assume that it must appear to the left of whatever form is in V° by simple structural facts.

With this in mind, Andersen (2001:280-281) states that in a VP “whose tensed verb is a primary or modal auxiliary,” the auxiliary must precede LIKE. As just discussed, this follows from the $\square P$ analysis. However, Andersen then continues to state that this restriction does not apply when the tensed verb is seem, wanna, and start as in (112), his (203)-(205).

- (112) a. They always **LIKE** seem to be impressing everyone.
 b. They **LIKE** wanna see like how we talk and all that.
 c. [...] when I **LIKE** start talking to Jenny and stuff.

Let us consider first the raising verb, seem. The structure of (112a) corresponds to

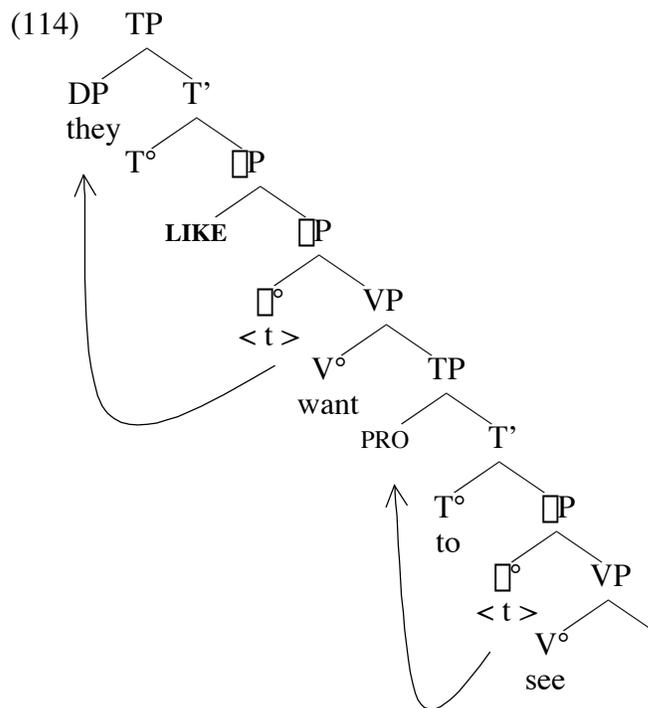
(113).⁷



⁷ For now I set aside the adverb as it forms part of a separate discussion, dealt with in section 7.2.3 below.

Example (113) contains two light verb projections. That LIKE precedes seem in Andersen's example means that it adjoins to the light verb of the main clause instead of targeting that of the nonfinite lower one. The current materials contain few raising structures in which the complement contains an infinitival predicate (N = 5), and so I can offer no examples of LIKE on the lower \square P in these types of structures. This possibility is not ruled out, however, by structural considerations. As such, it is unlikely that the position of LIKE in (112a)/(113) represents a selectional restriction on the particle, since the lower clause also contains a light verb. In other words, this example is predicted by the \square P analysis.

Support for this is also provided by control verbs such as want. The crucial difference between raising and control verbs is that in a control sentence, the subject is an argument of the main clause verb. This difference, however, has no bearing on LIKE; as shown in (114), corresponding to (112b), it continues to adjoin to the light verb.



Moreover, contra Andersen's data, LIKE can adjoin to either light verb in control structures: that of the main clause predicate, as in (115a), or that of the lower clause predicate, as in (115b). The data in (115c,d) demonstrate that this is not restricted to want, but holds of other control verbs as well. In fact, the structural analysis argued for here predicts that as long as a $\square P$ is present in the syntax, LIKE can target that projection. (115d) confirms this: here LIKE adjoins to both available light verbs.

- (115) a. I didn't want to LIKE walk up to them. (2/b/f/15)
 b. They LIKE want to get together. (3/Q/f/16)
 c. You're trying to LIKE pull it out of the water. (3/F/f/17)
 d. As long as they LIKE try to LIKE merge with Canadian culture. (1/J/m/22)

Now, what of Andersen's third exception, start, from (112c)? In the COLT data, LIKE precedes start. Of the 38 tokens of this type that were extracted from the Toronto interview materials, LIKE only occurs before the participle, never before the inceptive. This is exemplified in (116).

- (116) a. They started LIKE jumping around. (N/d/m/26)
 b. They kept LIKE trying to carry on a conversation. (3/T/f/18)

These types of structures have received little attention in the syntactic literature, but it seems that at least two analyses of (112c) and (116) are possible. If these structures are bi-clausal, then we again predict two positions for LIKE: on the light verb of the main clause and on that of the embedded clause. The higher position is corroborated by Andersen's example in (112c), and the lower position is corroborated by the Toronto examples in (116). If, on the

other hand, these structures are mono-clausal and start is a modal verb, then we can predict the data in (116), but Andersen's data remains a problem.

Nonetheless, the \square P analysis is also able to capture the contrast between the categorical failure of LIKE to precede the finite inflected copula and its ability to occur in nonfinite copular contexts, discussed in section 7.1.1 above, examples (102) and (103). In the surface structure, finite be is situated in T° , above the adjunction site for LIKE.⁸ In nonfinite contexts, however, be remains in V° (e.g., Becker 2004) and thus the particle may precede the copula in these instances. In short, (117a) is ruled out on structural grounds, while (117b) is not (see also Smith 2000 for syntactic explanations for categorical vs. variable contexts).

- (117) a. *They **LIKE** are just on my bed.
 b. You get to **LIKE** be smart.

To summarize, I have shown that it is the projection targeted by LIKE that determines its syntagmatic position within the VP. Thus, the fact that this particle appears to the immediate left of the lexical verb is not because LIKE is attracted to lexical material, but because it adjoins to the light verb. Moreover, I have argued that it is this syntactic constraint that accounts for the failure of LIKE to precede finite inflected be, since this form is in the head of TP, above the adjunction site for LIKE.

Having established the position of LIKE in the VP and that the only structural requirement for LIKE is the presence of a light verb, an interesting question is whether the

⁸ Whether be is merged directly in T° , as argued for by Becker (2004), or in V° and then raised to T° , as is traditionally held (e.g., Roberts 1998), is incidental to the facts presented here.

proportion of this particle is stable across different types of verbal complexes. If we first consider the overall distribution of LIKE according to the surface strings exemplified earlier in (106), as in Table 7.3, we see that there is a great deal of variability.

TABLE 7.3
Distribution of LIKE across different surface string types in the VP

	%	N
participle	15	188
auxiliary	14	768
infinitival 'to'	11	421
modal	6	496
finite verb	5	2254
'do' support	3	262
total		4389

Participial, auxiliary, and infinitival 'to' constructions are the three most frequent contexts for LIKE. Finite verbs barely present a variable context at all with an overall distribution of just 5 percent. These results cannot be interpreted as indicating that the proportion of LIKE is greatest when T° is lexically filled, because modal and 'do' support contexts pattern alongside bare verbs.

In fact, the inclusion of 'do' support requires comment, since LIKE is infrequent in this context (3%) and following standard variationist practice, would normally have been excluded from the multivariate analysis that follows (e.g., Guy 1988:132). As we will see in Figure 7.2, which tracks the use of LIKE in each string type across apparent-time, there is good reason for keeping this context in.

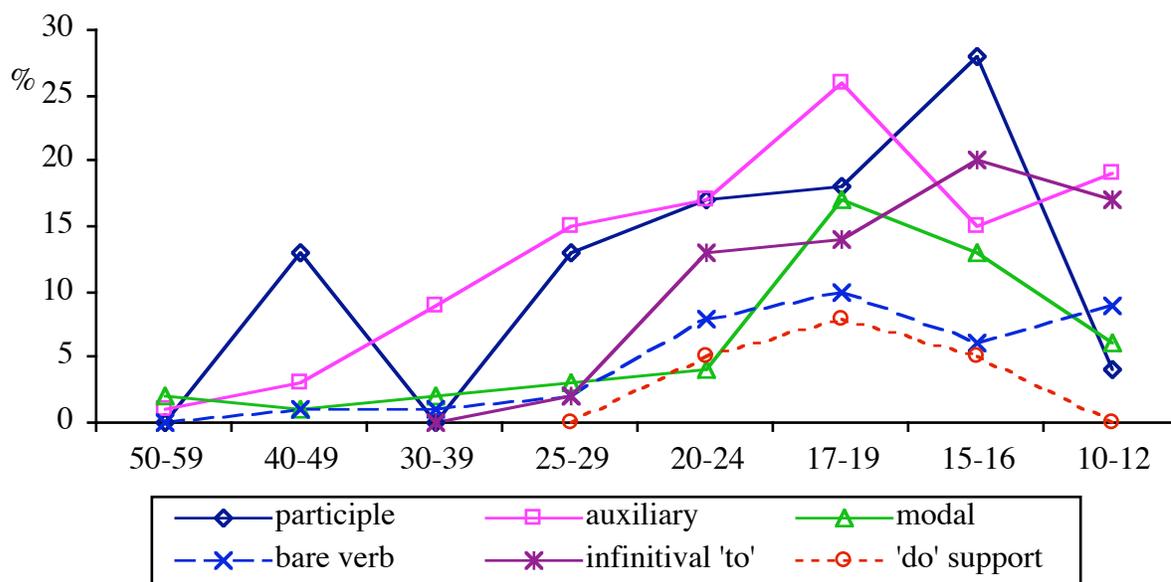


FIGURE 7.2 Type of surface string in which LIKE occurs according to age

Viewing the data from this perspective reveals that LIKE has not always been used in ‘do’ support contexts, typically negative declaratives (N = 248) but also including ‘emphatic do’ (N = 14): it first appears in these constructions beginning with speakers in their early twenties. For the older speakers, this context is invariant. However, among the 15-24 year olds, the proportion of LIKE in this context is 6 percent (N = 126). Consequently, ‘do’ support verbs have not been excluded because they provide a clear example of the development of LIKE within the VP. They are not a nearly categorical context; they are a new context.⁹

⁹ These can be contrasted with passive constructions, which with an overall distribution of 4 percent (N = 67), were excluded from the analysis (§7.1.1). This is because when cross-tabulated by age, LIKE occurs with the passive voice within the oldest generations of VP users, indicating that the spread of LIKE to these constructions is not recent, but simply one where the use of the particle is rare. I return to the issue of passives later in this section.

Figure 7.2 also reveals that LIKE first appeared in auxiliary and modal complexes before later spreading to bare verbs (the 40-49 year olds) and then to infinitival ‘to’ constructions (the 25-29 year olds).¹⁰ The spread of LIKE to ‘do’ support contexts appears to represent the most recent stage of development. It is this contextual generalization that is in part responsible for the ongoing increase in the overall frequency of LIKE that we saw in both Table 7.2 and Figure 7.1 (see, for example, Hopper & Traugott 2003; Ito & Tagliamonte 2003; Bybee 2003).

I have suggested that in each of these string types, LIKE adjoins to the light verb. If, however, following Hale and Keyser (1993, 2002) and Chomsky (1995), unaccusative verbs lack a \square P projection, then we would not expect this particle to occur with verbs like break, arrive, and fall; they lack an appropriate adjunction site. As can be seen in Table 7.4, however, this is not the case. Although infrequent, LIKE does occur with unaccusative verbs. Further, if a syntactic and/or semantic constraint is operative on LIKE, then unergatives and transitives should pattern together. This is not evident either.

TABLE 7.4
Distribution of LIKE according to verb type¹¹

	%	N
unergative	12	1257
transitive	7	1967
unaccusative	4	1043
total		4267

¹⁰ LIKE has probably been used in participle constructions from the outset as well. Its failure to occur in this context among the 50-59 year olds and the 30-39 year olds likely derives from low Ns (N = 9 in both groups).

¹¹ I do not discuss linking verbs (feel, look, etc.) here. They are infrequent (N = 122), and nothing hinges on their inclusion.

Let us tackle the first issue: unaccusatives. There are three possibilities:

- 1) LIKE adjoins to VP, not \square P
- 2) LIKE adjoins to VP with unaccusatives, \square P elsewhere
- 3) unaccusatives project a light verb

The analyses presented thus far in this dissertation suggest that within a constituent, LIKE first targets the functional level before expanding to the lexical level. Consequently, I rule out the first option. The second option remains a possibility, one to which I return below. The third choice is opted for here since, as discussed earlier, much recent work in syntactic theory argues that all verbs have a \square P layer, including unaccusatives (e.g., Levin & Rappaport Hovav 1995; Marantz 1997; Harley & Noyer 1998; Alexiadou 1999; Massam 2001). Thus, the fact that LIKE occurs with unaccusative verbs is predicted by the structure of the VP. The properties that distinguish the two classes of monadic verbs derive from 1) the position where the argument is merged and 2) the semantic properties of the subject. The former has no obvious bearing on LIKE, since it is unclear why merging the argument internally or externally should have consequences for this particle. What is clear, however, is that LIKE does not occur in equal proportions across intransitives. This is an interesting result, and lacking a structural explanation, it is one that supports the semantic distinction between unergatives and unaccusatives (Perlmutter 1978; Levin 1993). Specifically, the frequency of LIKE with unergatives (12%) in comparison to its low rate of occurrence with unaccusatives (4%) suggests that agentivity may be an important factor in the use of LIKE.

This may, in turn, help to explain the second issue, which is why LIKE does not occur with unergative and transitive verbs in equal proportions. While the subject of an unergative

verb generally carries the theta role of agent, this is not necessarily true of the subject of a transitive construction. Consider, for example, the transitive psych verbs (e.g., think, know, fear, hate, etc.). The subject of these verbs is usually an experiencer, as in (118).

- (118) a. Pat knows Sam.
 b. Chris fears snakes.

In neither case can Pat or Chris be said to be responsible for intentionally causing or carrying out some action (i.e., acting as agent). Instead, Pat and Chris experience the emotion or the state denoted by the verb.

If agentivity is correlated to the use of LIKE, then we should be able to find evidence for this in the distribution of LIKE according to the thematic role of the subject. To test this hypothesis, the data were reconfigured according to semantic criteria.

Table 7.5 displays the results, where the theta role of the subject is categorized as either an agent or ‘other’ (patient, theme, experiencer, etc.).

TABLE 7.5
 Distribution of LIKE according to semantic criteria

	unergative		transitive		unaccusative	
	%	N	%	N	%	N
agent	12	1069	10	1121	5	380
other	7	188	4	846	3	663
total		1257		1967		1043

As can be seen, the proportion of VPs containing LIKE is consistently higher with agentive subjects (119) than it is with those bearing other theta roles (120). Further, the difference

between unergatives and transitives has leveled: 12 percent vs. 10 percent respectively.

Both pieces of evidence support a semantic correlation in the use of LIKE.

(119) Agentive subjects

- a. We were **LIKE** playing this weird game. (2/e/f/11)
- b. Dan just **LIKE** stood up. (1/¶/m/22)
- c. Of course, all the Italians **LIKE** run out of their homes. (N/Δ/f/40)

(120) Non-agentive subjects

- a. I was **LIKE** shivering. (2/r/f/11)
- b. The rope just **LIKE** slips right out of his hand. (I/§/f/21)
- c. I'm not sure if my eight year old **LIKE** understands that. (N/X/m/46)

Moreover, note the relative proportions of agents and other types of theta roles across verb type: 85 percent of unergatives have agentive subjects compared to 57 percent of transitives and 36 percent of unaccusatives.¹² This hierarchy corresponds to the overall results seen in Table 7.4. When verb type is cross-tabulated by age, as in Figure 7.3, this pattern, unergative > transitive > unaccusative, is evident across the entire community of LIKE users:

¹² The relatively high proportion of agentive subjects within the class of unaccusatives requires comment. Although the subject of these verbs typically bears the role of patient or theme, a large number of tokens in the current data set consist of forms like go, whose subject is often an agent despite the syntactically unaccusative status of the verb.

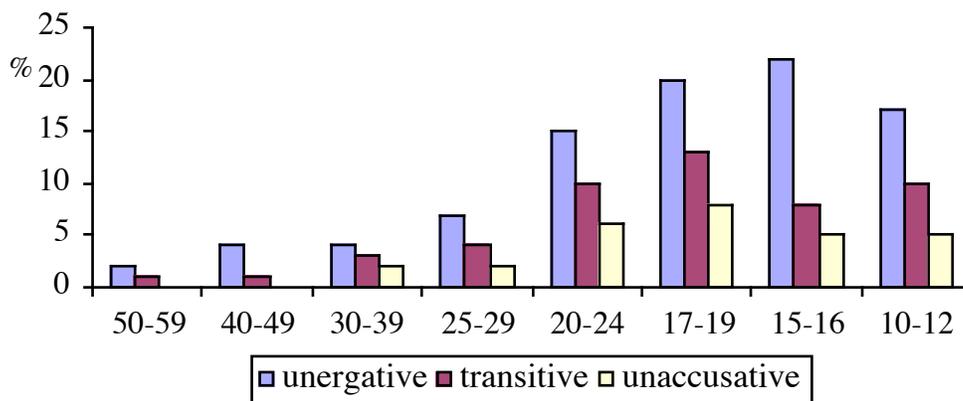


FIGURE 7.3 Type of verb with which LIKE occurs according to age

From the outset, LIKE is most frequent among unergatives, followed by transitives. In contrast, unaccusatives do not constitute a variable context for LIKE until the third generation of users, the 30-39 year olds. From this point, these verbs continue to lag behind. The regular correspondence between the proportion of subjects bearing an agent theta role and the proportion of LIKE that is evident across the sample strongly suggests that it is not verb type *per se* that conditions the use this particle: it is the availability of an agentive subject within each category that determines the frequency of LIKE. In other words, LIKE occurs most often with unergatives because these contain the highest proportion of agentive subjects. The second most frequent context is with transitive verbs because these contain the second highest proportion of agentive subjects, and LIKE occurs least often with unaccusatives because they contain the fewest number of agentive subjects. In sum, these results indicate that the semantic factor is a fundamental aspect of LIKE in the VP context.

At this point I would like to return to the issue of passive constructions. Recall from section 7.1.1 that LIKE is infrequent in this context, accounting for just 4 percent of the data

(N = 67). Crucially, the surface subject of a passive is not an agent. This DP originates in the deep structure as the object of the verb, where it is assigned the role of either patient or theme. Because passives cannot assign Case, however, the object must raise to the subject position in order to satisfy the Case Filter (Chomsky 1981:49). The rarity of LIKE in these constructions, therefore, falls out from semantic factors. As seen in Table 7.5, regardless of verb type, LIKE is highly infrequent with subjects bearing any theta role other than that of agent. This is precisely the context of passives, where the subject is never agentive.

This also factors into the low rate of LIKE in nonfinite copula constructions. As argued above, the particle can precede be in these structures because the copula is hosted in V, below the adjunction site for LIKE. Nonetheless, LIKE is infrequent (3%; N = 86). Note, however, that the subject of these predicates is not always an agent. Consider the data in (103), repeated here as (121).

- (121) a. So then it was cool ‘cause you get to LIKE be smart. (2/u/m/15)
 b. So it went from LIKE being like that, to like that. (2/h/m/18)

In both sentences, the subject of the nonfinite copular clause is an experiencer. This pattern holds across most of these constructions. The lack of an agent in the subject position suggests that the dispreference for LIKE in nonfinite copula contexts derives from semantic considerations.

Let me summarize to this point. The distributional analysis revealed that the frequency of LIKE is increasing in apparent-time, and that concomitantly, it is generalizing to a broader range of contexts (e.g., infinitival ‘to’, ‘do’ support, unaccusatives). This development appears to be constrained by both syntactic and semantic factors. For example, LIKE is more

frequent in participle, modal, and auxiliary constructions than it is in other strings such as with modal and bare verbs. It is also more frequent with unergative and transitive verbs than with unaccusative ones. I have suggested that this derives from a semantic effect, whereby LIKE is more frequent with agentive subjects. Finally, I have provided an account for why certain structures discussed in section 7.1.1 do not constitute part of the variable context, some of which derive from the semantic correlation (nonfinite copula, passives), while others can be explained on structural grounds (finite inflected copula). The crux of this syntactic analysis is that LIKE adjoins to the light verb, and this argumentation was used to explain the position of the particle in the VP context.

One of the implications of the \square P analysis that I have largely refrained from addressing is theoretical. If, following Hale and Keyser (1993, 2002) and Chomsky (1995) unaccusatives in fact lack a light verb, then does the occurrence of LIKE in unaccusative contexts present a problem for the analysis I present here? The answer to this is no. Recall the second option proposed in the discussion of LIKE with unaccusative verbs: LIKE targets VP with unaccusatives and \square P elsewhere. Of course, this is overly simplified. The analysis in Chapter 5 suggested that within a constituent, LIKE first targets the functional level before adjoining to the deeper lexical level. In the VP context, we saw in Figure 7.3 that LIKE does not immediately appear with unaccusatives when it first enters the verbal domain: among the first two generations, the particle appears uniquely with unergatives and transitives. This distributional effect was argued to derive from the semantics of the subject, a position I continue to maintain. However, if unaccusatives lack a light verb, then their later appearance as a variable context may also reflect syntactic factors: in the initial stage there is

no adjunction site. But, as LIKE develops in the verbal domain it begins to spread beyond $\square P$ to VP, and it is at this point that it can appear with unaccusatives. Importantly, this does not restrict LIKE to VP only in those cases where there is no light verb, but it does allow for two adjunction sites in the architecture of the verbal domain. I take no position on whether this is the correct analysis or not.¹³ I simply offer it here so that regardless of the theoretical framework one assumes regarding the structure of unaccusative verbs, the facts surrounding LIKE can be accounted for.

With regard to speaker sex, the overall results reveal a marginal male lead in the use of LIKE: males 9 percent (N = 2105) vs. females 7 percent (N = 2284). When cross-tabulated by age, as in Figure 7.4, this pattern is consistent across apparent-time. With the exception of 30-39 year olds, where LIKE is evenly distributed across the sexes, the use of this particle is more frequent among the males of each age cohort.

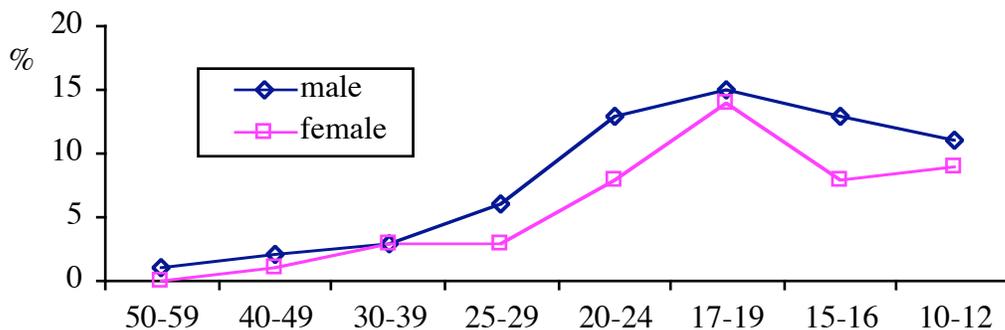


FIGURE 7.4 Cross-tabulation of LIKE by speaker sex and age

¹³ A consequence of this analysis is the possibility of ‘LIKE LIKE’ sequences (e.g., He was LIKE LIKE running), since there is nothing in the syntax to rule out simultaneous adjunction to both $\square P$ and VP (cf. (111), §7.2.1). Presumably this would be blocked by pragmatic factors, but it is an interesting question.

This is the same result we have seen in the two previous chapters as well (DP, Chapter 5 and PredAP, Chapter 6), and it is the opposite of the pattern that appears to be developing in the CP context, where a female association is emerging. I return to this issue in the next chapter, in section 8.2.

7.2.2 Multivariate analysis

The question at this point is whether any of the trends suggested by the distributional analysis are in fact significant factors in the use of LIKE. To shed light on this, I turn to the results of the multivariate analysis, which are reported in Table 7.6. A few words of explanation are necessary. First, the distributional results in Table 7.3 revealed a clear split between participial, auxiliary, and infinitival ‘to’ constructions on the one hand, and modal, bare, and ‘do’ support contexts on the other. Consequently, I have collapsed these into two groups, which I have labeled auxiliary and finite on the basis that each of these strings represents the greatest proportion of its category (auxiliary N = 768; finite verbs N = 2254). Second, following the practice adopted throughout this dissertation, I have grouped the speakers into four age groups: 60 and older (listed in Table 7.6 for the sake of completeness but not included in the regression analysis), middle age 30-59, young adulthood 17-29, and adolescence 10-16.

TABLE 7.6
 Contribution of external and internal factors on the probability of LIKE in the VP context

INPUT factors considered	FW	.05 %	N
Subject type			
agent	.59	10	2581
other	.38	4	1808
range	21		
String type			
auxiliary	.63	13	1377
finite	.43	5	3012
range	18		
Age			
10-16	.65	10	1242
17-29	.62	10	1780
30-59	.24	2	1367
60+	—	K/O [0%]	1094
range	41		
Sex			
male	.55	9	2105
female	.46	7	2284
range	9		
TOTAL N			4389

Each of the factors in Table 7.6 was selected by the step-wise procedure as significant. This is noteworthy in the case of speaker sex, since the difference in the overall proportions of LIKE across males and females is small. However, as seen in Figure 7.4, the sex effect is apparent even among the first generation of speakers who use LIKE in the VP context and from that point forward, it is constant across apparent-time. The significance of sex therefore reflects the stability of this factor across the sample.

The results for age are not surprising. The overall results in Table 7.2 indicate that once LIKE begins to appear in the VP its frequency increases incrementally before rising sharply

among the 17-24 year olds. The multivariate analysis confirms this jump, as speakers aged 30 and older are differentiated from the younger members of the community by a wide margin in the factor weights; note the range of 41. Even though age has an obvious effect, with the two youngest age groups very similar (proportionally and in terms of their factor weights) while the middle-aged speakers barely use LIKE at all (2%; N = 1367), its inclusion effectively highlights the large gulf in the rates of use across older and younger speakers.

Table 7.6 also confirms the effect of string type on the use of LIKE: it is strongly favoured in participial, auxiliary, and infinitival ‘to’ constructions (.63), while disfavoured with bare verbs and modals (.44).

Of all the results in Table 7.6, however, those for subject type are the most interesting. Setting aside age, the semantic status of the subject is the strongest constraint operating on LIKE (range = 21). The particle is favoured with agentive subjects at .59, and strongly disfavoured with those bearing any other thematic role, .38. This confirms the critical role of semantics in conditioning the use of LIKE in the VP context. As far as I am aware, this is the first time that evidence of a semantic constraint has been brought to bear on this feature.

Verb type was not included in the regression reported in Table 7.6. I have suggested that the hierarchy unergative > transitive > unaccusative falls out directly from the semantic effect: the rate of LIKE with each verb type reflects the proportion of available agents in each of these contexts. When verb type is run as a factor (not shown here), the difference between unergatives and transitives is not significant ($p > .10$ on a chi-square test). A “more general and efficient model of the variability” (Guy 1988:133), or better fit, is achieved when these two verb types are collapsed. This creates a binary division between

unaccusatives vs. unergatives/transitives. These two groups form natural classes in terms of the semantic status of the subject. Consequently, verb type and subject type thus configured become non-orthogonal factors: the relative proportion of agentive subjects distinguishes the verb types, and the factor group of subject type tests the effect of agentivity. In other words, verb type is a sub-category of subject type. To run the variable rule program effectively, factors must be orthogonal (Sankoff 1987), requiring one factor to be excluded from the analysis presented here. Comparison of the log likelihoods of two separate runs (also not shown here), one testing verb type and one testing subject type, confirms that subject type achieves a better fit of the data (compare -1069.698 with verb type to -1063.769 with subject type). This, in conjunction with the strength of this factor in Table 7.6, establishes the crucial role of semantics in constraining the use of LIKE in the VP.

7.2.3 The evidence from adverbs

Throughout this discussion, I have maintained that the adjunction site for LIKE in the VP is the light verb phrase. The data considered in section 7.2.1 support the $\square P$ hypothesis. Further support is provided by the syntagmatic position of LIKE in relation to adverbs.

Cinque (1999, 2004) argues that adverbs are merged in the specifier position of distinct functional projections such as mood, tense, aspect, voice, etc., and proposes a fixed universal hierarchy to account for the relative ordering of adverbs across languages. Much of the detail in this work exceeds the present purpose. I will simply note here that the ranking given in Cinque (1999:106) is compatible with Jackendoff (1972, 1997, 2002), where adverbs are grouped into the three classes identified in (122).

(122) speaker-oriented adverbs > subject-oriented adverbs > manner/degree adverbs

Following Jackendoff (1972:56-58), speaker-oriented adverbs relate the speaker's attitude toward the event. Subject-oriented adverbs express additional information about the subject, and manner/degree adverbs qualify the event/motion denoted by the verb itself. I have represented these classes as a hierarchy in (122) in order to demonstrate their structural relation to each other: each class to the left is merged higher in the syntax than the class on its right. The two crucial points are:

- 1) adverbs may be located hierarchically at different levels of structure
- 2) manner adverbs are ordered low

This predicts that the syntagmatic order of **LIKE** and adverbs may differ depending on the class to which the adverbial belongs. Specifically, an adverb that occurs higher in the structure will, in the unmarked case, precede **LIKE**, while one that occurs lower in the structure will follow **LIKE**.

As the data in (123) demonstrate, **LIKE** can appear on either side of an adverb:

- (123) a. And then he **LIKE** slowly added more and more things (3/A/m/15)
 b. He actually **LIKE** did something. (1/ϕ/m/21)

The sentence in (123a) contains a manner adverb. These are argued to occur low in the syntax. The sentence in (123b), on the other hand, contains a speaker-oriented adverb. In Cinque (1999), these are speech act, evidential, and epistemic adverbs, hosted in Mood projections. These are held to occur high in the syntax. Importantly, **LIKE** is situated to the

left of the manner adverbial in (123a), but to the right of the evidential adverbial in (123b).

Is this ordering accidental or is it systematic?

Close inspection of the data reveals that the position of LIKE in a VP is fairly systematic.

There are two patterns, as in Table 7.7. Examples follow in (124)-(127).

TABLE 7.7
Classification of adverbs

	I Adv LIKE		II LIKE Adv	
speaker-oriented	{ really (truly) literally actually honestly		{ slowly gradually	manner
subject-oriented	{ still always never		{ really (intensifier) totally slightly	degree

(124) Speaker-oriented

- a. I don't really LIKE judge people on what music they listen to. (2/u/m/15)
- b. We literally LIKE cooked all the food. (N/ð/m/26)
- c. He actually LIKE stood up. (I/ç/m/21)
- d. They honestly LIKE threatened me. (I/ç/m/21)

(125) Subject-oriented

- a. Andrea still LIKE comes to lunch with us. (2/a/f/16)
- b. Me and my friends, we always LIKE took rulers. (3/V/m/11)
- c. They like it but they never LIKE played. (3/F/f/17)

(126) Manner

- a. But people will LIKE slowly get into it. (2/n/f/19)
- b. And then they LIKE gradually changed like how they looked. (2/u/m/15)

(127) Degree

- a. A trade that I LIKE really like was the one they had got from Jersey. (3/H/m/12)
- b. Some people LIKE totally fell into the mold. (2/i/f/19)
- c. The glue LIKE slightly falls off your hair. (2/r/f/11)

In Group I, we find that with speaker-oriented (124) and subject-oriented (125) adverbs, LIKE follows the adverbial. If LIKE adjoins to $\square P$, below the projections that host these types of adverbs, then this ordering is predictable based on structural facts. Similarly, in Group II, LIKE precedes manner (126) and degree (127) adverbials, which again is explicable on structural grounds.

These patterns therefore add support to the $\square P$ analysis of LIKE, at least tentatively. First, a thorough search of the full Toronto English corpus resulted in just 10 tokens of LIKE with manner/degree adverbs within the VP context. Infrequency of manner/degree adverbs, on the one hand, and of LIKE in this context on the other, make the combination exceedingly rare. Added to these factors is the fact that these patterns are not categorical. Consider (128).

(128) She didn't learn sign-language until she got to University,
which I feel totally LIKE took away their culture and their language. (I/=/f/28)

There are two ways to interpret totally in (128). Either it is speaker-oriented, signalling that what follows is the speaker's perspective of the effect of not learning sign-language early, or it is a manner adverb (i.e., the way in which their culture was taken was total). In the first case, the position of LIKE is as predicted. In the second case, the more likely interpretation, the position of LIKE is potentially problematic.¹⁴

Further, we cannot rule out that LIKE could precede a higher-level adverb, since previous research has established that this particle can occur with adverb phrases (AdvP)

¹⁴ Of course, we can account for (128) straightforwardly if we allow for the generalization of LIKE beyond $\square P$ to VP.

(e.g. Andersen 2001). There is presumably no structural reason why LIKE could not, for example, adjoin to the left edge of the AdvPs in (124) and (125). In fact, this adjunction site is suggested by (129). I assume that in this instance, the pragmatic effect is that LIKE scopes over the adverbial, rather than over the full VP. I leave the issue open.

(129) It's amazing 'cause you put on music and you actually got to LIKE actually treat people. (I/=f/28)

Nonetheless, the available evidence suggests that it is possible to predict the position of LIKE when adverbs are also present based on structural factors. It appears that the particle follows adverbs that occur high in the syntactic structure of the VP (i.e., speaker and subject-oriented adverbs), but generally precedes those that occur low in the syntax (i.e., manner and degree adverbs).

7.3 DISCUSSION

This chapter had two primary goals: to demonstrate that when tracked accountably, the VP is no more a preferential slot for LIKE than other positions are (e.g., PredAP), but that when LIKE does occur in the VP, it has a fixed position in the string. In addressing these topics, I have also been able to show a number of other results. For example, I have argued that LIKE adjoins to the light verb, and this hypothesis was used to account for many of LIKE's distributional properties, such as the position it occupies in the domain of verbs, its inability to precede the finite inflected copula, and its syntagmatic position in relation to adverbs within the VP. These results argue strongly for syntactic constraints on this particle. I have also been able to show that in addition to these structural factors, semantic considerations

are significant in constraining LIKE: this particle is favoured when the subject bears the theta role of agent. I have suggested that the operation of this semantic effect explains the effect of verb type on its frequency of use, as well as the nearly categorical failure of LIKE to appear in passives and with the nonfinite copula. Thus, this chapter has established that the grammar has an important role, both syntactically and semantically, in the use of LIKE. In other words, whatever the function of LIKE, its consequences reach beyond pragmatics.

I have also demonstrated that the frequency of LIKE is increasing in this context, and that as it develops within the VP, it is slowly generalizing to new constructions. Finally, the VP provides evidence of a male association in the use of LIKE, one that has been in place since it first spread to the VP.

In the next chapter, I will bring together the various results I have reported throughout the investigation, and I will propose a model for the development of LIKE.

Chapter 8

THE DEVELOPMENT OF DISCOURSE ‘LIKE’

In Chapter 1, I suggested that when viewed across the generations, discourse uses of LIKE will be seen to have developed gradually and systematically, arriving at their current state through regular processes of language change. The logic behind this assumption is that language, despite inherent variability, is not haphazard. Consequently, I set out to examine LIKE in a contemporary variety of English, using apparent-time data from a single, cohesive speech community: Toronto, Canada. The framework I adopted is accountable, variationist, and sociolinguistic. This feature has never been investigated this way before.

Once the distinct adverbial function of LIKE was laid bare in Chapter 2, I set out to contextualize the development of its discourse functions within the grammaticalization models proposed in recent work by Traugott (1997 [1995]) and Brinton (forthcoming). However, these pathways explicitly make reference to the discourse marking function of pragmatic features, defined here as those forms that take wide scope over ‘units of talk’ (Schiffrin 1987:31). Particles, which occur within propositions, do not fit into this model in any obvious way. In this chapter, I will argue that this is because the development of the particle is not a simple continuation of the evolutionary pathway from preposition to discourse marker. Rather, I suggest that once LIKE grammaticalizes as a marker, this sets the scene for the beginning of a new pathway whereby it begins systematically to enter syntactic structure, gradually generalizing from one maximal projection to another.

Because the particle cannot be accounted for straightforwardly within the clines argued for in Traugott (1997 [1995]) and Brinton (forthcoming), it was necessary to suspend the discussion of the full evolutionary pathway that LIKE is currently undergoing until its use in various positions within an utterance was examined. In four successive chapters, I investigated the use of LIKE in four major contexts: CP, DP, PredAP, and \square P. These analyses were based on naturally occurring speech data from 97 speakers between the ages of 10 and 87. In all, over 20,000 tokens of structurally delimited and accountably circumscribed variable contexts were considered. These facts are summarized in Table 8.1.

TABLE 8.1
Summary of analyses

Chapter 4	clause-initial	CP, TP	5,737
Chapter 5	determiner phrase	DP, NP	4,408
Chapter 6	predicate adjectives	PredAP	4,298
Chapter 7	verbal domain	\square P	5,843
TOTAL			20,286

If LIKE is not a random and unconstrained linguistic device, then the social and linguistic conditions on its use should manifest themselves as regular patterns within the community. Poplack and Tagliamonte (2001:89) have stated that such patterns “can only be discerned by systematic and exhaustive quantitative analysis of the data, in conformity with Labov’s principle of accountability.” It is precisely these patterns that I have endeavoured to reveal, and to which I now turn. I first discuss their implications for the development of LIKE, section 8.1, and I then turn to the social embedding of this feature, section 8.2. Finally, in section 8.3, I offer some closing remarks.

8.1 THE DEVELOPMENT OF DISCOURSE ‘LIKE’

8.1.1 From whence it came

A growing body of work argues for the consideration of discourse markers as deriving from grammaticalization processes (e.g., Traugott 1997 [1995]; Brinton 1996, 2005, forthcoming). This is a stance that researchers on LIKE have adopted virtually from the outset, and which was most fully articulated by Romaine and Lange (1991). Equipped with advances to the theory from Traugott (1997 [1995]) and Brinton (forthcoming), I addressed this concept in Chapter 3, where I proposed the following developmental trajectory for LIKE as a discourse marker:

(130) preposition/conjunction > sentence adverb > discourse marker

Each stage in (130) is exemplified in (131) (see Chapter 3 for further examples):

(131) a	preposition	It looks LIKE a snail; it just is a snail. (I/@/f/19)
	conjunction	It felt LIKE everything had dropped away. (I TM /m/40)
	b. sentence adverb	We need to smarten it up a bit LIKE . (N/©/f/76)
	c. discourse marker	LIKE the week had already gone by. (N/Q/f/72)

The proposed pathway in (130) is identical to that undergone by many pragmatic forms throughout the history of English, exemplified in section 3.2 by soðlice and indeed. It differs from that argued for in previous accounts of LIKE (Meehan 1991; Romaine & Lange 1991; Andersen 2001) in that it incorporates the use of LIKE as a sentential adverb. The current model therefore has three main advantages.

First, it connects the formerly distinct ‘British’ (131b) and ‘American’ (131c) patterns as emblematic of different developmental stages. Thus, the former division, which was

defined along geographic parameters but remained unexplained within a broader semantic/pragmatic framework, is eliminated. Linking the sentential adverb and the discourse marker in this way is supported by the following factors: (1) the adverb pre-dates the discourse marker (cf. (41) in Chapter 3), which is exactly what we would expect if the discourse marker represents a later developmental stage, (2) the discourse marker takes wider scope than the adverb, again predicted by the pathway in (130), and (3) ‘American’ uses of LIKE are not restricted to North America, but as I have demonstrated (e.g., (7) in Chapter 1; (45c-e) in Chapter 3), are currently productive among older speakers of isolated regional varieties across the United Kingdom. This is not explicable on the basis of LIKE as an American innovation that developed during the 1960s among counterculture groups in New York (cf. Andersen 2001 and references therein), but it is straightforwardly accounted for as a consequence of the trajectory in (130).

The second advantage of the current proposal is that it provides the missing link in the archetypal increase in scope undergone by a grammaticalizing form (Traugott 1997 [1995]; Traugott & Dasher 2002; Brinton forthcoming). Consider again the model put forth by Romaine and Lange (1991) (Figure 3.1, Chapter 3). This view of the development of LIKE entails cataclysmic changes in the restrictions on this lexeme, since from its functions as a preposition and conjunction — with highly constrained syntactic positions — LIKE develops pragmatic functions and simultaneously becomes free to occur in a number of positions within a clause. This requires immediately assuming the ability to adjoin to a wide range of structures. Crucially, LIKE shifts from scope within propositions (preposition, conjunction) to scope over discourse (discourse marker) while also maintaining narrow scope (discourse

particle) without any intermediate stage. In contrast, if LIKE develops first as a sentential adverb, which, as I have just argued, there are good reasons for believing, then just such an intermediate stage is posited. In other words, whereas the previous model suggests a pathway of scope within the proposition to scope within the proposition/scope over discourse, the pathway I have argued for here posits a broadening of scope that progresses gradually from within the proposition to over the proposition to over discourse. This is precisely the pathway argued for by Traugott (1997 [1995]), Traugott and Dasher (2002), and Brinton (forthcoming).

The third advantage of the pathway in (130) is subtler than the previous two, but it is one to which I have already alluded. The current proposal does not assume that the development of LIKE as a particle and its development as a discourse marker were contemporaneous, nor does it suggest that these developments are one and the same thing. Indeed, as I will show, the sum of the apparent-time evidence suggests that these are different developments, justifying the treatment of the marker and the particle as separate, though related, discourse features.

8.1.2 Ongoing development

In each chapter, we have seen LIKE enter a particular context at a particular point in time. In Chapter 4, for example, the apparent-time data indicate that while LIKE was firmly entrenched as a marker on matrix CPs among the oldest speakers in the sample, it did not generalize to TP until the 40 year olds. In clause-internal positions, where I have referred to LIKE as a particle, it first appears on the periphery of DPs among speakers in their seventies,

APs among speakers in their sixties,¹ \square Ps among speakers in their fifties, and NPs do not enter the panoply of contexts until speakers in their thirties. This layering of contexts across apparent-time, captured so eloquently by the metaphor in (132), is displayed in Figure 8.1.

(132) It **LIKE** went **LIKE** seamlessly into it. (N/p/m/20)

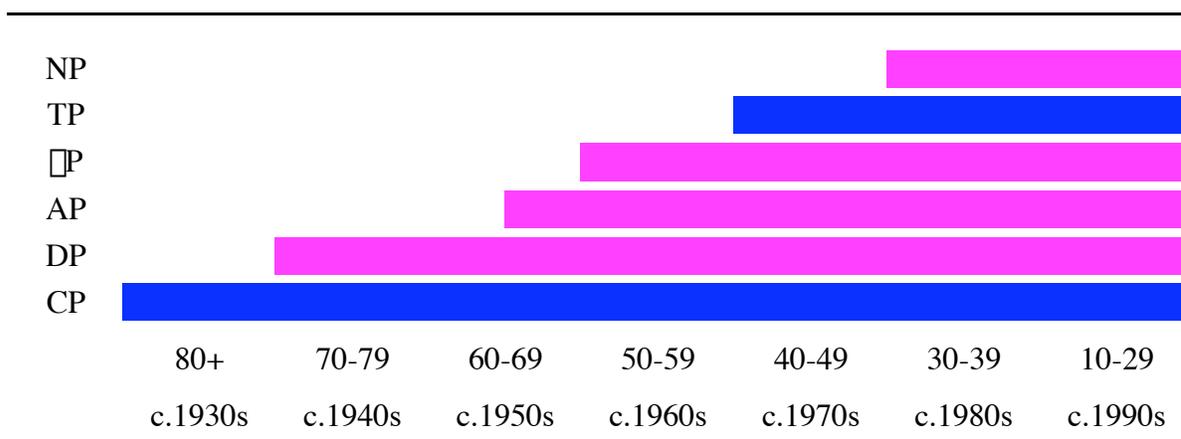


FIGURE 8.1 Generalization of LIKE across maximal projections apparent-time
(darker shading denotes the marker; lighter shading denotes the particle)

This figure provides a remarkable display of grammatical development, with LIKE slowly generalizing from one maximal projection to another within syntactic structure. These results strongly suggest regular, step-wise development. Further, it is clear that the use of LIKE in clause-internal positions arose after LIKE became established as a discourse marker, since here the first context in which LIKE functions as a particle, DP, evolves circa 1940. In contrast, as discussed in Chapter 3, the marker probably began to develop at some point around the end of the nineteenth or the beginning of the twentieth century, a

¹ Recall from Chapter 6 that the PredAP context emerges among speakers in their fifties, but the use of LIKE with attributive adjectives appears to be manifested first among those in their sixties.

hypothesis that is supported by the apparent-time evidence. We can therefore rule out that the particle and the marker evolved in tandem as a single development. It is also evident that the synchronic state of LIKE, where it appears in a broad range of syntactic positions, did not emerge all at once as a fully-fledged system. Instead, it developed one maximal projection at a time, beginning with matrix CPs, across the span of almost a century. Indeed, the layering of contexts reveals two patterns: (1) generalization from higher functional projections to lower functional projections (i.e., in terms of the hierarchical structure of syntax), and (2) generalization from functional projections to lexical projections within a domain.

Higher > lower functional projections

As a marker, LIKE first targets CP before generalizing to the lower TP. Both of these are functional categories. As a particle, the first projection to which LIKE adjoins is DP, also a functional category. It later generalizes to \square P, which is again functional. We therefore have evidence for the pathways in (133):

- (133) a. marker CP > TP
 b. particle DP > \square P

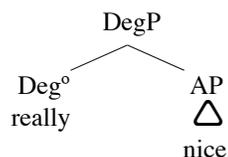
Thus, as both a marker and a particle, LIKE systematically appears in a context that occurs higher in the structure before spreading to a lower projection. It appears, therefore, that after initially developing as a discourse marker, adjoining to the outermost level of the syntax, LIKE slowly makes its way down into syntactic structure.

Functional > lexical projections

This same pattern is also evident within constituents. In this case, the spread is not from one functional category to another, but from the functional category to the lexical one that forms its complement. The best evidence for this was provided in Chapter 5, where we saw that LIKE first targets DP, and only later begins appearing on the periphery of NP. A similar possibility was discussed in Chapter 7, where it was suggested that the later rise of unaccusatives as a variable context may stem from the generalization of LIKE from \square P to VP (assuming, that is, that unaccusatives do not project a light verb; e.g., Hale & Keyser 1993, 2002; Chomsky 1995).² We therefore have evidence for the pathways in (134):

- (134) a. DP > NP
 b. \square P > VP

² Though not explored in Chapter 6, the AP context may also provide evidence for the pathway of functional > lexical. I have assumed a structure in which adverbs fill the specifier slot of AP (see (96), page 157). However, an alternative analysis is available. Abney (1987) argues for the following structure, in which all APs are complements of Degree Phrases (DegP), regardless of whether the head of DegP is lexically specified or not:



In this structure, LIKE would adjoin to DegP, the higher functional category; recall that it precedes any adverbs that modify the adjectival head. Under this view, the problematic cases in (97), in which LIKE follows the adverb, may simply represent the spread of LIKE beyond DegP to AP, since these rare examples were all produced by speakers 30 years old and under (i.e., approximately 35 years after LIKE first appears in adjectival contexts). Note also that this type of analysis allows us to expand the pathway in (133b) to DP > DegP > \square P, a trajectory which continues to maintain the higher > lower direction of generalization.

To summarize, the systematicity of the development of LIKE along the pathways in (133) and (134) indicates regular and rule-governed change, change that has not proceeded in an *ad hoc* manner. This conclusion is further corroborated by Figure 8.2, which tracks the progression of LIKE not only by the linguistic contexts in which it appears, but by its frequency across apparent-time as well.

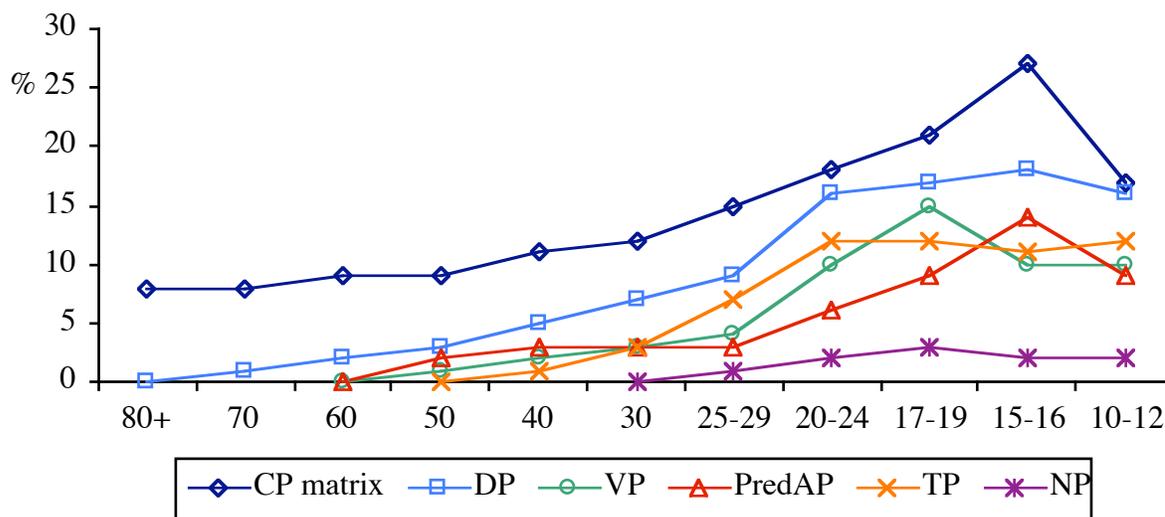


FIGURE 8.2 The rise and development of LIKE as a pragmatic device

This expanded view of the layering of LIKE within the grammar demonstrates how, after it appears in a given context, it continues to do so, and further, that its frequency consistently rises. Thus, once LIKE spreads to a new projection, that particular X^{\max} is established as a possible adjunction site for successive generations of speakers. Such a finding is strongly suggestive of the operation of grammar-internal factors in the development of LIKE. Moreover, when the various contexts are viewed in this way, it is immediately apparent that the youngest speakers in the community, the adolescents, are not using LIKE in contexts in which older segments of the population do not; the difference is that they use it at higher

rates. This latter result falls out from the gradual slope of change observable for every context in which LIKE appears: this is the monotonic association of frequency with age that is characteristic of change in progress (Labov 2001:460). In other words, the characterization of LIKE as nothing more than a marker of adolescent inarticulacy (e.g., Diamond 2000) is not only simplistic, it is simply false.

To this end, consider the regularity of internal constraints across the population of LIKE users for each of the contexts discussed in Chapters 4 through 7. The distributional and multivariate findings are summarized in Table 8.2.

TABLE 8.2
Summary of linguistic constraints on LIKE

Context	Population	Constraints	Direction of effect
CP (matrix)	10 – 87	Status Position	bare > conjoined/marked follow up > medial > initial
DP	10 – 79	Status Definiteness Modification	complement > argument indefinite > definite unmodified > modified
□P	10 – 59	Subject String	agent > other auxiliary > finite
PredAP	15 – 24	Lexical verb	other > copula

The uniform operation of these constraints, regardless of speaker age, provides clear and unambiguous evidence that LIKE undergoes ordered and systematic change. As hypothesized, these conditions are manifested as regular patterns of use within each age group that uses LIKE in the contexts examined here. Importantly, these constraints are visible even among adolescents, that group who has borne the brunt of negative attention for their use of this feature. Clearly, however, they are not the sole proprietors of LIKE, nor do

they use it randomly. Rather, while the rate of use is visibly higher among adolescents than among adults, as predicted by the logistic incrementation of linguistic change (Labov 2001:455), this use is constrained identically across all age groups. In other words, the community shares a single variable grammar for LIKE (see Poplack & Tagliamonte 2001:93).

Since this grammar consists of a range of constraints but lacks a single over-arching factor that operates across all the contexts in which LIKE may be used, an intriguing question raised by this research is: Does each context involve a 'different' LIKE? My response to this is, quite simply, no. The constraints on LIKE are context-specific, not function-specific. In other words, setting aside the CP context where LIKE functions as a discourse marker, the LIKE that appears within propositions shares a number of pragmatic functions, and these are independent of the context in which this particle appears (see Andersen 2001). Thus, the factors that operate on LIKE when it adjoins, for example, to the left edge of DP reflect contextual effects that are specific to the DP and not a unique instantiation of this form.

This type of linguistic conditioning is common across all levels of grammar. Allophonic variation provides a straightforward example. The phonetic output is constrained by the phonological context in which a phoneme is located. Consider the voiceless stops in English, /p,t,k/. Whether one of these is the onset to a stressed syllable or the coda of an unstressed syllable, the underlying phoneme remains the same despite the fact that each linguistic environment affects its realization in different ways. Extending this one step further, English t/d deletion also applies in certain contexts (e.g., C__#C, C__#/j,w/, /C__#V, etc.) but not in others (e.g., *#C__V). We would never suggest, however, that the

/d/ that undergoes deletion is somehow different from the one that does not. This is the type of linguistic conditioning that obtains with LIKE. We have just one discourse feature, and the probabilistic conditions on whether it is used or not vary according to linguistically defined factors that prevail in a given context.

Of course, not all of the constraints on LIKE are variable. For example, its syntagmatic position in the string falls out directly from the level of structure that it targets: X^{\max} . Thus, Andersen's (2001:276) finding that the type of phrase crucially constrains the likelihood that LIKE will appear within the construction does not reflect any special properties of the phrase itself, but derives directly from the category to which LIKE adjoins. In the case of the verbal domain, for example, LIKE appears between any functional material in T° and the main verb not because it is attracted to lexical material but because it targets $\square P$. Further, the lower incidence of LIKE within a DP reflects the later development of NP as a possible adjunction site. The higher DP projection was established first, and therefore contains a higher rate of LIKE than the more recent NP. This is evident in Figure 8.2, where LIKE is consistently more frequent on the periphery of DP than it is on the periphery of NP.

In fact, observe once again the layering of contexts in Figure 8.2. Each layer correlates with the point at which LIKE generalized to the respective projection. The oldest context, CP, is consistently the most frequent position for LIKE across apparent-time. The DP context, the second oldest, is consistently the second most frequent position for LIKE. There is slight fluctuation among the $\square P$ and PredAP contexts, but these two contexts generally rise together. Finally, the most recent context, NP, consistently contains the lowest proportion of LIKE.

Returning then to the three preferential positions for LIKE that have been reported in the literature, CP, DP, and \square P, we can make the following observation: the two most frequent slots for LIKE are CP and DP, where its overall distribution is 14 percent (N = 3363) and 10 percent (N = 4047) respectively. Thus, the current results appear to support previous findings. However, assessment of the proportional results in light of text frequency cannot be made in such a straightforward manner. As defined here, the CP and DP contexts reflect more limited sets than what these contexts constitute in previous research, where ‘DP’ entails both DP and NP and ‘CP’ includes both CP and TP (e.g., Andersen 2001; Levey 2003; Wolgemuth 2003; Tagliamonte 2005). In each case, the frequency of LIKE is much lower in the second and more recent context than it is in the first and older one. If combined, the overall distribution would drop. Thus, the text frequency results probably reflect the raw frequency of these types of elements in grammar more generally, while the current findings reflect developmental factors.

What of the \square P? It is the third most frequent context for LIKE, where the overall distribution is 8 percent (N = 4389). Recall from the last chapter, however, that the proportion of LIKE in this context is not significantly different from its frequency in the PredAP context, which is 6 percent (N = 3455). This too suggests that the earlier findings reported in the literature reflect the overall text frequency of clauses, nouns, and verbs, which are the three most frequent elements of grammar (e.g., Altenberg 1990:185). Further, although CP, DP, and \square P are recurrently reported to co-occur with LIKE most often, the ranking differs from study to study (see Chapter 7), while APs are consistently reported to contain few instances of LIKE. In contrast, the proportional results reveal that PredAPs are

not quantitatively differentiated from \square Ps, and the ranking CP > DP > \square P/PredAP reflects the sequentially ordered emergence of each of these contexts within the grammar as possible adjunction sites for LIKE.

Strikingly, this developmental trajectory provides further corroboration for the hypothesis that discourse uses of LIKE were not purely a North American innovation. The DP and \square P contexts are in evidence in the Toronto materials, collected in 2003 and 2004, beginning only with the seventy and sixty year olds respectively. The following examples from Northern Ireland and Scotland in 2001, however, demonstrate that these contexts had some currency at least two decades earlier in isolated communities outside of North America (Tagliamonte to appear; see fn.5, Chapter 1):

(135) DP

- a. It was only **LIKE** a step up to this wee loft. (CLB/c/m/91)
- b. Oh, it was **LIKE** boots we wore. (CLB/a/f/89)
- c. It left **LIKE** a form of a chimney. (CLB/l/f/86)

(136) \square P

- a. We were **LIKE** walking along that Agohill road. (CLB/l/f/86)
- b. And it was just **LIKE** doing what I telt you about. (AYR/f/f/78)
- c. They were just **LIKE** sitting waiting to dies. (AYR/c/m/75)

Again, this is not explicable on the basis of diffusion from North America to the United Kingdom. It does follow, however, if the discourse particle represents a later development, one that began, as I will argue in the following section, with the marker. The marker in turn developed from the use of LIKE as a sentential adverb, a form present in both the British and Canadian data among older speakers (e.g., (43) in Chapter 3). Thus, while the current

results are specific to Toronto, they are generalizable to other Englishes in that they provide a model for the development of LIKE.

8.1.3 Implications for grammaticalization

As discussed above, the grammaticalization clines in Traugott (1997 [1995]) and Brinton (forthcoming) refer to the development of discourse markers. The apparent-time trajectory in Figures 8.1 and 8.2 indicates that among the oldest members of the current Toronto population, LIKE is well established in initial position on matrix clauses, where it functions as a marker. Clearly, however, the clause-internal uses of LIKE, where it functions as a particle, were later developments. This suggests that, following the diachronic pathways posited by Traugott and Brinton, LIKE has followed the cline of the discourse marker and thus, the marker develops first. The particle can only come later. The question is: Is the evolution of LIKE as particle the next step along the pathway in (130), repeated here as (137), or does it represent something else? If so, what would this be?

(137) preposition/conjunction > sentence adverb > discourse marker

I would like to suggest that the development of the particle is not simply an extension of (130/137), but that it is the beginning of a new trajectory, one that begins with the marker. Two pieces of evidence point in this direction. First, in developing as a marker, LIKE moved outward to the edge of syntactic structure. As we have seen, however, the particle develops in the opposite direction. Once established on the periphery of the clause, LIKE begins to work its way back into the syntax, spreading from one maximal projection to another.

Second, the development of the marker entails gradually increasing scope. In contrast, the particle takes narrow scope, modifying whatever element occurs to its right (Andersen 2001:273). These observations make it possible to hypothesize that the differences between the two developments are diagnostic of distinct evolutionary pathways. Moreover, the later rise of the particle, also a discourse feature with specialized pragmatic functions (e.g., focus, ‘loose talk’, etc.; see Chapter 1), suggests that the marker presents the starting point for the particle. Consequently, I propose that this pathway begins with the marker, and it is from it that the particle develops, as in (138):

(138) discourse marker > discourse particle > _____ ?

I do not show (138) as ending with the particle, but allow instead for another stage. What form this stage may take remains an open question, but the possibility for further development is suggested by Sankoff and Brown’s (1976) analysis of the Tok Pisin relativizer, *ia*. There, a pragmatic form — used in discourse as a focusing device as well as for interactive functions such as checking or negotiating understanding — develops a syntactic role, coming to mark relative clauses, topic-comment structures, and cleft sentences. It is by no means certain that LIKE will also develop non-pragmatic functions in the grammar, but Sankoff and Brown’s analysis establishes that the evolution of syntactic functions from discourse features is neither unprecedented nor impossible.

It should be noted that the way in which LIKE is developing, generalizing downward one maximal projection at a time, differs substantially from the model of grammaticalization argued for by Roberts and Roussou (2003). Using extensive cross-linguistic data, the

authors propose that grammaticalization involves local reanalysis that moves upward along functional heads. Clearly this is not what happens with LIKE, which not only progresses downward from the periphery of CP but also does not (yet) entail reanalysis. Rather, it continues to exhibit generalization, shedding the restrictions on its distribution. However, Roberts and Roussou do not consider discourse-level aspects of language, focusing instead on developments within the inflectional, complementizer, and determiner domains. As an adjunct, LIKE is best characterized as external to these systems. Thus, these differences may be emblematic of the types of change under consideration.

At this point we may ask ourselves whether the development of the particle continues to constitute grammaticalization, or whether it is better accounted for in some other model of linguistic change. To address this question, let us consider the characteristics of the ongoing development of LIKE. The current analysis has demonstrated that LIKE comes to be used in a broader range of contexts as new maximal projections come available as possible adjunction sites, and that concomitantly, its frequency continues to rise. Further, the pragmatic literature has established that there is a proliferation of pragmatic functions (e.g., Schourup 1983; Underhill 1988; Andersen 1997 et seq.). Thus, in addition to its duties in the textual component where it may signal elaboration and clarification, LIKE comes to signal metalinguistic focus, hesitation, ‘loose talk’, etc. In sum, there is evidence of generalization (Hopper & Traugott 2003; Heine 2003), pragmatic strengthening (Traugott 1988; Traugott & König 1991; Hopper & Traugott 2003), and the increase in the overall rate that is characteristic of all types of linguistic change (e.g., Hook 1991; Hopper & Traugott 2003). All of this is highly suggestive that grammaticalization is continuing.

Finally, what of the narrower scope of the particle? Given that I have suggested the particle is evolving along a cline distinct from that of the marker, this need not be reconciled within the model proposed by Traugott (1997 [1995]) and Brinton (forthcoming). Indeed, unlike markers, I would argue that scope is not a criterial element in the development of particles. Rather, the reduction of scope from marker to particle is epiphenomenal, deriving entirely from the smaller domain over which a particle presides. In the case of LIKE, it functions as a marker to link units of discourse (Schiffrin 1987; Fraser 1988, 1990), but as a particle, it adjoins to individual projections within syntactic structure to signal pragmatic implicatures within the utterance. In so doing, its scope must narrow; wide scope is ruled out by the clause-internal level that it targets.

In sum, the results presented here suggest that once LIKE grammaticalizes as a discourse marker, it begins a new path and develops as a particle. Thus, in the initial stage, scope broadens, allowing the form to move from within the proposition to its edges, ultimately ending up in the syntactic adjunct slot where it links sequences of discourse. From this position, a marker may begin a new cline and start moving inward where its scope necessarily narrows again. In short, this research suggests that the place where discourse features return 'into' the grammar is from their position on the periphery of the clause.

8.2 SOCIAL EMBEDDING

An interesting question at this point is whether the range of developments undergone by LIKE has social correlates. We know, for example, that both as a marker and a particle, the frequency of LIKE is increasing in all contexts in apparent-time (Figure 8.2). Thus, as far as

we can tell, this change continues to progress. With this in mind, is there evidence that speaker sex is a factor in the ongoing rise of LIKE? After all, sex is an important concomitant of linguistic change, with females typically at the forefront (e.g., Labov 1990, 2001; Chambers 1992). In fact, Labov (2001:462) states that the acceleration of changes associated with males is “dampened by mechanical factors,” since the primary input is from more conservative female speakers.

Attitudinal studies reveal that all forms of LIKE — marker, particle, and quotative be like — are overtly associated with females (e.g., Dailey-O’Cain 2000; Buchstaller 2003). Quantitative research corroborates this finding in so far as the quotative is concerned: it is clearly favoured by females (e.g., Tagliamonte & D’Arcy 2004, under review). As discussed in section 1.3.1, however, there has been little agreement in the literature regarding the effects of speaker sex and the use of LIKE as a discourse feature.

The findings from the current analysis reveal divergent sex effects, which may help to clarify why previous research has produced mixed results. As a marker, LIKE is significantly correlated with females. Although the distributional margin between the two sexes is narrow, the association with females is largely uniform across apparent-time (Figure 4.3, Chapter 4). By contrast, as a particle, LIKE is significantly correlated with males, and this pattern is also consistent across apparent-time (Figure 5.1, Chapter 5; Figure 6.2, Chapter 6; Figure 7.4; Chapter 7). These overall patterns are displayed in Figure 8.3, which charts the distribution of LIKE according to speaker sex in each of the four contexts considered in this analysis.

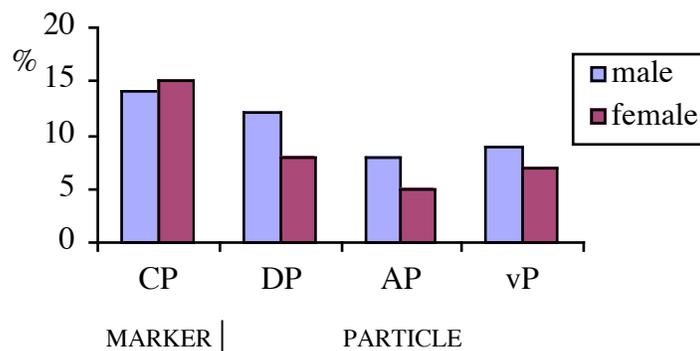


FIGURE 8.3 Distribution of LIKE in four contexts according to speaker sex

The clear split between the effect of sex on LIKE as a marker and as a particle suggests a functional division along the lines argued for by Dubois (1992) and Erman (1992). Although it is clear that quantitative differences are in place that correlate with sex, the fact that these obtain in opposite directions depending on the function of LIKE indicates that what differentiates the sexes is primarily qualitative: females favour the marker while males favour the particle. This finding adds further support to the hypothesis that the marker and the particle reflect separate developmental pathways, since the social effects operate in the opposite directions for each function. This in turn may explain why previous studies have reached different conclusions, since the marker and the particle are usually collapsed as a single option when calculating the number of occurrences produced on average by each of the sexes (typically determined based on the ratio per 1000 words; though see Dailey-O’Cain 2000).

It is precisely the association of the particle with males that creates an interesting picture for the social embedding of linguistic changes. The locus of change is usually attributed to females (Labov 1990:240), who favour incoming prestige forms and lead in the

advancement of system-internal innovations (i.e., change from below). In revealing a consistent and significant favouring effect of males in the use of LIKE as a particle, this analysis has therefore unearthed a rare exception to the general sociodemographic trend.

Labov (1990:218-219) discusses the minor tendency for males to be in the advance, remarking that all such cases involve “relatively isolated changes.” As discussed above, the development of LIKE reflects the gradual generalization of this form across functional projections as well as spread from the higher functional to the lower lexical X^{\max} within a domain. In other words, it is systemic, not isolated. Thus, not only is LIKE exceptional in being favoured by males, it is also noteworthy in that it involves the full extent of syntactic structure.

LIKE is exceptional in other ways as well. It is not prestigious and it draws negative social commentary, but it is precisely the overt attention it garners that disqualifies it as change from below, since change from below progresses with little or no social awareness (Labov 1990:215). For these reasons we should not expect to find females in the lead. Yet, the cases where males typically display higher frequencies of use are with stable, stigmatized variants (1990:210). The rising frequency of LIKE in Figure 8.2 rules this out as well. In short, LIKE does not fit into the typical categories of linguistic variability, whether the variation signals change or not.

Add to these observations the fact that in the case of male-dominated changes, females do not abstain from using the incoming forms, but tend instead to follow closely behind males (Labov 2001:460-461). This pattern is visible for each of the clause-internal contexts examined here: DP, PredAP, and \square P (Figure 5.1, Chapter 5; Figure 6.2, Chapter 6; Figure

7.4, Chapter 7). Labov (2001:460-461) suggests a number of ways in which women may follow behind at a lower level (e.g., they may have a slower rate of change or a lower limit on the amount of change in a given year, or they may adopt the change at a later age). Factors such as these account for lower frequencies of use. They were not intended, however, to explain why women should adopt a form that receives such overt negative evaluation within the community. As Labov (1990:240) points out, women are the chief agents of differentiation not only because they adopt prestige features more rapidly than men do, but also because they react more sharply against the use of stigmatized variants. It is therefore conspicuous that multivariate analysis consistently reveals speaker sex to be the weakest factor. Thus, while males lead in the move toward LIKE, the females are not far behind. This seems to corroborate Dubois (1992:198-199), who suggests that proportional variation in the use of discourse-level features is “at most weakly correlated with sociodemographic factors.” Here, we find a *bona fide* sex effect, but it is the least robust of the factors that operate on LIKE.

The paradox in this is that the data do not manifest the overt social evaluation that associates LIKE with females (Dailey-O’Cain 2000; Buchstaller 2003). In other words, the probabilistic patterns do not match the social attitudes. Why not? A distinct possibility is that LIKE is evaluated as a single entity, when in fact there is a complex juxtaposition of forms: quotative be like, marker LIKE, and particle LIKE, as well as the use of LIKE as an approximative adverb. Each is functionally distinct, but they are homophonous. Moreover, although some have been around for quite some time (e.g., adverb, discourse marker), the critical developmental period for each appears to have occurred in the past 15 to 20 years.

Thus, if we consider each form individually, we observe a recurrent pattern in which LIKE is most strongly associated with speakers below the age of 30. For example, we saw in Chapter 2 that the point where LIKE replaces about is between the 30-59 year olds, who use about more frequently, and the 17-29 year olds, where LIKE is supreme. For the CP context, where LIKE is a marker, a favouring factor weight is first attained for the 17-29 year olds (see Table 4.4). Similarly, for the DP, PredAP, and □P contexts, where it is a particle, the 17-29 year olds are significantly differentiated from the 30-59 year olds, whether this is evident through inputs (DP, Table 5.5), or factor weights (PredAP, Table 6.3; □P, Table 7.6). Finally, Tagliamonte and D’Arcy (under review) demonstrate that it is among 25-29 year olds that be like breaks away and becomes the most frequent form within the quotative paradigm. In sum, the data suggest that each LIKE has recently undergone a period of vigorous development. This has likely increased the saliency of LIKE as a lexeme in the vernacular. Combined with the fact that they all sound the same, each use probably resounds in the ears of members of the speech community simply as ‘LIKE’. That only two of the forms — the quotative and the discourse marker — are in fact favoured by females, while the adverb exhibits no social conditioning and the particle is favoured by males, is thus moot. In short, I would like to suggest that the social evaluation of LIKE reflects a folk belief that LIKE is just LIKE (and not LIKE_{adverb}, LIKE_{marker}, LIKE_{particle}, be like_{quotative}). As a result, the overt social associations, which lump all forms together, do not match the covert probabilistic patterns in the data that emerge when the particle is considered distinct from the others.

8.3 CLOSING REMARKS

“In high school and that I never learned, like, grammar.”

Interviewer, IN-TO-VATION Project, Toronto 2004

As discussed in the first chapter, the vast majority of research investigating LIKE has been framed within the tradition of discourse-pragmatics, both diachronic (Meehan 1991; Romaine & Lange 1991) and synchronic (e.g., Schourup 1983; Romaine & Lange 1991; Miller & Weinert 1995; Andersen 1997 et seq.; Buchstaller 2001; Hasund 2003; Sharifian & Malcolm 2003). This body of work has yielded insights into the pragmatic functions of this feature (e.g., metalinguistic focus, hesitation, ‘loose talk’, etc), and has suggested a number of avenues for research regarding the nature of social and linguistic constraints that may be operative.

The goal of the current investigation has been to examine LIKE from a different perspective — that provided by accountable variationist methodology — to determine the interaction of this feature with the syntactic component of the grammar. Following Andersen (2001:275), the hypothesis underlying this work has been that LIKE is not random, but that its use is constrained. In so doing, I have discovered that grammatical factors, syntactic structure in particular, are major constraints on its use. I addressed this issue by asking a series of questions such as: Can LIKE in fact occur “anywhere”? If not, are there structural considerations that govern the distribution of LIKE, and if so, what are they? Are these considerations constant across apparent-time or is there evidence for change in progress? Is there evidence that LIKE is rising in frequency and/or that it is spreading into new contexts?

A frequently discussed aspect of discourse features is their ‘loose attachment’ to syntactic structure (Brinton 1996:34). This characterization derives from their optionality, since they have no impact on the propositional content of an utterance and they do not affect grammaticality (Quirk 1972; Svartvik 1979; Brown & Yule 1983; Quirk et al. 1985; Schiffrin 1987; Fraser 1988). This holds true of LIKE. As an adjunct, it is not part of argument structure, nor does it interfere with basic operations such as merge or with more complex aspects of the syntax such as C-command, negative polarity, agreement, etc. In other words, a sentence such as (139a) is equally “good” whether we adjoin LIKE to any maximal projection contained within, as demonstrated in (139b), while a sentence such as (140a) remains bad regardless of LIKE-insertion, shown in (140b).³

- (139) a. Bob_i was thinking of himself_i.
 b. (LIKE) Bob_i was (LIKE) thinking (LIKE) of (LIKE) himself_i.
- (140) a. * Bob_i was thinking of him_i.
 b. * (LIKE) Bob_i was (LIKE) thinking (LIKE) of (LIKE) him_i.

However, to assume that discourse features such as LIKE function independent of the syntax is clearly false. As an adjunct, LIKE targets maximal projections. It therefore occurs at syntactically delimited positions of structure, interfacing with the syntax in regular and predictable ways. This was best demonstrated in Chapter 7, where we saw that by adjoining to the light verb, the position of LIKE in relation to elements hosted in T° such as auxiliaries, periphrastic ‘do’, and infinitival ‘to’ falls out from the lower position occupied by the particle. Moreover, this analysis captures the ordering of LIKE and adverbs, some of which

³ Though see Siegel (2002) for a discussion of sluicing constructions.

target projections above \square P, others, below. Thus, these results suggest that we must allow for the interaction of discourse particles with structural aspects of grammar.

What is most striking about LIKE is that in each context in which it appears (e.g., CP, DP, \square P, etc.), the same factors constraining its use among 10 year olds are also apparent in the speech of older sectors of the population who use this feature. There is therefore evidence for constancy of constraints (Kroch 1989a, 2001), a result which highlights the systematicity of LIKE, while the incremental increase in frequency across apparent-time indicates ongoing change (Labov 2001).

In sum, this research has demonstrated that a method that combines aspects of historical pragmatics, synchronic pragmatics, and variationist traditions permits observations that are not possible using just one approach. Indeed, with very few exceptions (e.g., verb type/subject type in Chapter 7), the analyses presented here were based on hypotheses gleaned from the existing literature. I have relied most heavily on the work of Andersen, and his 2001 book in particular, because it presents the most thorough analysis of LIKE in the pragmatic tradition.

Finally, the current investigation has corroborated previous work in demonstrating that, contrary to popular stereotypes, LIKE is not random; its use is not free (e.g., Underhill 1988; Andersen 1997 et seq.). Rather, it represents what Weinreich, Labov and Herzog (1968:100) refer to as “orderly heterogeneity.” The added insight of the analysis presented here is that through regular processes of language change, this feature has spread from one maximal projection to another, and this evolution has been guided by factors internal to the syntax. Indeed, the variable grammar is shared by the community of LIKE users. Thus, not

only have younger speakers not appropriated LIKE as a youthful crutch for lexical indecision (cf. Diamond 2000; Siegel 2002), neither do they use it with disregard to grammar, which I use in the technical sense outlined in Chapter 3 (§3.1). On the contrary, the model upon which their use is based is the one established in the vernaculars of successive generations of speakers. Frequencies of use may differ, but the constraints hold constant. In sum, we confirm yet again the axiomatic truth that lies at the very foundation of research into language variation and change: “The grammars in which linguistic change occurs are grammars of the speech community” (Weinreich, Labov & Herzog 1968:188).

References

- Abney, Steven. 1987. The English noun phrase in its sentential aspect. Doctoral dissertation, MIT.
- Adger, David. 2003. *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.
- Aijmer, Karin. 1988. 'Now may we have a word on this': The use of 'now' as a discourse particle. In M. Kytö, O. Ihalainen & M. Rissanen (eds.), *Corpus Linguistics, Hard and Soft: Proceedings of the Eighth International Conference on English Language Research on Computerized Corpora*. Amsterdam: Rodopi. 15-34.
- □ . 2002. *English Discourse Particles: Evidence from a Corpus*. Amsterdam and Philadelphia: John Benjamins.
- Alexiadou, Artemis. 1999. On the syntax of nominalization and possession: Remarks on patterns of ergativity. Doctoral dissertation, University of Potsdam.
- Allerton, David John and Alan Cruttenden. 1974. English sentence adverbials: Their syntax and their intonation in British English. *Lingua* 34:1-30.
- Altenberg, Bengt. 1987. *Prosodic Patterns in Spoken English. Studies in the Correlation Between Prosody and Grammar for Text-to-Speech Conversion*. Lund: Lund University Press.
- Altenberg, Bengt. 1990. Spoken English and the dictionary. In J. Svartvik (ed.), *The London-Lund Corpus of Spoken English: Description and Research*. Lund: Lund University Press. 193-211.
- Andersen, Gisle. 1997. 'They gave us these yeah, and they like wanna see like how we talk and all that': The use of 'like' and other discourse markers in London teenage speech. In U.-B. Kostinas, A.-B. Stenström & A.-M. Karlsson (eds.), *Ungdomsspråk i Norden*. MINS 43. Stockholm: Stockholms universitet, Institutionene för nordiska språk. 83-95.
- □ . 1998. The pragmatic marker 'like' from a relevance-theoretic perspective. In A. Jucker & Y. Ziv (eds.), *Discourse markers: Descriptions and theory*. Amsterdam and Philadelphia: John Benjamins. 147-170.
- □ . 2001. *Pragmatic Markers and Sociolinguistic Variation*. Amsterdam and Philadelphia: John Benjamins.

- Bailey, Charles-James. 1973. *Variation and Linguistic Theory*. Washington, DC: Center for Applied Linguistics.
- Becker, Misha. 2004. Is isn't be. *Lingua* 114:399-418.
- Beghelli, Filippo and Tim Stowell. 1997. Distributivity and negation: The syntax of 'each' and 'every'. In A. Szabolcsi (ed.), *Ways of Taking Scope*. Dordrecht: Kluwer. 29-71.
- Belletti, Adriana. 2001. Inversion as focalization. In A. Hulk & J. Pollock (eds.), *Inversion in Romance and the Theory of Universal Grammar*. Oxford: Oxford University Press. 60-90.
- Bernstein, Judy. 2001. The DP hypothesis: Identifying clausal properties in the nominal domain. In M. Baltin & C. Collins (eds.), *The Handbook of Contemporary Syntactic Theory*. Malden and Oxford: Blackwell.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finnigan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Blakemore, Diane. 1988. 'So' as a constraint on relevance. In R. M. Kempson (ed.), *Mental Representations: The Interface Between Language and Reality*. Cambridge: Cambridge University Press. 183-195.
- Blyth, Carl, Sigrid Recktenwald and Jenny Wang. 1990. I'm like, 'say what?!': A new quotative in American oral narrative. *American Speech* 65:215-227.
- Bopp, Franz. 1816. *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprachen*. Frankfurt: Andreäische Buchhandlung.
- Boskovic, Zeljko. 1994. Categorical status of null operator relatives and finite declarative complements. *Language Research* 30:387-417.
- Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24:591-656.
- □ . 2002. Transitivity. *Linguistic Inquiry* 33:183-224.
- Boyland, Joyce Tang. 1996. Morphosyntactic change in progress: A psycholinguistic approach. Doctoral dissertation, University of California.
- Brinton, Laurel. 1996. *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin and New York: Mouton de Gruyter.

- □ . 2005. Processes underlying the development of pragmatic markers: The case of '(I) say'. In J. Skaffari, M. Peikola, R. Carroll, R. Hiltunen & B. Wårvik (eds.), *Opening Windows on Texts and Discourses of the Past*. Amsterdam and Philadelphia: John Benjamins. 279-299.
- □ . forthcoming. Pathways in the development of pragmatic markers in English. In A. van Kemenade & B. Los (eds.), *Handbook on the History of the English Language*. London: Blackwell.
- Brown, Gillian. 1977. *Listening to Spoken English*. London: Longman.
- Brown, Gillian and George Yule. 1983. *Discourse Analysis*. Cambridge: Cambridge University Press.
- Buchstaller, Isa. 2001. An alternative view of 'like': Its grammaticalization in conversational American English and beyond. *Edinburgh Working Papers in Applied Linguistics* 11:21-41.
- □ . 2003. Putting perception to the reality test: The case of 'go' and 'like'. Paper presented at NWA 32, Philadelphia, University of Pennsylvania.
- □ . 2004. The sociolinguistic constraints on the quotative system: British English and US English compared. Doctoral dissertation, University of Edinburgh.
- Burzio, Luigi. 1986. *Italian Syntax: Studies in Natural Language and Linguistic Theory*. Dordrecht: Reidel.
- Butters, Ronald. 1982. Editor's note to Schourup (1982). *American Speech* 57:149.
- Bybee, Joan. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.
- □ . 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14:261-290.
- □ . 2003. Mechanism of change in grammaticalization: The role of frequency. In B. Joseph & R. Janda (eds.), *The Handbook of Historical Linguistics*. Malden: Blackwell. 602-623.
- Bybee, Joan and Paul Hopper (eds.). 2001. *Frequency and the Emergence of Linguistic Structure*. Amsterdam and Philadelphia: John Benjamins.
- Bybee, Joan and William Pagliuca. 1985. Cross-linguistic comparison and the development of grammatical meaning. In J. Fisiak (ed.), *Historical Semantics, Historical Word-Formation*. Berlin: Mouton de Gruyter. 59-83.
- Bybee, Joan, Revere Perkins and William Pagliuca. 1994. *The Evolution of Grammar: Tense, aspect and modality in the Languages of the World*. Chicago: University of Chicago Press.

- Cameron, Deborah, Fiona McAlinden and Kathy O’Leary. 1989. Lakoff in context: The social and linguistic functions of tag questions. In J. Coates & D. Cameron (eds.), *Women in their Speech Communities*. London: Longman. 74-93.
- Cedergren, Henrietta. 1984. Panama revisited: Sound change in real time. Paper presented at NWAV 12, Philadelphia, University of Pennsylvania.
- Chambers, J.K. 1992. Linguistic correlates of gender and sex. *English World-Wide* 13:173-218.
- □ . 1995. The Canada-U.S. border as a vanishing isogloss: The evidence of ‘chesterfield’. *Journal of English Linguistics* 23:155-166.
- □ . 2000. Region and language variation. *English World-Wide* 21:169-199.
- □ . 2002. Patterns of variation including change. In J.K. Chambers, P. Trudgill & N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change*. Malden: Blackwell. 349-372.
- □ . 2003. *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. 2nd edition. Oxford and Malden: Blackwell.
- Chambers, J.K. and Peter Trudgill. 1998. *Dialectology*. 2nd edition. Cambridge: Cambridge University Press.
- Cheshire, Jenny. 2002. Who we are and where we’re going: Language and identities in the new Europe. In M. Holt & P. Gubbins (eds.), *Beyond Boundaries: Language and Identity in Contemporary Europe*. Clevedon: Multilingual Matters. 19-34.
- Chomsky, Noam. 1971. Deep structure, surface structure, and semantic interpretation. In D. Steinberg & L. Jakobovits (eds.), *Semantics. An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*. New York: Cambridge University Press.
- □ . 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- □ . 1995. *The Minimalist Program*. Cambridge: MIT Press.
- □ . 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels & J. Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honour of Howard Lasnik*. Cambridge: MIT Press. 89-115.
- □ . 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*. Cambridge: MIT Press. 1-52.

- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- □ . 2004. Issues in adverbial syntax. *Lingua* 114:683-710.
- Coopmans, P. 1989. Where stylistic and syntactic processes meet: Locative inversion in English. *Language* 65:728-751.
- Cukor-Avila, Patricia. 2002. 'She say, she go, she be like': Verbs of quotation over time in African American Vernacular English. *American Speech* 77:3-31.
- Dailey-O'Cain, Jennifer. 2000. The sociolinguistic distribution of and attitudes toward focuser 'like' and quotative 'like'. *Journal of Sociolinguistics* 4:60-80.
- D'Arcy, Alexandra. 2005. The development of linguistic constraints: Phonological innovations in St. John's English. *Language Variation and Change* 17:327-355.
- Diamond, S.J. 2000. Like it or not, 'like' is probably here to stay. *Los Angeles Times*. August 21:2.
- van Dijk, Teun. 1982. Episodes as units of discourse analysis. In D. Tannen (ed.), *Spoken and Written Language: Exploring Orality and Literacy*. Norwood, NJ: Ablex. 177-195.
- Dines, Elizabeth. 1980. Variation in discourse 'and stuff like that'. *Language in Society* 9:13-31.
- Dubois, Sylvie. 1992. Extension particles, etc. *Language Variation and Change* 4:179-203.
- □ . 1995. Structural processes in enumeration. *Language Variation and Change* 7:113-137.
- □ . 1997. *L'analyse variationniste du discours en sociolinguistique: l'approche modulaire pour décrire l'usage et la formation des procédés discursifs*. American University Studies XIII. New York: Peter Lang.
- Dubois, Sylvie and Isabel Crouch. 1975. The question of tag questions in women's speech: They don't really use more of them, do they? *Language in Society* 4:289-294.
- Dubois, Sylvie and David Sankoff. 2001. The variationist approach toward discourse structural effects and socio-interactive dynamics. In D. Schiffrin, D. Tannen and H. Hamilton (eds.), *The Handbook of Discourse Analysis*. Malden and Oxford: Blackwell. 252-303.
- Eckert, Penelope. 1988. Adolescent social structure and the spread of linguistic change. *Language in Society* 17:183-207.

- □ . 2000. *Language Variation as Social Practice*. Oxford: Blackwell.
- □ . 2003. Language and gender in adolescence. In J. Holmes & M. Meyerhoff (eds.), *The Handbook of Language and Gender*. Oxford: Blackwell. 381-400.
- Elsness, Johan. 1997. *The Perfect and the Preterite in Contemporary and Earlier English*. Berlin and New York: Mouton de Gruyter.
- Erman, Brit. 1986. Some pragmatic expressions in English conversation. In G. Tottie & I. Bäcklund (eds.), *English in Speech and Writing: A Symposium*. Stockholm: Almqvist and Wiksell. 131-147.
- □ . 1987. *Pragmatic Expression in English: A Study of 'you know', 'you see' and 'I mean' in Face-to-Face Conversation*. Stockholm: Almqvist and Wiksell.
- □ . 1992. Female and male usage of pragmatic expressions in same-sex and mixed-sex interaction. *Language Variation and Change* 4:217-234.
- Ernst, Thomas. 1984. *Toward an Integrated Theory of Adverb Position in English*. Indiana University Linguistics Club.
- Ernst, Thomas. 2002. *The Syntax of Adjuncts*. Cambridge: Cambridge University Press.
- Even-Zohar, Itamar. 1982. The emergence of speech organizers in a renovated language: the case of Hebrew void pragmatic connectives. In N.E. Enkvist (ed.), *Impromptu Speech: A Symposium*. Åbo: Åbo Akademi. 179-193.
- Faerch, Claus and Gabriele Kasper. 1982. Phatic, metalingual and metacommunicative functions in discourse: Gambits and repairs. In N.E. Enkvist (ed.), *Impromptu Speech: A Symposium*. Åbo: Åbo Akademi. 71-103.
- Ferrara, Kathleen and Barbara Bell. 1995. Sociolinguistic variation and discourse function of constructed dialogue introducers: The case of be + like. *American Speech* 70:265-290.
- Fidelholtz, James. 1975. Word frequency and vowel reduction in English. *Chicago Linguistics Society* 11:200-213.
- Fox Tree, Jean. 1995. The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language* 34:709-738.

- Fleischman, Suzanne. 1982. *The Future in Thought and Language: Diachronic Evidence from Romance*. Cambridge: Cambridge University Press.
- Fraser, Bruce. 1988. Types of English discourse markers. *Acta Linguistica Hungaria* 38:19-33.
- □ . 1990. An approach to discourse markers. *Journal of Pragmatics* 14:383-395.
- □ . 1996. Pragmatic markers. *Pragmatics* 6:167-190.
- Gabelentz, George von der. 1891. *Die Sprachwissenschaft. Ihre Aufgaben, Methoden, und bisherigen Ergebnisse*. Leipzig: Weigel.
- Givón, Talmy. 1971. Historical syntax and synchronic morphology: An archaeologist's field trip. *Chicago Linguistic Society* 7:394-415.
- □ . 1973. The time-axis phenomenon. *Language* 49:890-925.
- Grant, William and James Main Dixon. 1921. *Manual of Modern Scots*. Cambridge: Cambridge University Press.
- Guy, Greg. 1988. Advanced VARBRUL analysis. In K. Ferrara, B. Brown, K. Walters & J. Baugh (eds.), *Linguistic Change and Contact*. Austin: University of Texas at Austin. 124-136.
- Haiman, John. 1994. Ritualization and the development of language. In W. Pagliuca (ed.), *Perspectives on Grammaticalization*. Amsterdam and Philadelphia: John Benjamins. 3-28.
- Hale, Ken and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In K. Hale & S.J. Keyser (eds.), *The View from Building 20: Essays in Honor of Sylvain Bromberger*. Cambridge: MIT Press.
- □ and □ □ . 2002. *Prolegomenon to a Theory of Argument Structure*. Cambridge: MIT Press.
- Halliday, M.A.K. 1967. Notes on transitivity and theme in English, part 2. *Journal of Linguistics* 3:199-244.
- Han, Chung-Hye. 2000. *The Structure and Interpretation of Imperatives*. New York and London: Garland.
- Harley, Heidi and Rolf Noyer. 1998. Licensing in the non-lexicalist lexicon: Nominalizations, vocabulary items and the encyclopedia. *MIT Working Papers in Linguistics* 32:119-137.

- Hasund, Ingrid Kristine. 2003. The discourse markers 'like' in English and 'liksom' in Norwegian teenage language: A corpus-based, cross-linguistic study. Doctoral dissertation, University of Bergen and Agder University College.
- Heine, Bernd. 2003. Grammaticalization. In B. Joseph & R. Janda (eds.), *The Handbook of Historical Linguistics*. Malden: Blackwell. 575-601.
- Heine, Bernd and Mechtild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske.
- Heine, Bernd, Ulrike Claudi and Friederike Hünemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: Chicago University Press.
- Hoekstra, Teun. 1984. *Transitivity: Grammatical Relations in Government-Binding Theory*. Dordrecht: Foris.
- Hoekstra, Teun and R. Mulder. 1990. Unergatives as copular verbs: Locational and existential predication. *The Linguistic Review* 7:1-79.
- Hölker, Klaus. 1991. Fransösisch: Partikelforschung. In *Lexikon der Romantischen Linguistik*, Vol. 5. Tübingen: Niemeyer. 77-88.
- Holmes, Janet. 1984. Hedging your bets and sitting on the fence: Some evidence for hedges as support structures. *Te Reo* 27:47-62.
- □ . 1986. Function of 'you know' in women's and men's speech. *Language in Society* 15:1-22.
- □ . 1995. *Women, Men and Politeness*. London: Longman.
- Hook, Peter. 1974. *The Compound Verb in Hindi*. Ann Arbor: Center for South Asian Studies.
- □ . 1991. The emergence of perfective aspect in Indo-Aryan languages. In E. Traugott & B. Heine (eds.), *Approaches to Grammaticalization*, Vol. 2. Amsterdam and Philadelphia: John Benjamins. 59-89.
- Hooper, Joan. 1976. Word frequency in lexical diffusion and the source of morpho-phonological change. In W. Christie (ed.), *Current Progress in Historical Linguistics*. Amsterdam: North Holland. 95-105.
- Hopper, Paul. 1991. On some principles of grammaticalization. In E. Traugott & B. Heine (eds.), *Approaches to Grammaticalization*, Vol. 1. Amsterdam and Philadelphia: John Benjamins. 17-35.
- Hopper, Paul and Elizabeth Traugott. 2003. *Grammaticalization*. 2nd edition. Cambridge: Cambridge University Press.

- Horn, Laurence. 1980. Affixation and the unaccusative hypothesis. In J. Kreiman & A. Ojeda (eds.), *Papers from the 16th Regional Meeting of the CLS*. Chicago: Chicago Linguistic Society. 134-146.
- Horn, Merle, Petra Hansson, Gösta Bruce, Johan Frid, and Marcus Filipsson. 2001. Cue words and the topic structure of spoken discourse: The case of Swedish 'men' 'but'. *Journal of Pragmatics* 33:1061-1081.
- Hornstein, Norbert. 1995. *Logical Form*. Oxford: Blackwell.
- Humboldt, Wilhelm von. 1825. Über das Entstehen der grammatischen Formen und ihren Einfluss auf die Ideenentwicklung. *Abhandlungen der Königlich Akademie der Wissenschaften zu Berlin*: 401-430.
- Ito, Rika and Sali A. Tagliamonte. 2003. Well weird, right dodgy, very strange, really cool: Layering and recycling in English intensifiers. *Language in Society* 32:257-279.
- Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.
- □ . 1997. *The Architecture of the Language Faculty*. Cambridge: MIT Press.
- □ . 2002. *Foundations of Language*. Oxford: Oxford University Press.
- Jespersen, Otto. 1942. *A Modern English Grammar on Historical Principles. Part VI. Morphology*. Copenhagen: Ejnar Munksgaard.
- □ . 1974 [1933]. *Essentials of English Grammar*. London: George Allen and Unwin Ltd.
- Jucker, Andreas. 1993. The discourse marker 'well': A relevance-theoretical account. *Journal of Pragmatics* 19:435-452.
- Jucker, Andreas and S. Smith. 1998. 'And people just you know like 'wow''': Discourse markers as negotiating strategies. In A. Jucker & Y. Ziv (eds.), *Discourse Markers: Descriptions and Theory*. Amsterdam and Philadelphia: John Benjamins. 171-201.
- Kiparsky, Paul. 1995. Indo-European origins of Germanic syntax. In A. Battye & I. Roberts (eds.), *Clause Structure and Language Change*. Oxford: Oxford University Press. 140-169.
- Kotsinas, Ulla-Britt. 1994. *Ungdomsspråk*. Uppsala: Hallgren & Fallgren.
- Kroch, Anthony. 1989a. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199-244.
- □ . 1989b. Function and grammar in the history of English: Periphrastic do. In R. Fasold & D. Schiffrin (eds.), *Language Variation and Change*. Washington: Georgetown University Press. 133-172.

- □ . 2001. Syntactic change. In M. Baltin & C. Collins (eds.), *The Handbook of Contemporary Syntactic Theory*. Malden: Blackwell. 699-729.
- Krug, Manfred. 2000. *Emerging English Modals: A Corpus-based Study of Grammaticalization*. Berlin: Mouton de Gruyter.
- Laberge, Suzanne. 1980. The changing distribution of indefinite pronouns in discourse. In R. Shuy & A. Shnukal (eds.), *Language Use and the Uses of Language*. Washington: Georgetown University Press. 76-87.
- Labov, William. 1963. The social motivation of a sound change. *Word* 19:273-309.
- □ . 1966. *The Social Stratification of English in New York City*. Washington: Center for Applied Linguistics.
- □ . 1972. *Sociolinguistic Patterns*. Philadelphia: University of Philadelphia Press.
- □ . 1973. The meaning of words and their boundaries. In C. Bailey & R. Shuy (eds.), *New Ways of Analyzing Variation in English*. Washington: Georgetown University Press. 340-373.
- □ . 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2:205-251.
- □ . 1994. *Principles of Linguistic Change: Internal Factors*. Oxford and Cambridge: Blackwell.
- □ . 2001. *Principles of Linguistic Change: Social Factors*. Oxford and Malden: Blackwell.
- Labov, William and Joshua Waletzky. 1967. Narrative analysis: Oral versions of personal experience. *Journal of Narrative and Life History* 7:3-38.
- Lakoff, Robin. 1973a. Language and woman's place. *Language in Society* 2:45-80.
- □ . 1973b. Questionable answers and answerable questions. In B. Kachru (ed.), *Papers in Honor of Henry and Renee Kahane*. Urbana: University of Illinois Press. 453-467.
- Langacker, Ronald. 1990. Subjectification. *Cognitive Linguistics* 1:5-38.
- Lauwereyns, Shizuka. 2002. Hedgges in Japanese conversation: The influence of age, sex, and formality. *Language Variation and Change* 14:239-259.
- Lehti-Eklund, Hanna. 1990. *Från adverb till markör i text: Studier i semantisk-syntatisk utveckling i äldre svenska*. Helsingfors: Svenska litteratursällskapet.

- Lehmann, Christian. 1982. *Thoughts on Grammaticalization: A Programmatic Sketch*, Vol. 1. Köln: Universität zu Köln, Institut für Sprachwissenschaft.
- □ . 1985. Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile* 20:303-318.
- Lenker, Ursula. 2000. 'Soðlice' and 'witodlice': Discourse markers in Old English. In O. Fischer, A. Rosenbach & D. Stein (eds.), *Pathways of Change: Grammaticalization in English*. Amsterdam and Philadelphia: John Benjamins. 229-249.
- Levey, Stephen. 2004. The sociolinguistic distribution of 'like' in preadolescents. Queen Mary College, University of London, ms.
- Levin, Beth. 1993. *English Verb Classes and Alternations*. Chicago and London: The University of Chicago Press.
- Levin, Beth and Malka Rappaport. 1986. The formation of adjectival passives. *Linguistic Inquiry* 17:623-661.
- Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity. At the Syntax-Lexical Semantics Interface*. Cambridge, MA: MIT Press.
- Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Macauley, Ronald. 1977. *Language, Social Class, and Education: A Glasgow Study*. Edinburgh: The University Press.
- □ . 2002. Discourse variation. In J.K. Chambers, P. Trudgill & N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change*. Malden: Blackwell. 283-305.
- Marantz, Alec. 1997. No escape from syntax: Don't try a morphological analysis in the privacy of your own lexicon. MIT, ms.
- Maschler, Yael. 2001. On the grammaticalization of of ke'ilu 'like' in Hebrew talk-in-interaction. *Language in Society* 31:243-276.
- Massam, Diane. 2001. The status of subject in Niuean. In W. Davies & S. Dubinsky (eds.), *Objects and Other Subjects*. Amsterdam: Kluwer. 225-246.

- Matsumoto, Yo. 1988. From bound grammatical markers to free discourse markers: History of some Japanese connectives. In S. Axmaker, A. Jaisser & H. Singmaster (eds.), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society. 340-351.
- Meehan, Theresa. 1991. It's like, "What's happening in the evolution of 'like'?: A theory of grammaticalization. *Kansas Working Papers in Linguistics* 16:37-51.
- Meillet, Antoine. 1948 [1912]. *Linguistique Historique et Linguistique Générale*. Paris: Édouard Champion.
- Meyerhoff, Miriam. 2002. All the same? The emergence of complementizers in Bislama. In T. Güldemann & M. von Roncador (eds.), *Reported Discourse: A Meeting Ground for Different Linguistic Domains*. Amsterdam and Philadelphia: John Benjamins. 341-359.
- Meyerhoff, Miriam and Nancy Niedzielski. 1998. The syntax and semantics of 'olsem' in Bislama. In M. Pearson (ed.), *UCLA Occasional Papers in Linguistics* 20:235-243.
- Miller, Jim and Regina Weinert. 1995. The function of 'like' in dialogue. *Journal of Pragmatics* 23:365-393.
- Munn, Alan. 1993. Topics in the syntax and semantics of coordinate structures. Doctoral dissertation, University of Maryland.
- OED = Simpson, J.A. and E.S.C. Weiner (eds.). *The Oxford English Dictionary*. 2nd edition. Oxford: Clarendon Press.
- Onodera, Noriko. 1995. Diachronic analysis of Japanese discourse markers. In A. Jucker (ed.), *Historical Pragmatics: Pragmatic Developments in the History of English*. Amsterdam: Benjamins. 393-437.
- Östman, Jan-Ola. 1981. *'You know': A Discourse Functional View*. Amsterdam and Philadelphia: John Benjamins.
- □ . 1982. The symbiotic relationship between pragmatic particles and impromptu speech. In N.E. Enkvist (ed.), *Impromptu Speech: A Symposium*. (Meddelanden från Stiftelsens för Åbo Akademi Forskningsinstitut 78.) Åbo: Åbo Akademi. 147-177.
- Owen, M. L. 1983. *Apologies and Remedial Exchanges*. The Hague: Mouton de Gruyter.
- Pei, Mario. 1973. *Double-speak in America*. New York: Hawthorn.

- Perlmutter, David. 1978. Impersonal passives and the unaccusative hypothesis. In *BLS 4: Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society. 157-189.
- Pesetsky, David. 1982. *Paths and categories*. Doctoral dissertation, MIT.
- Phillips, Betty. 1980. Old English an ~ on: A new appraisal. *Journal of English Linguistics* 14:20-23.
- Pintzuk, Susan. 1993. Verb seconding in Old English: Verb movement to Infl. *The Linguistic Review* 10:5-35.
- □ . 1996. Cliticization in Old English. In A. Halpern & A. Zwicky (eds.), *Approaching Second*. Stanford: CSLI Publications. 375-409.
- □ . 1999. *Phrase Structures in Competition: Variation and Change in Old English Clause Structure*. New York: Garland.
- □ . 2003. Variationist approaches to syntactic change. In B. Joseph & R. Janda (eds.), *The Handbook of Historical Linguistics*. Malden and Oxford: Blackwell. 509-528.
- Pintzuk, Susan and Anthony Kroch. 1989. The rightward movement of complements and adjuncts in the Old English of *Beowulf*. *Language Variation and Change* 1:115-144.
- Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20:365-424.
- Poplack, Shana and Sali A. Tagliamonte. 2000. The grammaticalization of 'going to' in (African American) English. *Language Variation and Change* 11:315-342.
- □ and □ □ . 2001. *African American English in the Diaspora: Tense and Aspect*. Malden: Blackwell.
- Quirk, Randolph. 1972. *The English Language and Images of Matter*. London: Oxford University Press.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London and New York: Longman.
- Rand, David and David Sankoff. 1990. *GoldVarb: A Variable Rule Application for the Macintosh. Version 2*. Montréal, Canada: Centre de recherches mathématiques, Université de Montréal.
- Redeker, Gisela. 1990. Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics* 14:367-381.

- Reinhart, Tanya. 1978. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20:335-397.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*. Dordrecht: Kluwer. 281-337.
- □ . 2001. On the position Int(errogative) in the left periphery of the clause. In G. Cinque and L. Salvi (eds.), *Current Studies in Italian Syntax*. Amsterdam and London: Elsevier. 287-296.
- □ . 2002. Locality and left periphery. In A. Belletti (ed.), *Structures and Beyond. The Cartography of Syntactic Structures*, Vol. 3. Oxford: Oxford University Press.
- Roberts, Ian. 1998. Have/Be raising, move F, and procrastinate. *Linguistic Inquiry* 29:113-125.
- Roberts, Ian and Anna Roussou. 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge and New York: Cambridge University Press.
- Roberts, Julie. 1994. Acquisition of variable rules: (-t,d) deletion and (ing) production in preschool children. Doctoral dissertation, University of Pennsylvania.
- □ . 1997. Hitting a moving target: Acquisition of sound change in progress by Philadelphia children. *Language Variation and Change* 9:249-266.
- Roberts, Julie, and William Labov. 1995. Learning to talk Philadelphian: Acquisition of short *a* by preschool children. *Language Variation and Change* 7:101-112.
- Rochemont, Michael and Peter Culicover. 1990. *English Focus Constructions and the Theory of Grammar*. New York: Cambridge University Press.
- Romaine, Suzanne and Deborah Lange. 1991. The use of 'like' as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech* 66:227-279.
- Ross, J.R. and W.E. Cooper. 1979. Like, syntax. In W.E. Cooper & E.C.T. Walker (eds.), *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*. Hillsdale, NJ: Erlbaum. 343-418.
- Sacks, Harvey, Emanuel Schegloff, and Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking in conversation. *Language* 50:696-735.

- Sankoff, David. 1982. Sociolinguistic method and linguistic theory. In L. Cohen, J. Los, H. Pfeiffer & K. Podewski (eds.), *Logic, Methodology and Philosophy of Science*, Vol. 6. Amsterdam: North-Holland. 677-689.
- □ . 1987. Variable rules. In A. Ulrich, D. Norbert & K. Mattheier (eds.), *Sociolinguistics: An International Handbook of the Science of Language and Society*, Vol. 2. Berlin: Mouton de Gruyter. 984-997.
- □ . 1988. Sociolinguistics and syntactic variation. In F. Newmeyer (ed.), *Linguistics: The Cambridge Survey*. Cambridge: Cambridge University Press. 140-161.
- Sankoff, David and Pierrette Thibault. 1981. Weak complementarity: Tense and aspect in Montreal French. In B. Johns & D. Strang (eds.), *Syntactic Change*. Ann Arbor: University of Michigan Press. 205-215.
- Sankoff, David, Pierette Thibault and H. Bérubé. 1978. Semantic field variability. In D. Sankoff (ed.), *Linguistic Variation: Models and Methods*. New York: Academic Press. 23-43.
- Sankoff, Gillian and Penelope Brown. 1976. The origins of syntax in discourse. *Language* 52:631-666.
- Sankoff, Gillian, Pierette Thibault, Naomi Nagy, H el ene Blondeau, Marie-Odile Fonollosa and Lucie Gagnon. 1997. Variation in the use of discourse markers in a language contact situation. *Language Variation and Change* 9:191-218.
- Santorini, Beatrice. 1993. The rate of phrase structure change in the history of Yiddish. *Language Variation and Change* 5:257-283.
- Schiffirin, Deborah. 1986. Functions of 'and' in discourse. *Journal of Pragmatics* 10:41-66.
- □ . 1987. *Discourse Markers*. Cambridge: Cambridge University Press.
- □ . 1992. Anaphoric 'then': Aspectual, textual and epistemic meaning. *Linguistics* 30:753-792.
- Schourup, Lawrence. 1983. *Common Discourse Particles in English Conversation*. Ohio State Working Papers in Linguistics 28.
- □ . 1999. Discourse markers. *Lingua* 107:227-265.
- Sharifian, Farzad and Ian Malcolm. 2003. The pragmatic marker 'like' in English teen talk: Australian Aboriginal usage. *Pragmatics and Cognition* 11:327-344.
- Siegel, Muffy. 2002. 'Like': The discourse particle and semantics. *Journal of Semantics* 19:35-71.

- Smith, Jennifer. 2000. Ye ø na hear that kind o' things: Negative 'do' in Buckie Scots. *English World-Wide* 21:231-259.
- Sperber, Dan and Deirdre Wilson. 1995. *Relevance. Communication and Cognition*. 2nd edition. Oxford: Blackwell.
- Strang, Barbara. 1963. *Modern English Structure*. New York: St. Martin's.
- Stubbe, Maria and Janet Holmes. 1995. 'You know', 'eh' and other "exasperating expressions": An analysis of social and stylistic variation in the use of pragmatic devices in a sample of New Zealand English. *Language and Communication* 15:63-88.
- Svartvik, Jan. 1979. 'Well' in conversation. In S. Greenbaum, G. Leech & J. Svartvik (eds.), *Studies in English Linguistics for Randolph Quirk*. London and New York: Longman. 167-177.
- Szabolsci, Anna. 2001. The syntax of scope. In M. Baltin & C. Collins (eds.), *The Handbook of Contemporary Syntactic Theory*. Malden and Oxford: Blackwell. 607-633.
- Tabor, Whitney and Elizabeth Closs Traugott. 1998. Structural scope expansion and grammaticalization. In A. Giacalone Ramat & P. Hopper (eds.), *The Limits of Grammaticalization*. Amsterdam and Philadelphia: John Benjamins. 229-272.
- Tagliamonte, Sali A. 2005. So who? Like how? Just what?: Discourse markers in the conversations of English speaking youth. In A.-B. Stenström & K. Aijmer (eds.), *Special Issue of the Journal of Pragmatics*.
- □ . to appear. *Roots of English: Exploring the History of Dialects*. Cambridge: Cambridge University Press.
- Tagliamonte, Sali A. and Alexandra D'Arcy. 2004. 'He's like, she's like': The quotative system in Canadian Youth. *The Journal of Sociolinguistics* 8:493-514.
- □ and □ □ . under review. Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change*.
- Tagliamonte, Sali A. and Rachel Hudson. 1999. Be like et al. beyond America: The quotative system in British and Canadian youth. *Journal of Sociolinguistics* 3:147-172.

- Tagliamonte, Sali A. and Jennifer Smith. 2002. Either it isn't or it's not: NEG/AUX contraction in British dialects. *English Word-Wide* 23:251-281.
- Tagliamonte, Sali A., Jennifer Smith and Helen Lawrence. 2005. No taming the vernacular! Insights from the relatives in northern Britain. *Language Variation and Change* 17:75-112.
- Taylor, Ann. 1994. The change from SOV to SVO in Ancient Greek. *Language Variation and Change* 6:1-37.
- Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In W. Lehmann & Y. Malkiel (eds.), *Directions for Historical Linguistics*. Austin: University of Texas Press. 245-271.
- □ . 1988. Pragmatic strengthening and grammaticalization. In S. Axmaker, A. Jaisser & H. Singmaster (eds.), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society. 406-416.
- □ . 1995. Subjectification in grammaticalization. In D. Stein & S. Wright (eds.), *Language, Subjectivity and Subjectivisation*. Cambridge: Cambridge University Press. 31-54.
- □ . 1997 [1995]. The role of the development of discourse markers in a theory of grammaticalization. Paper presented at the Twelfth International Conference on Historical Linguistics, University of Manchester.
- □ . 1997. The discourse connective 'after all': A historical pragmatic account. Paper presented at the Sixteenth International Congress of Linguists, Paris.
- □ . 1999. From subjectification to intersubjectification. Paper presented at the Workshop on Historical Pragmatics, Fourteenth International Conference on Historical Linguistics, Vancouver, Canada.
- □ . 2001. Legitimate counterexamples to unidirectionality. Paper presented at Freiburg University.
- □ . 2003. Constructions in grammaticalization. In B. Joseph & R. Janda (eds.), *The Handbook of Historical Linguistics*. Malden: Blackwell. 624-647.
- Traugott, Elizabeth Closs and Richard Dasher. 2002. *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

- Traugott, Elizabeth Closs and Bernd Heine (eds.). 1991. *Approaches to Grammaticalization*. Amsterdam and Philadelphia: John Benjamins.
- Traugott, Elizabeth Closs and Ekkehard König. 1991. The semantics-pragmatics of grammaticalization revisited. In E. Traugott & B. Heine (eds.), *Approaches to Grammaticalization*, Vol. 1. Amsterdam and Philadelphia: John Benjamins. 189-218.
- Trudgill, Peter. 2002. *Sociolinguistic Variation and Change*. Edinburgh: Edinburgh University Press.
- Underhill, Robert. 1988. Like is like, Focus. *American Speech* 63:234-246.
- Vincent, Diane. 1986. Que fait la sociolinguistique avec l'analyse du discours et vice-versa. *Language et Société* 38:7-17.
- Vincent, Diane and David Sankoff. 1992. Punctors: A pragmatic variable. *Language Variation and Change* 4:205-216.
- Wang, William S-Y. 1977. *The Lexicon in Phonological Change*. The Hague: Mouton de Gruyter.
- Watts, Richard. 1989. Taking the pitcher to the 'well': Native speakers' perception of their use of discourse markers in conversation. *Journal of Pragmatics* 13:203-237.
- Weinreich, Uriel, William Labov and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In W.P. Lehmann & Y. Malkiel (eds.), *Directions for Historical Linguistics*. Austin: University of Texas Press. 97-188.
- Wolgemuth, Tracy. 2003. It's like discourse 'like': The story of 'like' in Canadian English in 1995. MA thesis, University of Toronto.
- Wootton, A. 1981. The management of grantings and rejections by parents in request sequences. *Semiotica* 37:59-89.
- Wright, Joseph (ed.). 1902. *The English Dialect Dictionary*. London: Frowde.