

**The Sociolinguistic constraints  
on the Quotative System-  
British English and US English compared**

**Isabelle Buchstaller**

**A thesis submitted in fulfilment of requirements for the degree of  
Doctor of Philosophy**

**to  
Theoretical and Applied Linguistics  
School of Philosophy, Psychology and Language Sciences  
University of Edinburgh**

**March 2004**

## **Declaration**

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Isabelle Buchstaller

## Abstract

The recent advent and rapid spread of two new vernacular options, *go* and *like*, within the (say) variable has attracted a growing body of research in variationist sociolinguists. This thesis examines the synchronic functions of these new quotatives and considers pragmatic, discourse, and social factors. The investigation is based on an analysis of very large corpora of spontaneous spoken British and American English. This cross-variety comparison gives me the opportunity (i) to investigate a case of rapid language change that is happening concurrently with the time of research and (ii) to consider to what extent social and linguistic constraints hold globally.

A variationist study of the constraints which govern the quotative system is valuable for the following reasons: By investigating the patterning of the (say) variable as a whole, we gain insights into the rule-governed variability of innovative features and their rival variants (*say, tell, think, cry, ...*). A look at the entire quotative system reveals the intimate interplay of competing choices within the (say) variable. My project aims at understanding how the system as a whole reacts to the intrusion of newcomer variants. A sharply delimited and hitherto stable set of variants - such as the (say) variable - presents the unique opportunity to investigate the restructuring of all variants as new ones come in. This is especially interesting when we look at competing choices which have the same [-canonical] underlying semantic feature, such as unframed quotes. The data show that far from ousting the unframed or *say*-variant, *like* and *go* add options within the vernacular category.

A comparative study on the patterning of non-canonical variants within the quotative pool produces important insights into phenomena such as reallocation, competition within one socio-pragmatic field, as well as interaction of variants

within the same variable. In this light, the restructuring of all quotative variants during the rapid intrusion of the robust new variant *like* gives us an important test-case for various explanatory parameters for language change that have been raised in the literature, i.e. Romaine's claims about language and gender (2003), the reallocation hypothesis as raised by Britain (2002) and Britain and Trudgill (1999), the reformulation of the standard vs. non-standard dichotomy into local vs. supra-local variants by Foulkes and Docherty (1999) and the Milroys' findings (1998) concerning network and class.

Furthermore, research in variationist sociolinguistics has revealed the importance of intralinguistic constraints in situations of competing grammars (Meyerhoff 2000, Preston 1991, Rickford and MacNair-Knox 1994). Using a synthetic approach which looks at both intralinguistic and extralinguistic constraints, I attempt to explain the variability at all levels within the variable in order to account for as much of the variability as possible. Investigation into an array of linguistic factors reveals that linguistic constraints on the members of the (say) variable are indeed very important and quite robust.

A cross-variety comparison gives insights into how much we can generalize our findings: do locally separate systems handle the situation in the same way? The comparison of pragmatic and sociolinguistic factors reveals some interesting but subtle differences in *go*'s and *like*'s development in different locales. This raises more fundamental linguistic questions, such as are whether we are indeed comparing the same variable in the US and in Britain. My findings give evidence of the restructuring processes in the quotative system as a whole that accompany the arrival of a newcomer variant, *like*, in two varieties, British English and US English. There is evidence that different systems find idiosyncratic solutions to similar problems. This finding constitutes an important contribution to the growing body of research on globalisation phenomena and supra-local trends (Buchstaller 2003, Kerswill 2003, Meyerhoff and Niedzielski 2003, Tagliamonte and Hudson 1999, Trudgill 1983, 1994, and many others).

## Acknowledgements

As my postgraduate years have been somewhat a journey around the world, I should first of all thank the University of Konstanz for raising the money to let me go to Hawai'i, which turned out to be the big bang for the genesis of this thesis.

At Konstanz, I am indebted to Elizabeth Couper-Kuhlen for raising my interest in variability and for including me in a research group where my fascination with CA and grammaticalisation was kindled. I also owe a great deal to Aldo DiLuzio, who was my first source of inspiration in the field of creolistics and who sent me out to dig my teeth into the Hawaiian Pacific collection.

In Hawai'i, I want to send a big thanks to Michael Forman who actually raised the idea in me to go ahead and do a PhD, at the time in Creolistics. Mahalo to Mike, Wendy Onishi and Anatole Lyovin for their support and their open ear to any problems a newly arrived student might have who radically changes her PhD (topic, location, everything) within 6 weeks. Which brings me to my "Doktormutter" Miriam Meyerhoff.

Miriam not only helped me realise and continue to believe (in spite of whatever anybody else may say) that *like* is really worth a PhD. She also persuaded me that Scotland is a great place to live in. She succeeded in talking my funding body into transferring my scholarship from the US to the UK and the University of Edinburgh into accepting me. This thesis owes so much to the phenomenal knowledge and insight she has had the kindness to share with me. She has had the patience, the sharp wit and the warmth to listen, kick ass and give a cry shoulder whenever any of these were necessary. She helped me find my track and has given me a home, intellectually and topographically. I am certainly the PhD student who can say that without my supervisor, I would not be where I am now (God knows where I'd be). Thank you so much, Miriam.

I would also like to express my sincerest gratitude to my other supervisor, Caroline Heycock, to whom I owe a great deal of different perspective readings of my output. Especially the statistical methodology section would not be what it is without her. Thank you for putting fingers on wound spots and helping to find solutions. At the linguistics department here, John Joseph and Hugh Trappes-Lomax,

Dharshi Santhakumaran and Heather Hewitt have bounced off ideas with me. Thanks also go to Heather for doing all the proofreading so very swiftly and diligently and for providing me with loads of “odd examples”. I think I now know all the 80 year olds in Britain who are using *like* in spite of their age. I also owe a lot to all the other members of the ‘Language in Context’ research group who shared ideas with me and who patiently listened to many many talks of mine. I want to especially thank Therese Lindstrom for numerous discussions on grammaticalisation and for being a great friend. And thanks go to John Victor Singler, my overseas mentor in things *like*, who was a constant and wise discussion partner every year again at the NWAVE.

I am furthermore obliged to the DAAD for sponsoring me and to Rhetorical Systems and the Universities of Edinburgh and Hawai’i for making this thesis financially possible. This applies especially to Cassie Mayo and Alice Turk, who have given me the categorical perception experiments to run.

The University of Edinburgh has been a very inspiring and accommodating research environment and what it lacked in Pacific charm, it certainly made up in competence and helpfulness. I would like to express my sincerest gratitude to all those people who have helped me grapple with the everyday pitholes – academic, administrative and emotional – of being a PhD student. Penny Earle, the good soul of the department. Eddie Dubourg for his advice and help in the weird and wonderful world of word processing. And Michael Bennet for best-humoured and most linguistically aware UNIX, data and web-support. Presiding over this are Morag and Cedric whose presence was only felt because things worked. Thanks for all that.

This thesis would not have been possible without the data that was I was very generously offered by incredibly kind people. The LDC team at the University of Pennsylvania in Philadelphia ([www ldc upenn edu](http://www ldc upenn edu)) have given me permission to access their files. They have set up whole pages for me, the **un**subscribed user, and have been incredibly patient and swift in answering questions and responding to queries. Great job!

In search of a British English corpus and in total desperation, I came across two people who had the kindness to give me access to their data. Thank you very much to Joan Beal for offering the Newcastle Electronic Corpus of Tyneside English, which I have yet to mine. And an especially big thank you goes to Paul Foulkes. I do not

know what I would have done had Paul not offered the Derby and Newcastle corpus (cf. Chapter 2). He and Ghada opened up their house for me and made the endless hours of data mining a pleasant enterprise. Thank you for a wonderful time in York and for a great contribution to my thesis.

At this point I would also like to express my sincerest gratitude to all those subjects whose speech I have used and whom I have never been able to thank personally. Without their contribution, research such as this would just not be possible. Thanks also to the group of Ling2 students who helped me set up the Language Attitude Questionnaire and to all the 180 or so informants who took the time to answer it.

Apart from academia, there are very important people whose amicable support throughout this project has made life just so much more fun. There are my parents and my little brother, who have given me unwaivering support and a steady “ich weiss ganz genau, dass Du das schaffst”-kind of attitude. Thanks so much for believing in me and for providing the happy and antidotal environment where I could unwind entangled minds and regain sanity. What would I have been without such a home? I would further like to thank all those wonderful people in the lab and in the common room – too many to give names – who have given me collegial spirit and friendship and have made so much more than a working place out of the university. I have to mention the Edinburgh Tango Society ([www.edinburghtango.org.uk](http://www.edinburghtango.org.uk)), where I could leave analytical minds behind and let latino spirit grip my heart. The tango people up here are not only fabulous dancers but they are also brilliant people, who create a great family atmosphere. There is also nothing more relaxing than indulge in passionate rhythms and dance tango after a day of linguistics! Thanks to all of you for keeping me fit and mentally healthy.

I owe an enormous portion of my emotional wellbeing to Wojtek Skut, who has given me a wonderful life outside my thesis and who has constantly reinforced my sometimes wavering belief that I, too, can do it. He has stood by my side, has listened without tiring, asked important questions, and was always ready to discuss quotatives, linguistic theory or life in general. He believed in me and cooked for me and has given me all the support in the world. Kocham cię.

# List of Figures

2.1	Multiply embedded spaces as of Fauconnier and Sweetser (1996) . . . . .	23
3.1	The Hypotheticality Continuum . . . . .	45
3.2	Co-occurrence of contextual factors with degrees of hypotheticality . . . . .	55
4.1	Different conceptualisations of the grammaticalisation process . . . . .	87
4.2	Romaine and Lange's (1991) Model . . . . .	91
4.3	The Radial Structure Model of <i>LIKE</i> . . . . .	95
4.4	The functional extensions underlying <i>like</i> 's semantic core field . . . . .	133
4.5	The radial structure model of <i>go</i> . . . . .	139
4.6	The metaphorical base for <i>go</i> as an introducer of reported speech . . . . .	158
4.7	The radial structure model of <i>go</i> . . . . .	162
4.8	The substructures of <i>go</i> 's radial structure model . . . . .	163
5.1	The metaphorical base for <i>go</i> 's use as a quotative . . . . .	183
5.2	100-year progress of a change with incrementation (from Labov 2001:450) . .	185
6.1	100-year progress of a change with incrementation (from Labov 2001:450) . .	205
6.2	Distribution of canonical and non-canonical variants per gender . . . . .	217
6.3	Distribution of <i>like</i> -ratio over numerical age . . . . .	225



6.4	Distribution of <i>like</i> –frequency over age-groups	227
6.5	Incrementation of <i>like</i> -frequencies in 7 years	227
6.6	Distribution of <i>go</i> across age groups	229
6.7	<i>Go</i> 's patterning across age (data: Singler's NYC corpus)	232
6.8	Classical age grading pattern (simplified)	236
6.9	Generational change captured at two stages	237
6.10	Chronological development over several generations	239
6.11	Distribution of unframed quotes across age	243
6.12	Contrastive patterning of non-canonical and canonical (say) quotatives across age	246
6.13	A possible scenario for contact-induced <i>like</i> -adoption	252
6.14	Real-time comparison between say, <i>like</i> , <i>go</i> in 1994/5 and 2001 (data extrapolated from Johnson 2001) for the youngest age-group	256
6.15	Distribution of <i>go</i> across class and gender	263
6.16	Distribution of <i>like</i> across class and gender	263
6.17	Comparison of younger speakers' production of quotatives in 1994/5 (Derby and Newcastle), and 1996 and 2001 (York)	266
6.18	Distribution of <i>like</i> among young and old speakers in the Texas, York and Derby/Newcastle corpus	269
6.19	<i>Like</i> -ratio use per chronological age	270
6.20	Distribution of <i>like</i> per chronological age by brackets	271

6.21	Distribution of <i>go</i> by chronological age .....	274
6.22	Distribution of <i>go</i> -frequency by chronological age .....	275
6.23	Distribution of unframed as a ratio by chronological age .....	278
6.24	Patterning of non-canonical and <i>say</i> -variables across apparent time .....	280

# List of Tables

2.1	Frequencies of occurrence of variants of the (say) variable .....	39
3.1	ANOVA Post Hoc table (LSD) for the hypothetical categories .....	58
3.2	Correlation between <i>like</i> and <i>go</i> and degrees of hypotheticality .....	61
3.3	Distribution of degrees of hypotheticality as a percentage of all quotatives ...	61
3.4	Co-occurrence of <i>like</i> and <i>go</i> with mental/speech expressions (in %) .....	65
3.5	Co-occurrence of (say) variants with mental/speech expressions (in %) .....	67
3.6	Frequency of occurrence of quotative strategy per quotative .....	69
3.7	Frequency of occurrence of previous lexical quotative for <i>like</i> and <i>go</i> .....	70
3.8	Sequential triples of quotative verbs .....	71
3.9	Co-occurrence of 1s <i>go</i> and 1s <i>like</i> with preceding and following 3s <i>say</i> ....	72
3.10	Co-occurrence of quotatives with mimesis .....	75
3.11	ANOVA post hoc (Bonferroni) results for significances patterning with mimesis .....	76
4.1	Cross-linguistic distribution of <i>like</i> -lexemes .....	131
5.1	Frequency of distribution of <i>go</i> (in %) with degrees of hypotheticality by variety .....	174
5.2	Frequency of distribution of <i>like</i> (in %) with degrees of hypotheticality by variety .....	174
5.3	Frequency of occurrence of V <sub>quot</sub> with level of hypotheticality per variety ...	175
5.4	Co-occurrence of quotatives with mental/speech expressions (in %) .....	176

5.5	Co-occurrence of quotatives with Mimesis .....	178
5.6	Co-occurrence of quotatives with Exclamation. ....	180
5.7	Collocations of <i>like</i> with quotatives across varieties .....	189
5.8	Verb-person Correlation for <i>like</i> in both varieties (in %) .....	190
6.1	Frequencies of occurrence of variants of the (say) variable .....	206
6.2	Production of quotatives per gender .....	207
6.3	Production of quotatives per class .....	207
6.4	Production of quotatives per age .....	207
6.5	Vertical distribution of quotatives per gender .....	209
6.6	Horizontal distribution of all quotatives per gender (in % and N) .....	210
6.7	Logistic regression for <i>like</i> and <i>go</i> with the social categories age, gender, and class in US English (significant results in bold) .....	213
6.8	Logistic regression for <i>like</i> and <i>go</i> with the social categories age, gender, and class in British English (significant results in bold) .....	214
6.9	Distribution of quotatives in US English per gender, (significant differences bold) .....	216
6.10	Distribution of quotatives in US English per educational level, (significant differences in bold) .....	218
6.11	Distribution of <i>like</i> , <i>go</i> , <i>say</i> and unframed across gender and class .....	220
6.12	Distribution canonical versus non-canonical quotatives spilt up by gender and educational level .....	221
6.13	Distribution of quotatives in US English per age-group, (significant differences in bold) .....	223

6.14 Comparison of quotative systems: Like-frequency in the NYU corpus (Singler 2001) and the SWB .....	224
6.15 Distribution of quotatives per gender .....	254
6.16 Distribution of <i>like</i> per gender and variety (in % frequency) .....	258
6.17 Distribution of <i>go</i> with respect to gender per variety (in % frequency) .....	259
6.18 Distribution of quotative strategies according to class .....	261
6.19 Distribution of quotatives per age-group .....	266
6.20 Distribution of <i>like</i> with respect to age per variety (in % frequency) .....	272

# Contents

Declaration	i
Abstract	ii
Acknowledgements	iv
List of Figures	vii
List of Tables	x
Chapter 1     Introduction	1
Chapter 2     Data and Definitions	9
1. Attestations of <i>like</i> and <i>go</i> . . . . .	9
2. The Data . . . . .	12
2.1 American English . . . . .	14
2.2. British English . . . . .	17
3. What is a quote? . . . . .	20
3.1. Indirect versus direct quotes . . . . .	32
3.2. The problem of <i>think</i> . . . . .	36
3.3. Unframed quotes . . . . .	37
Chapter 3     Intralinguistic Factors	41
1. Epistemicity and Quotation . . . . .	43
2. Categories of Realisation . . . . .	46
2.1 Realis . . . . .	47
2.2. Hypothetical . . . . .	48
2.3. Situational . . . . .	50
3. Contextualisation and Hypotheticality . . . . .	52

4. Comparison between the Quotatives . . . . .	60
5. Priming . . . . .	63
5.1. Contextual Semantic Priming Effects . . . . .	63
5.2. Birds of a Feather Effect . . . . .	68
6. The Co-occurrence with Mimetic Performances . . . . .	72
7. Conclusion . . . . .	77
 Chapter 4      The Grammaticalisation of <i>like</i> and <i>go</i>	 80
1. Conceptualising grammaticalisation . . . . .	84
2. The role of ambiguity . . . . .	86
3. The <i>like</i> -Scenario . . . . .	89
3.1. The radial structure . . . . .	94
3.2 The hedging function . . . . .	97
3.3 The focus function . . . . .	101
3.4 <i>Like</i> as a filler . . . . .	104
3.5 The quotative function . . . . .	107
3.6. Discussion . . . . .	130
3.7. Conclusion . . . . .	134
4. The <i>go</i> -Scenario . . . . .	136
4.1. The radial structure . . . . .	137
4.2. The temporal function . . . . .	140
4.3. The modal function . . . . .	145
4.4. The quotative function . . . . .	150
4.4.1. <i>Go</i> as an introducer of voice effects, sounds and gestures . .	151
4.4.2. <i>Go</i> as an introducer of reported speech . . . . .	157

4.5. The substructures of the grammaticalization schema . . . . .	163
4.6. Conclusion . . . . .	164
5. General Conclusion . . . . .	164
Chapter 5      The new quotatives in US English and British English	168
1. Emergent Grammar . . . . .	171
2. Separate but equal . . . . .	173
2.1. The Co-occurrence with Levels of Hypotheticality . . . . .	173
2.2. Priming . . . . .	176
2.3. Verbs of Quotation, Mimetic Effect and Exclamation . . . . .	177
3. The Crucial Differences . . . . .	182
3.1. Telicity with <i>go</i> . . . . .	182
3.2. Indirect quotes with <i>go</i> . . . . .	186
3.3. The collocation of <i>like</i> with other verbs of quotation . . . . .	188
3.4. Verb-person correlation for <i>like</i> . . . . .	190
4. Conclusions . . . . .	191
4.1. Descriptive . . . . .	191
4.2. Methodological . . . . .	193
Chapter 6      Extralinguistic Factors	195
1. Literature overview . . . . .	198
2. A note on statistics . . . . .	205
3. A regression analysis . . . . .	212
4. US English . . . . .	215
4.1. Gender . . . . .	215
4.2. Socio-economic Class . . . . .	218



4.3. Age .....	222
5. British English .....	249
5.1. Gender .....	254
5.2. Socio-economic Class .....	261
5.3. Age .....	265
6. Conclusion .....	281
 Chapter 7: Conclusion and Outlook	 284
 Appendix 1: Sociolinguistic Information	 296
Appendix 2: Transcription Conventions	298
 References	 299

[The introduction and expansion of *like* is] one of the most striking and dramatic linguistic changes of the past three decades, offering sociolinguists an opportunity to study language change in progress on a large scale in order to address the general questions on the mechanism, the causation, and the consequences of change. (Labov 2000)<sup>1</sup>

# Chapter 1

## Introduction

The ‘principle of linguistic economy’ was originally devised by Martinet (1955) in order to account for the position and extension of phonemes, which are dictated or limited by the configuration of neighbouring phonemes. According to this principle, we expect the introduction of an item into a stable system to result in the restructuring of the organism as a whole. We know of such phenomena in the field of phonology (Lindblom 1998). The semantic equivalent to this process is when a new concept gets introduced into a semantic field and causes the members of this field to regroup and consequently readjust their territories. This has been shown with respect to colour systems (cf. Mackeigan, in progress, MacLaury 1998). This thesis investigates the introduction of a new member into a functional morpho-syntactic category. If it occurs, the entire system has to be re-organized in order to accommodate the incoming option. This will consequently lead to each member acquiring a new functional niche. We are witnessing such a large-scale restructuring right now in the pool of quotative introductory devices. Sparked by the advent of the newcomer *like*, as in example (1.1), the hitherto stable functional distribution within the pool of *verba dicendi*, viz. verbs of quotation, has become imbalanced and been forced to reorganize.

### (1.1) The computer buff

- L: she’s talking about ROM and RAM and [you know  
A: ]((snore snore))  
L: and she ha ha ha I know she’s passionate about it.  
ha **I’m like** ha “I can’t relate.  
leave me alone”.  
A: I can’t relate at all,

---

<sup>1</sup> Excerpt from an e-mail message sent by Labov to linguists attending the October 2000 NWAVE (New Ways in Analyzing Variation (in English)) in Michigan. He also outlined a large scale cross-linguistic study of the quotative verbs which he titled “The Tsunami Project”.

There is only one other case we know of in the history of English where an item has been introduced into the quotative cohort. The intrusion of quotative *tellan* (attested since 888, OED) into the hitherto stable quotative system is comparable to *like* being introduced into the pool of quotatives today (cf. Marckwardt (1967:122). Ferrara and Bell (1995:286) write that “the ongoing grammaticalization of a third alternative is altering the balance between traditional forms of dialogue introduction (*viz.* *say* and *go*) ... The obsolence of quotatives (e.g. *quoth*)<sup>2</sup> has occurred before in English when a three-way competition emerged and may be repeated if *be+like* competes successfully with *go*”. The fact that a collective reallocation is occurring at present gives us the exciting opportunity to investigate a phenomenon of language change that is happening concurrently with the research.

By modern times, *pre-like*, the quotative system seems to have balanced out with *say* as the major introducer for reported speech, *think* for reported thought and *go* for the introduction of non-lexical sounds, voice effects and gestures (see Ferrara and Bell 1995, Johnson 2001, Tagliamonte and Hudson 1999). It is of great interest to see how the advent of a new member, *like*, imbalances this previously stable system and how speakers resolve this new choice they are presented with. Due to the rarity of this process, the arrival of a new option into a formerly stable system such as the quotative pool, and the subsequent competition amongst the variants, is an interesting field of investigation. Sociolinguists around the globe have been investigating innovative *like* with an eye on the question of whether there is any indication that a competitor variant will be pushed out, which would parallel the introduction of *tellan* and the consequent obsolence of *cweðan*. It has been claimed that one of the possible outcomes of *like*-intrusion is that the newcomer will drive out one or more of the older members of the quotative cohort, *go* (Bakht-Rofheart 2002), or *say* (Johnson 2001), or unframed quotes. An alternative possibility is that the system will stabilise in its richer form.

---

<sup>2</sup> *Cweðan* > Scots and obs. *quoth* was according to Mackwardt (1967:118-20) the most frequent quotative in “A Grouped Frequency Word-List of Anglo-Saxon Poetry” (430 times versus 380 times *secgan*).

Ever since its advent, there have been a number of studies which investigate the distribution and social spread of the new quotative *like* (most notably Blyth et al 1991, Ferrara and Bell 1995, Fleischman and Yaguello to appear, Underhill 1988). And since the late 1990s, most N.W.A.V.E. conferences have featured at least one paper which has topicalized it in some way (Bakht-Rofheart 2002, Blyth et al 1988, Buchstaller 2001b, 2002a, 2003, Cukor Avila 2003, Dougherty and Strassel 1998, Ferrara and Bell 1990, Fuller 2000, Igoe, Lamb, Gilman, and Kim 1999, Sanchez and Charity 1999, Singler 2001, Singler and Woods 2002). But no paper ever specifically addressed the patterning of the other vernacular quotatives, *go* or unframed quotes (for a definition of this variant cf. Chapter 2). Also, as few studies actually tried to locate *like* within the whole quotative system, the relationship between all members of the quotative pool on the social and the functional planes is still under-researched. A notable exception is Tagliamonte and Hudson (1999), who write that “its [*like*’s] relationship to other members of the English quotative cohort in each context, presents an interesting test site for the examination of linguistic change in progress.” (149, cf. also Tagliamonte and D’Arcy 2003). Indeed, the quotative system provides us with a fascinating opportunity to investigate the interaction amongst numerous competing variants within one variable. In order to understand *like*’s patterning, we have to examine the balance between all members of the entire quotative system. It will be especially revealing to compare *like*’s functional and social behaviour with that of other non-standard members. Hence, this study examines the systematic patterning of the quotative system as a whole after the advent of quotative *like*. I will investigate its interaction with all the quotative strategies which speakers have at their disposition when they want to report. This thesis analyses how the quotative system reacts to such a new form coming in.

The corpora used in this study present us with the opportunity to look at the composition of the quotative system at two different stages of this change, viz. in different diachronic “slices” (Chapter 2 outlines the corpora used in this study). The quotative system of two urban varieties of British English in 1994 and 1995 gives evidence of the patterning of *like* when it first came into being as a quotative. Results from this corpus can help us answer important questions such as the following: Which environments does *like* first get used in? Who uses *like* and when? Is *like* born in a niche which it can enlarge or does it originate as an across-the-board quotative?

A second corpus of US English data, where *like* had already had a lifespan of roughly 10 years and where it was already considerably widespread at the point in time when the recordings took place (1988-90) shows how the quotatives system has settled in its larger inventory. The investigation of the US quotative system reveals which functional and social niches *like* has appropriated for itself. It will become evident that the quotative system has become fully restructured in order to accommodate the additional item. This corpus will show which quotatives have to “give way” for *like* to be able to establish itself. I will argue that, in fact, there is no indication of a “pushing out” of older variants. Rather, the distributive evidence points to a reallocation or refunctionalisation of all quotatives as the new one is introduced (Singler 2001, Buchstaller 2001a). That is, different quotatives occupy different functional niches.

A comparison between the two corpora will also reveal whether the two varieties have reacted in the same way to the additional quotative option or whether we find locally idiosyncratic developments on both sides of the Atlantic. In this respect, this study constitutes an important contribution to the growing body of research on globalisation phenomena and supra-local trends. (Buchstaller 2003, Kerswill 2003, Meyerhoff and Niedzielski 2003, Tagliamonte and Hudson 1999, Trudgill 1986, 1994, and many others).

In the following I will give an outline of the thesis:

Chapter 2 presents an overview of the corpora which form the basis of this study. The methodology for the collection, transcription and coding of the data will be laid out. I will define the variable under investigation and determine the envelope of variation. The variants’ attestations in earlier corpora will be examined and borderline cases will be discussed.

Chapter 3 examines the interaction between the quotative options with respect to linguistic factors in US English. I will evaluate the quotative system under pragmatic, semantic, discursal and syntactic aspects. By examining how the functional load is distributed amongst the members of the quotative cohort, we get an insight into the intimate interplay between the pool of variants. I will show their functional distribution with respect to several linguistic variables such as (i) the epistemic relation between the quotative variant and its complement, (ii) the

“delivery aspects” (Clark and Gerrig 1990), viz. enquoted non- or para-linguistic features (iii) the co-occurrence and interaction with inner state markers or “response cries” (Goffman 1981) and (iii) the interaction with the syntactic system and possible “birds of a feather” priming effects (Scherre and Naro 1991, 1992).

Using two case studies, Chapter 4 discusses *go* and *like*’s synchronic status quo. I will argue that we can understand and explain *like* and *go*’s ambiguity and polyfunctionality (comparative preposition, discourse marker, and quotative for *like* and movement verb, future marker, and quotative for *go*) if we conceptualise functions in terms of family resemblances (Wittgenstein 1978), or prototypes (Berlin 1992 and Rosch 1975, 1978). The manipulation of such concepts is creative with one meaning giving rise to one or more others. I will give a model for the grammaticalisation of *like* and *go* which is based on the radial structure model originally proposed by Lakoff (1987). It depicts their synchronically occurring functions as structured in a field around one focal member. Later functions, such as the quotative use, are conceptualized as functional extensions from this common core and are interrelated via cognitive links. *Like* and *go*’s manifold synchronic functions can hence be conceived of as a network of relations. It will be argued that *like* and *go*’s multifunctionality and ambiguity are highly motivated on a semantic–pragmatic level.

The question as to whether *like* in British English and multiple other varieties (Canada, India, Singapore, Australia, Namibia, ...) is an independent linguistic development or whether the trend was picked up from the *like*-use in the US is not easily resolved. The occurrence of quotative *like* in British English might have been an independent process (especially given the naturalness of the underlying process, as I will argue in Chapter 3), or a development influenced by US English, where the grammaticalization of *like* in its quotative function was first attested (cf. Fleischman and Yaguello to appear). In Chapter 5, I will argue that before we make claims about the borrowing of one variant from one into another variety, we first have to ascertain its fundamental sameness in different locales. Indeed, the comparison of pragmatic and syntactic factors will reveal some interesting but subtle differences between *go*’s and *like*’s development in different locales. This raises fundamental linguistic

questions, such as whether we are indeed comparing the same variable in the US and in Britain, and how we can ratify calling something one and the same variant? Surely a mere surface resemblance is not enough.

The study of grammaticalisation, which attends to the single case of incipient language change and investigates the spread in frequency and linguistic versatility, is a useful tool for answering such questions. In this thesis, I propose a combinatory approach whereby the investigation of (socio)linguistic variables proceeds in two steps. First, I quantify the intra-linguistic behavior of *like* and *go* with respect to various linguistic variables. Statistical methods yield correlations and lead to a frequency-based definition of the variants' properties in both varieties. If there is no statistically significant difference in their patterning with respect to the distributional and functional parameters across varieties, I can ratify my variables as the same across varieties, viz. entities the patterning of which is linguistically non-significantly different. Only in a second step can I proceed to correlate *go* and *like* with extralinguistic, social variables.

But before proceeding to the social factors involved, Chapter 5 will ask whether locality might actually have an effect on grammaticalisation. As we are looking at incipient language change, numbers are small. Therefore, quantitative methods yield unsatisfactory, statistically non-significant, results. In order to account for my data as fully as possible and to trace language change in situ nascendi, the second part of Chapter 5 pays attention to the single case, looking at what is different in the two varieties. A closer look provides evidence that British English indeed constitutes an independent development. This fact fuels the discussion about the importance of social dimensions in grammaticalisation (Lass 2000, Nolde 2001).

After having ratified *like* and *go*'s fundamental sameness in the two varieties under investigation in Chapter 5, Chapter 6 gives an overview of the patterning of the members of the quotative complex with the social variables age, gender, and socio-economic standing. Following Tagliamonte and Hudson's (1999:149) statement that "the diffusion of *be like* beyond the United States presents a possible test case for the examination of putative mega-trends currently underway as English increasingly becomes a global language", the sociolinguistic distribution of the quotative pool in two varieties of English (US and Southern British English) is investigated.

The variant *like* is introduced into the linguistic system of both varieties by typical proponents of language change, young speakers. These linguistic innovators are also the very social group we can assume to be most likely to pick up a new lexical variant through TV programmes such as soaps, talkshows, etc. While there is unanimous agreement that the strong quantitative presence of TV in young peoples' lives is bound to have some influence, the exact shape of this supposed effect is more disputed. In Chapter 6, I will give a possible transmission scenario for *like* from the US to the UK. On the basis of my data, I will examine which social groups carry the newcomer variant in the US and which ones introduce it into the British English quotative system. I will further indicate possible social trajectories that *like* could follow in its spread with an eye on the other non-canonical variables *go* and unframed quotes and also in contrast to canonical *say*.

There is evidence that discourse patterns such as “narrative style” in storytelling (Johnstone 1990) and other norms of narration (Aukrust and Snow 1998) are culture-specific. We would consequently expect discourse features such as quotative introductory items to occur in different frequencies overall and proportionally in different varieties. And indeed, this has been shown by Tagliamonte and Hudson (1999). I will argue that while the social constraints on the (say) variable are generally not very important, *like* and *go* have a very different social reality in the US and in Britain. In addition, my investigation of language attitudes (Buchstaller 2003) have shown that British speakers have not borrowed social attitudes along with the surface item *like*. Rather, just as reallocation of linguistic forms is well attested, my research shows that reallocation of social attitudes also constrains the outcomes of language contact.

In conclusion, this thesis examines the patterning of the quotative pool in its syntactic, pragmatic, discourse and social aspects. The investigation uses the quantitative methodology of variationist sociolinguistics and pairs it with the more qualitative approaches of grammaticalization research and discourse analysis. The fact that no single method can fully explain the patterning of the data speaks for an eclectic, synthetic approach.

Furthermore, variationist sociolinguistics has been looking at how the variability of a given variable can be explained by correlating it with social and



linguistic variables. Of the large body of work that has productively investigated the importance of linguistic factors on linguistic variation, many have found that the linguistic constraints are more important than the social constraints (cf. Baugh 1979, Nagy and Blondeau 2000, Meyerhoff 1999, 2000, Rickford and McNair-Knox 1994). This study gives evidence that while the social constraints are only moderately exciting, the linguistic constraints on the quotative pool are quite substantial. This leads me to conclude that only when we look at variability at all levels, quantitative, and qualitative, extralinguistic and intralinguistic, can we fully describe the behaviour of our variable. Only then can we fully understand the role of the vernacular variants *go* and *like* within the whole mechanism of quotative options. *Go* and *like* have for a long time been dismissed as functionally vacuous variants in the quotative pool. The following pages will show that contrary to claims in the sociolinguistic literature, they are not just intrusive items in the paradigm of reporting devices but fulfil a number of important roles, which are governed by linguistic constraints.

Furthermore, the study of the quotative system provides a fascinating site for research into the rule-governed variability of innovative features and their rival variants. A sharply delimited and hitherto stable set of variants presents a unique opportunity to investigate the restructuring of all variants as new ones come in. This thesis reveals important insights into phenomena such as reallocation, competition within the same socio-pragmatic field, as well as into the interaction of variants within the same variable.

In this light, the restructuring of all quotative variants during the rapid intrusion of the robust new variant *like* gives us an important test-case for various explanatory parameters for language change that have been raised in the literature i.e. Romaine's claims about language and gender (2003), the reallocation hypothesis as raised by Britain (2002) and Britain and Trudgill (1999), the reformulation of the standard vs. non-standard dichotomy into local vs. supra-local variants by Foulkes and Docherty (1999) and the Milroys' findings (1998) concerning networks and class.

No linguistic investigation can proceed without reference to concrete data of some sort, no matter how informal or intuitive the collection method used. Equally, an analyst necessarily and unavoidably brings to bear certain theoretical assumptions to an analysis of data. Work within the sociolinguistic tradition has always emphasized the principle of **accountability**, i.e. the explicit reporting of how data have been assembled, analyzed, presented and interpreted. (Foulkes and Docherty 1999: 23)

## Chapter 2

### Data and Definitions

#### 2.1. Attestations of *like* and *go*

Macaulay's (2001:5) claim that "the quotatives *go* and *be like* are relatively recent innovations, apparently starting among younger speakers, possibly in California" exemplifies many researchers' view that they are rather new additions to the quotative pool (cf. also Butters 1980, Schourup 1982b). Quotative *go* has been treated by sociolinguists as if it were another newcomer quotative, a slightly older but equally innovative non-canonical co-variant of *like*. But these studies fail to consider historical evidence. Looking up *go* in the OED yields a number of borderline cases for quotative use, where it is ambiguous between a quotative introducer and its older use as a story introducer (*that's the way the story goes*). Absolutely certain cases of quotative-introductory function, first with onomatopoeic sound effects only, are attested from 1791 onwards (Cowper Reddy "*and his noble heart goes pit-a-pat*"). Thus, *go* has to be considered a member of the quotative pool that has been around a while, if largely unnoticed.

But we note that in the early beginnings, *go* is only used in the sense "with imitative interjections or verb-stems used adverbially, e.g. *to go bang, clatter, cluck, crack, crash, patter, smash, snap, tang, whirr*, etc" (no 10 in the OED). The first account for *go* as a full quotative, viz. an item which can introduce linguistic material as well as sounds, is much later. The sociolinguistic literature usually refers to Butters (1980) as the first mention of *go* as a quotative that can take linguistic

complements in US English. While Butters was certainly the first to point out its existence as something noteworthy, the first study I know of to mention it in this function is Partee (1973). Schiffrin, in her 1981 study, reports 10% *go* within the pool of quotative verbs. I will henceforth refer to *go* in this larger functional coverage, viz. when it can take the whole range of quotative complements, voice, sounds, gesture as well as linguistic elements, as a ‘full’ quotative. Consequently, when reference is made to *like* and *go* as ‘new quotatives’, the attribute ‘new’ pertains to slightly different degrees for the two quotative options. Whereas *like* has only recently enlarged its functions to the framing of quotations per se, for *go*, the ‘newness’ holds with respect to the recent broadening of complement type, from para- and non-linguistic to linguistic.

In British English, the general view seems to be that *go* has been around for a while (Macaulay 2001) but there is no explicit account for its arising as a fully fledged quotative verb. The first study I could locate that mentions *go* in the UK as a full quotative is Cheshire (1982) with data collected in Reading in the late 1970s. Andersen (1996) finds it widely used in the Corpus of London Teenage English (COLT)<sup>1</sup> collected in 1993. Tagliamonte and Hudson (1999) are the first to analyse the British quotative system as a whole with a corpus collected in 1995-6 containing *go*. For Scottish English, Macaulay (2001) reports its existence in 1997 data. While my corpus shows that quotative *go* was alive as a fully fledged quotative in BrE in 1994, it is an outstanding question how far back we can trace its full quotative function.

It was Butters (1982:149) who first noted *like*’s new quotative function in the USA. “Many speakers who use narrative *go* also have narrative use of *to be* (usually followed by *like*) where what is quoted is an unuttered thought, as in *And he was like ‘Let me say something’ or I thought I was going to drown and I was (like) ‘Let me live, Lord’*”. In a 1986 article, Tannen finds 4% *be like* in her corpus but does not give any indication of its date of collection. In the years to follow, a flood of studies on quotative *like* followed, most of them on its sociolinguistic behaviour, and notably

---

<sup>1</sup> “The Bergen Corpus of London Teenage Language (COLT) is the first large English Corpus focusing on the speech of teenagers. It was collected in 1993 and consists of the spoken language of 13 to 17-year-old teenagers from different boroughs of London. The ... corpus is a constituent of the British National Corpus.” (University of Bergen, <http://www.hit.uib.no/colt>).

two on its grammaticalisation (Romaine and Lange 1991, Meehan 1991). Within the UK, Miller and Weinert (1995) report for Scottish English no occurrence of *like* as a quotative verb prior to 1980. Also, my investigations have revealed that the British National Corpus, collected up until 1993, does not yield any tokens of quotative *like* (this excludes the COLT). In the same vein, Tagliamonte and Hudson (1999) claim that quotative *like* is unattested in Britain until the early 1990s. We know that by 1993, *like* has found its way into the use of London teenagers because Andersen (1996) is the first to report its occurrence in the COLT. Tagliamonte and Hudson (1999) attest *like* for their 1995-6 data in York. In turn, Macaulay (2001) shows that Glasgow Scots speakers use *like* in 1997.

In all evidence, and contrary to *go*, *like* really seems to be a newcomer to the quotative pool. The following evidence corroborates this claim: (i) *like* was not mentioned anywhere in the scientific literature before Butters (1982), (ii) it only recently appeared in dictionaries. For example, the 2nd edition of the Random House Webster (1999:768) is the first to incorporate the new use of *like*:

informal (used esp. after forms of 'to be' to introduce reported speech or thought) (3) She's like "I don't believe it," and I'm like "No, it's true".

Furthermore, (iii) the OED does not yet mention *like* in its full quotative use with dummy *be* and only gives an entry where *like* serves as a complementiser with a full quotative verb, a stage which most probably led to *like*'s grammaticalisation of a full quotative. With respect to *like*, a full quotative verb means a verb that can carry the force of introducing quotation independently of any graphic dialogue introducer such as *say*, *think* etc (cf. Chapter 4 for *like*'s grammaticalisation). I will give the OED entry for *like* in the context of quotation in full:

*N. Amer. colloq.* Followed by an adj.: in the manner of one who is —. Cf. *like crazy* (CRAZY a. 4c), *like mad* (MAD a. 1c). **Also in less analysable constructions.** [...] 1970 *Time* 31 Aug. 19 Afterward, a girl came up to me and said, 'You kinda look interested in this; did you know there are civil rights for women?' And I ***thought like*** wow, this is for me. (OED online, **emphasis mine**).

Given this information, my British English data represents the very point in time when *like* was first introduced into the British quotative system. Examining the

1994/5 corpora hence yields interesting insights into the patterning of a new quotative item at its very emergence. The data from the US was collected at a point in time when *like* was already sufficiently widespread.

I will now discuss the composition of the two corpora used in this study. I will lay out their speaker selection, social composition and the transcription and coding methodologies used.

## 2.2. The Data

The data for this thesis is drawn from 4 corpora of two varieties. These data had already been collected for purposes other than this present thesis. The reasons for basing my investigation on already existing corpora are the following: (i) One of the aims of this study was to explore the earliest possible recording of British English where *like* had just become introduced into the quotative system. As this time predates the start of my PhD, I could not possibly have collected the data myself. I had to rely on recordings that had already been made during the time slot I was interested in. (ii) Furthermore, I wanted to do a cross-variety comparison of a low frequency discourse feature, which required a huge amount of data for each variety. Given the restricted time and resources of a PhD, I had to weigh the cost of collecting my own data (with the time needed for speaker selection, data collection, transcription etc.) against the benefits of working with pre-existing corpora (and hence being able to devote more time to the analysis and implications of the data). I opted for the latter.

The work with data coming from different sources inevitably leads to some slight inconsistencies across the corpora which I will discuss under the respective data sources. However, I suggest that the disparities are minor enough to make the data comparable. This is for the following reasons: the corpora of both varieties were collected within sociolinguistic research programmes, which documented and controlled the social factors involved. The sociolinguistic matrix was filled in both cases with equal amounts of male and female speakers coming from 2 distinct social strata (for the definition of social “class” see the following sections) and from a wide

range of age groups which I grouped into young (<38) and old speakers (>38)<sup>2</sup>. Appendix 1 gives a break-down of the social profile of the speakers used in this investigation. Furthermore, all data consists of unmonitored, spontaneous talk-in-interaction from speakers paired in dyads. The speakers were native speakers of the respective varieties and there was no or next-to-no intrusion from the fieldworker.

Note that this study uses quantitative variationist methodology in order to investigate sociolinguistic patterns, viz. the way membership in social categories correlates with the presence or absence of linguistic features. While this perspective has been the prevailing one in variationist sociolinguistics to date, where social factors are considered constraints on linguistic variability, it is in order to mention important criticisms that have been raised against the operationalisation of prefabricated, given, social categories. One shortcoming of such an approach is that the notions of gender and sex are not disengaged. For a problematisation of the notion of gender with respect to features of homosexual speech see Labov (2001:263), Gaudio (1994), Leap (1996), Schilling-Estes (2002). For the delineation of the concepts sex and gender see Giddens (1989). Research in the social constructivist framework has questioned the view that gender and race are de facto categories (Holmes and Meyerhoff 2003, Eckert and McConnell-Ginet 2003). Rather, they suggest an approach whereby “the relation between person and society is [considered as] dynamic and mediated by language“ (McElhinny 2003:36). In such a performative approach, gender and race are considered accomplishments. They are emergent, viz. the products of linguistic and other interaction. Acknowledging such justified criticism, this thesis takes a quantitative, variationist, approach in order to investigate the extralinguistic factors that play into the patterning of the quotative system during the introduction of a newcomers variant. Such an exploration into the question to which extent certain speakers adopt and further non-standard variants can provide a useful backdrop for further, more detailed, work on the gendered practices that have a bearing on the re-enactment of self and other past performance (Holmes and Meyerhoff 2003). Furthermore, as Eckert (1989, 2003) points out, chronological

---

<sup>2</sup> In the Derby data, the older speakers were >45. But as there are no speakers in the age range from 38-45 in the Derby corpus, suffice it to say that overall, young speakers were below 38 and old ones above.

age has often been divided into equal chunks without consideration of the importance of these chunks with respect to socially significant life stages such as adolescence. This study takes the commonly cited “temporal isogloss” of 38 years for *like* (Singler 2001) as the cut-off point for age-classification and later splits up the two rough groups into chronological age.

In the following paragraphs, I will discuss the composition and the coding mechanisms used in the corpora in more detail.

### **2.2.1 American English**

The analysis of US English is based on the Switchboard Telephone Speech Corpus (SWB) available online through the University of Pennsylvania Linguistic Data Consortium ([www ldc.upenn.edu](http://www ldc.upenn.edu)). The data were collected by Texas Instruments in 1988-92 and first released by the LDC in 1992-3. They comprise of roughly 2400 two-sided telephone conversations, recorded from 543 speakers (302 male, 241 female) from all areas of the United States. The telephone calls were handled automatically by a computer-driven system. The caller was given recorded prompts while the "robot operator" facilitated the connection with the callee and gave a pre-selected topic for discussion. Interlocutors were told to talk about one of 70 topics, of which about 50 were used more than once. The topic and speaker selection was regulated in order to ascertain that (i) no two speakers would converse together more than once and (ii) no one spoke more than once on a given topic. Speakers were told to talk for about 5 minutes. The ensuing conversations lasted for varying lengths of time and were cut after a certain time if speakers exceeded the pre-given time limit. The speech from the two subjects was recorded into separate channels and transcribed and tagged with respect to grammatical categories. The speakers were asked where they grew up during the first 10 years of life. They come from 7 major linguistic regions in the US, South, North, New England, New York, West, South Midland and North Midland (Lance 1999: 311). There were also some speakers of mixed origins.<sup>3</sup>

---

<sup>3</sup> This information is taken from the Switchboard information services provided via the LDC webpage (<http://www ldc.upenn.edu/doc/switchboard/manual.html>).

For the analysis of *like*'s and *go*'s intralinguistic factors and for the case studies of grammaticalisation (Chapters 3 and 4), I used the whole corpus. For the extralinguistic investigation, I limited the data to a sociolinguistically balanced subset of 136 speakers. I made sure that I included equivalent numbers of speakers from each social category for each dialectal region (for a break-down consider Appendix 1).

Because I was aiming at depth in regional coverage while keeping token numbers per cell within reasonable limits, the only other social variables controlled for were age, gender and educational level. The SWB provided the birth year of the speakers. In order to achieve general comparability with the age range chosen in the Derby and Newcastle corpora, I subdivided the US English speakers into two broad age groups +/- 45 (with no speaker between ages 38 and 45). To counterbalance this very broad categorization, I will later align each speaker on an age-scale and correlate quotative production with numerical years of age.

As Labov (2001) noted, the division of a population into socioeconomic classes tends to trigger a feeling of unease in many sociolinguists in the western world. He suggests that rather than trying to avoid the issue because of ideological reservations, sociolinguistic enquiry should aim at an accurate description of all social factors that play into variability within the linguistic system. When looking at all the potential social constraints, socio-economic stratification certainly is a possible candidate. Whether or not it plays a significant role in the patterning of the quotative system can only be determined post hoc, by using it as a predictor variable in a statistical analysis. Research in anthropology has shown that social categorizations need to be justified from within the system of the society under investigation (Pike 1954/1967, Harris 1964, 1968). Only a socio-economic subdivision which makes sense in the emic structure of the speech community will yield authentic criteria and provide useful results. Hence, the criteria for the subcategorisation of speakers within the socio-economic hierarchy in Great Britain and the US have to be tailored to the respective situations. They are necessarily different if we want them to be meaningful when applied to their particular societies. The United States of America is and has been a highly mobile society, regionally as well as within the social



hierarchy. A nation-wide large-scale study which bases the socio-economic ranking on residential area does not make much sense.<sup>4</sup> This is in contrast to studies such as the ones conducted by Lippi-Green (1989) and the Milroys (L. Milroy 1980, Milroy and Milroy 1985), which are based on the longevity of settlement patterns and which base social stratification on residence in a specific, sharply delineated locality only. But given the short-lived nature of settlement patterns in the US and the large number of speakers from all over the nation in the American corpus, the criterion 'neighbourhood' used for British English as the basis of a class-based categorisation cannot be adopted for the American variety. The much greater fluidity of the American society defies a priori categorization of individual entities by their ephemeral occupation or residence patterns (see also Ganzeboom, Luijkx and Tremain 1989).

As Chambers notes (1995:7), the most important elements in the formation and perception of social class are education, occupation, area and form of residence. It would certainly be desirable to have such multiplex parameters for the determination of the placement of the individual speaker within the socioeconomic hierarchy.<sup>5</sup> Unfortunately, the build-up of the US English corpus on which this thesis is partly based, is such that educational level is the sole determiner of socio-economic standing (cf. Labov 2001:115 et passim for a discussion of educational indicators for his SEC category). I am aware that this constitutes a simplification and would have preferred a social index based on at least a score for occupation, income, and residential neighbourhood (cf. Labov 1994) as well as educational level. But it has to be pointed out that education, if not only and primarily, is at least causally linked to social class via most of the related concepts: The degree of schooling will determine a person's qualifications. These have, in most cases, a pre-selective effect on the range of professional choices. Occupation, in most cases, has a direct causal effect on remuneration and financial standing. Income then determines one's residency with urban patches of much-sought after i.e. expensive areas and less-sought after,

---

<sup>4</sup> This is not to say that studies based on specific neighbourhoods cannot be implemented. Given that necessary care is taken to select areas with stable residencies and that the individual area is carefully selected, such a design is a highly valid enterprise. See Labov 1994, 2001 for the study "Linguistic Change and Variation in Philadelphia", especially Labov (1994: 35-73) for speaker selection and methodology. Cf. also Edwards (1992).

<sup>5</sup> For complex multiple indices see Duncan (1961), Nakao and Treas (1992). For an employment of social indices in sociolinguistic studies see Cedergren (1973), Labov (2001) and Trudgill (1974).

cheaper areas. Thus, while a reduction of the determinant social class to educational level is certainly a gross underspecification of the complexity of societal reality, it certainly is the one element that subsumes and causally determines most other ones (for an in-depth discussion of the determining factors of social class see Labov 2001). In Macaulay's (1976) survey of Glaswegian speakers, occupation was the sole defining criterion of social class.<sup>6</sup>

Also the number of social groupings that are distinctly different within a society is an emic concept determined by the build-up of the particular culture. With respect to the US English corpus, I chose to classify the speakers very broadly into two socio-economic groups, the less educated and the more educated speakers (with a cut-off point after a finished college education).

### **2.2.2. British English**

The BrE data on which this study is based were extracted from the project *Phonological Variation and Change in Contemporary Spoken British English* (UK ESRC grant no. R00234892; 1994 -1997 to J. Milroy, L. Milroy & G. Docherty). It consists of spontaneous spoken interaction collected from two urban varieties of English, Derby and Newcastle. The Derby Corpus was recorded in 1994. The Newcastle Corpus was collected in 1995 in order to match the Derby corpus. While Derby lies at the transition point between several major dialectal areas and has salient features of northern as well as south midland varieties, Newcastle has a quite strong northern influence (see Docherty and Foulkes 1999). Both corpora were collected within a sociophonological research programme (Milroy, Milroy and Docherty 1997). The multi-million word recordings span 32 45-minute cassette or DAT tapes. Of these recordings, which usually contained a word list at the end, I used the first side of a 45 minutes audio tape in Newcastle and one side of a DAT tape in Derby. I transcribed them orthographically.

---

<sup>6</sup> But Macaulay himself pointed out that Glasgow's peculiar situation might be the factor causing this one-to-one correlation between occupation and social standing. Also, it has to be conceded that the speakers in my corpus are of various age groups and much more mobile than the Glaswegian speakers both in the social as and in geographical sense.

In both corpora, speakers from two socially contrasting areas were paired up. The determination of middle class and working class speakers was via their residential area. This method was chosen because of the more neatly stratified class-system in Britain (mentioned above). The fact that residential areas in the two locations were quite hermetic made a grouping according to socio-economic status determined by neighborhood practical (consider also Chambers' (1995:38-39) discussion of the advantages of selecting people from well-defined neighborhoods). Derby and Newcastle indeed constitute closely-knit, socially clustered residential areas of manual workers' communities and other, also areally delimited clusters of white-collar families (for other neighbourhood studies see especially L. Milroy 1980 in Belfast and Labov 2001 in Philadelphia).

I will now shortly motivate the choice of residential areas in the two corpora: The speakers were selected so as to conform to two broad socio-economic settlement areas. The classification of neighbourhoods as contrasting on the socio-economic dimension rely on the experimenter's native intuitions and are underlined by demographical facts (such as housing, unemployment rate and occupation patterns) given in the General Census (1991). In Newcastle, the two residential areas are the electoral wards of Woosington and Westerhope, which are judged by the inhabitants themselves as broadly working class and broadly middle class. In Derby, the suburbs chosen were Spondon for the middle class (+ 4 speakers from Little Eaton) and Chaddesden for the working class speakers.

As recommended in Feagin (1979:25-26), the experimenters were natives of the towns (in Derby) or used a native as a broker (in Newcastle). They could hence exploit their community-internal knowledge of settlement patterns, social strata and their personal acquaintances. Britain (1999:22) comments on this method as follows "Taking full advantage of my insider credentials as a native to the area, I used community-internal techniques for selecting speakers, drawing on peoples' pre-existing social relationships". Speaker selection was hence informed by in-place social relationships, through the friend-of-a-friend method.

Generally, the fieldworkers gave as the reason for the recording the intention to research the perceptions of older and younger locals regarding how they felt about growing up and living in the city. There was no set topic for conversation and

linguistics was not mentioned at all, as the interviewers “aim was to collect a corpus of reasonably relaxed conversation from the speakers involved” (Oxley 1994). With the exceptions of a room in the Conservative Club and a classroom in a school, all interviews took place in the speakers’ homes with the speakers sitting comfortably face-to-face. There was a visible microphone on the table. The interviewers made it clear that they did not intend to participate in the conversations and usually stayed in the same room or an adjacent room during the tape-recording. All subjects embarked on lively talk-in-interaction with the only exception of tape Derby16, where the subjects were clearly microphone-shy. Only once, when the conversation “dried up”, did the interviewer intervene and suggest lines of talk by asking leading questions.

In order to balance the speakers sociolinguistically, they were evenly subdivided by age and sex, which resulted in a matrix of 4 cells per group (see Appendix 1). The Derby and Newcastle corpora came tagged for age in two categories, young (16-24 years) and old (45-65 years) in Newcastle and (14-27) and (38-69) in Derby.

In Newcastle, mostly single-sex pairs were used. In Derby, six of the tapes contain mixed-sex dyads. As one of the experimenters pointed out (Oxley 1994), this set up resulted in one peculiarity depending on age. Many older mixed-sex dyads addressed their speech to the experimenter rather than chatting amongst themselves. As they directed their speech to a different interlocutor, viz. the fieldworker (who often only participated non-verbally), they tended to produce a different form of talk. They produced quite a number of stretches of speech where one speaker tells a story. Due to the fact that only one speaker engages in longish monologues, these tapes contain longer stretches of narratives by one person. This is in contrast with the highly interactive speech we witness with most of the younger subjects, who, almost without exception, talked between themselves and did not pay much attention to the experimenter. But, as both forms of talk are (i) situated within the realm of spontaneous talk-in-interaction and (ii) largely consist of the same genre, viz. narratives, this effect should not affect the comparability of the data with respect to quotative use. Overall, I suggest that the fact that the recorded speech has different levels of collaborative production does not limit comparability.

One other irregularity needs to be addressed at this point. In most older mixed-sex couples, the female speaker produced a far larger amount of speech than the male. The older women tended to take over the role of the principal story-teller, which is especially noticeable with couple Derby10. There were two ways in which the corpus design tried to get around this problem: (i) the experimenter in Derby tried to counterbalance uneven production output by starting to address questions explicitly to the male speakers. Furthermore (ii) many of the old speakers were grouped into same-sex dyads, which eliminates the gender-bias of the opposite-gender grouping. While for the older speakers gender thus played a huge role, amongst the younger speakers, men and women produced roughly the same amount of quotations. This results overall in a slight bias towards female speakers in the number of quotations produced (cf. Table 6.2).

Before I turn to the investigation, I will have to clarify a few theoretical concepts involved in this study. I first need to describe the envelope of variation of my analysis and to define my variable. I have to ask the question ‘What is quotation?’<sup>7</sup> or rather ‘What constitutes quotation for me, in this analysis?’ I will now give the working definitions I have adopted for this investigation and will justify these decisions in the light of the current literature.

### **2.3. What is a quote?**

As Cameron (1998:48) shows, research on reported speech has been carried out in many subdisciplines of linguistics such as anthropology and ethnology (Boeder 2003, Briggs 1992, Hymes 1975), discourse analysis and literary stylistics or pragmatics (Bakhtin 1986, Bolden to appear, Chafe 1994, Clark and Gerrig 1990, DuBois 1989, Faircloth 1992, Ferrara 1992, Goffman 1981, Golato 2003, Johnson 2001, Johnstone 1987, 1990, 1993, Leech and Short 1981, Lucy 1993, Mathis and Yule 1994, Mayes 1990, Maynard 1996, Myers 1999, Page 1973, Polanyi 1982, Sanders and Redecker 1996, Schiffrin 1981, Tannen 1988, 1986, Thompson 1996,

---

<sup>7</sup> I prefer the term ‘quotation’ to ‘reported speech’ because it is not necessarily speech that is reenacted.

Vološinov 1973, Yule 1993, Yule and Mathis 1992), historical linguistics and language change (Buck 1915, Kammerzell and Peust 2003), kinesics (Kendon 1992), language typology (Coulmas 1985, 1996, Li 1986, Massamba 1986, Roeck 1994, Güldemann 2001), phonetics and intonation (Fonagy 1986, Klewitz and Couper-Kuhlen 1999, Kvavik 1986), syntax and semantics (Declerck and Tanaka 1996, Munro 1982, Partee 1973, Steever 2003, Wierzbicka 1974), and variationist sociolinguistics (Johnstone 1987, Vincent and Dubois 1996). Especially the newly-arrived *like* has triggered a great deal of interest in the quotative system (Blyth et al. 1990, Cukor-Avila 2002, 2003, Ferrara and Bell 1995, Fuller 2000, Johnson 2001, Singler 2001, Tagliamonte and Hudson 1999). Recently, more and more researchers have tried a grammaticalisation approach in order to capture the emergence of quotative function in constructions which previously did not have this meaning (Cohen, Simeone-Senelle and Vanhove 2003, Golato 2000, Güldemann 2001, 2003, Klamer 2003, Meyerhoff 2003, Meyerhoff and Niedzielski 1995, 1998, Romaine and Lange 1991, Waksler 2001).

I feel that I need to first lay out how I, in this study, define my subject of analysis. We need to come to terms with the fact that quotation contains multiple voices, the voice of the narrator and the voice of the person quoted. Talbot (1992) refers to a “tissue of voices”, which is closely related to the concept of “multivoicedness” and “polyphony” raised by Bakhtin (1986) and Vološinov (1973). The question is how can we account for this “relationship between an inset and a frame within an utterance” (Maynard 1996:210)?

Looking at a data-base of talk-in-interaction yields the result that quotes defy syntactic rules as they seem to be freely plugged into linguistic structure (cf. Dailey-O’Cain 2000). As Clark and Gerrig (1990:771/2) put it, quotations can occur “embedded in many type of external structure”. In this vein, Cameron (1998:53) gives the examples of a quotation embedded as a noun, as a verb and completely free:<sup>8</sup>

---

<sup>8</sup> The first line gives the original Puerto Rican Spanish, the second line (*in italics*) gives Cameron’s translation.

- embedded as a noun: (2.1) Y se formó un “¡Vente tú!”...  
 ‘and a “come here, you!” got going’
- embedded as a verb: (2.2) cada vez veía a uno tenía que “Vi a fulano, Toco el palo”  
 ‘every time he’d see one he had to “I saw so-and-so. I tag the tree”’
- completely free: (2.3) estaba “yeee” llorando allí  
 ‘and she was “yeee” crying there’

Hence, syntactic parameters do not help in defining a quote.

A useful approach for defining quotation that I would like to discuss here in slightly more detail is Fauconnier’s and Sweetser’s (1996) mental space model. According to their model, we can conceptualise a quote as a person-bound, embedded space within a narrative. It marks the introduction of a subjective point of view and restricts the validity of the presented information to a particular person or entity. Hence, quotation involves the creation of spaces within embedded spaces, which are in turn entailing a restriction of the validity or factuality of the embedded material (cf. also Fauconnier 1985 and Dinsmore 1991). Every space has its own truth-conditions, different degrees of epistemic stance involved in it, etc. For a similar concept of framing consider Goffman (1974), (1981), for the concept of contextual frames see Mandler (1984), Schank and Abelson (1977) and Tannen (1993). Linguistic markers (such as quotatives) or prosodic cues help the speaker to establish and conversely the hearer to interpret the insertion of embedded mental spaces within narrative. In Fauconnier and Sweetser’s terms (1996:295), those markers are space builders for embedded spaces. Let me apply this model to an example:

(2.4) BrE – Derby 11: *The good old days at Parcon Electronics. Told by Kate*<sup>9</sup>

- |    |   |
|----|---|
| K: | Johnny Mardo and Neill Archer,  |
| T: | they’d gone to Parcon.  |
| K: | I think they’re at Parcon,<br>cause Anne said - she <b>says</b> “I wonder if Archie knows”,<br>(...)<br>h ah ah<br>she <b>says</b> “ohhh dear”,<br>she <b>says</b> “they were the good old days weren’t”,<br>she <b>says</b> “we really had a laugh didn’t we”,<br>I <b>says</b> “yeah” I <b>says</b> “that was when none of us had got any problems ha ha ha”, |
| T: | only than the ones you created yourselves or amongst each other,  |

<sup>9</sup> Please consider Appendix 2 for the transcription conventions.

In the above excerpt, the

**BOX**

represents the narrative sphere, as told by Kate at the moment of narration  
represents an embedded sphere of a 3<sup>rd</sup> person sg. female speaker (Anne)

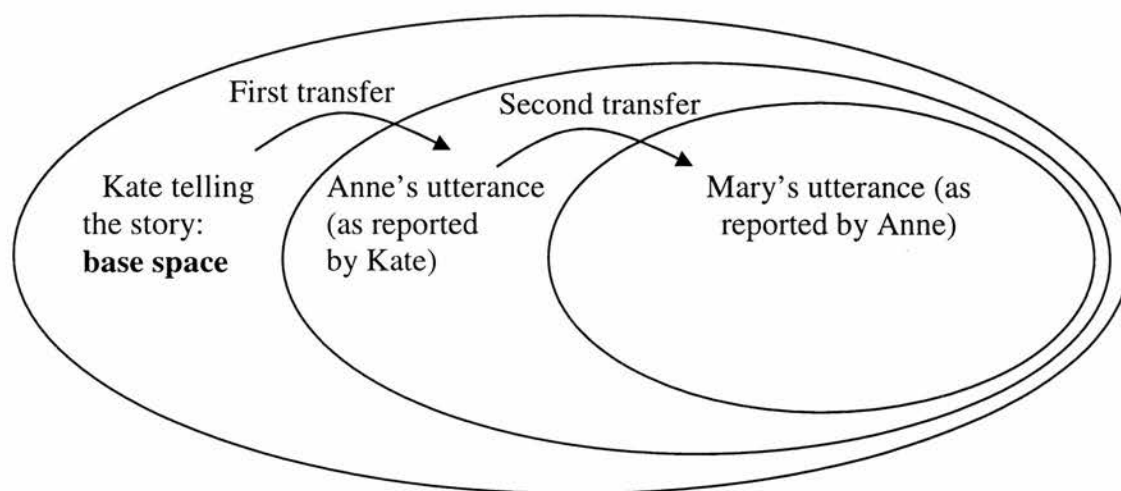
**□**

represents another embedded sphere of a 1<sup>st</sup> person sg. female speaker. This speaker is the same person as narrator (Kate) but in a different consciousness at the time of speaking.

In example (2.4), there are 3 different spheres of experience, the narrative surround, which is the ground zero of the narrative structure (called the “basic mental space” by Fauconnier and Sweetser 1996:295), and two embedded spaces which feature reported speech. Linguistic markers, in this case the quotative *says*, help the speaker to establish and conversely the hearer to interpret the insertion of embedded mental spaces within the narrative. In Fauconnier and Sweetser’s terms (295), those markers are termed “space builders” for the embedded spaces (cf. Jakobson’s (1971) shifters, Gumperz’ (1989, 1992) contextualization cues) which mark a ‘transfer’, viz. a jump from space to space.

The mental space model has the advantage that it helps us to understand cases of multiple stacked quotations. We can envision them as spaces within spaces, usually indicated by space builders. In the example above, if Anne, the reported speaker had in turn quoted the utterance of their friend Mary, we could conceptualize this as shown in figure (2.1) below.

**Figure 2.1: Multiply embedded spaces as of Fauconnier and Sweetser (1996).**



The mental space model makes it easy to parse structures like the above where we find quotes within quotes: We consider them as embedded spaces of a second order.



This cognitive model is of great value for the definition of quotation. (i) It has the advantage that it is independent of syntactic structure. (ii) It is also flexible enough to allow for recursion. (iii) It allows for jumps back and forth between spaces. (iv) Furthermore, as Bal (1990) has pointed out, it can capture the fact that narrating and enquoting (what she calls 'vision') coincide in the person of the narrator, whose voice is always present. Hence, a mental space model is able to conceptualise the fact that quotation consists not only of changes of perspective. It lets us understand the infusion of one perspective into another – for example in cases where a narrator superimposes evaluative devices into a quote in order to characterize the person quoted i.e. by representing the quoted person with a high nagging voice, a French accent etc. (cf. Bakhtin's (1981) dialogic principle, Fludernick's (1993) echoing, Labov's 1972 concept of embedded evaluation, Stanzel's (1984) contamination).

The conceptualization of quotes as mental spaces is reminiscent of Bakhtin's (1981, 1984, 1986) notion of "dialogism" or "polyvocality" (cf. also Bakhtin/Medvedev 1978), Goffman's "lamination" (1981) and Vološinov's (1986) concept of "heterosemy of voices". What all have in common is the concept that when quoting, there are (at least) 2 voices superimposed in narration. Hence, we have to posit (at least) 2 types of subjectivity, the narrator (=reporting speaker) and the reported speaker(s), which pertain to different perspectives. In Genette's (1980) terms there is a difference in the experiencing subject and the narrating agent. The interpretation of the story depends not only on who is quoted but also on whose voice verbalizes it (cf. Goffman's (1981:226ff) production roles of 'animator' and 'author').

In this thesis, I will adopt a definition which relies heavily on the mental space model (and its precursors) and combines it with the 'speakers as actors' metaphor that is found throughout Goffman's work. This conceptualization has given rise to the powerful theatre metaphor used also by Wierzbicka (1974). When quoting, the narrator "“plays his part”, that is to say, imagines himself as the other person and for a moment behaves in accordance with this counter-factual assumption" (Wierzbicka 1974:272). Later researchers have picked it up showing that reported speech is enactment (Haberland 1986), demonstration and "non-serious action" (Clark and

Gerrig 1990). In Wierzbicka's words: "the author of the quotative .... does not say what the content of the quote is (i.e. what was said) – instead he does something that enables the hearer to see for himself what it is, that is to say, in a way, he shows this content" (1974: 282). If we take on board this image, we assume that quotation is not a word for word (verbatim) re-production (pace Coulmas 1985:42, Leech 1974:353, Li 1986:40). Rather, the quoting speaker assumes the role of the original speaker and selectively depicts the crucial aspects of earlier behaviour.

Contingent on this concept of selective re-enactment is the notion of "constructed dialogue" introduced by Tannen (1989), which takes into account the loose equivalence between original behaviour and quotation (called token mimicry by Rimmer 1988). When reproducing another's verbal and bodily behaviour, the quoting speakers are faced with their own personal limitations with respect to voice quality, language, etc. Consider the following example:

(2.5) USE - File 1137: *The joys of teaching EFL*

- S: I tried to get her to say hello,  
and **she'd be like** "(CHOKE)="
- C: [ha ha ha ha
- K: [ha ha ha ha] ha ha ha ha ha ha ha ha ha ha [ha ha ha ha ha ha ha
- S: [no puedo no puedo]".
- I'd be like** "yes you can,  
just[say 'hello].
- K: [ha ha ha ha ha ha ha ha ha ha ha ha
- S: [hello' Amaria]".
- K: [ha ha ha ha ha ha ha ha ha ha ha ha ha ha ha ha
- S: ["(CHOKE)"]
- K: [ha ha ha ha ha

The depictions of S's student's attempts to speak English are clearly an approximation of the original output.

Tannen's (1986:313) and Lehrer's work (1989) as well as much of Goffman's (1980) insightful research, has shown that speakers are not actually able to remember and consequently faithfully quote even short utterances. And Chafe (1980) illustrates that even if speakers could technically reproduce a chunk of speech verbatim - because they were asked to memorize a story- when quoting, they usually refrain from doing so (for the discussion and demotion of the verbatim assumption see also Clark and Gerrig 1990). Hence, quotes are to be considered approximative reproductions of original behaviour. Romaine and Lange (1991:230) speak of

“adequate representations” of the original utterance. What exactly adequate means is defined by the context as well as negotiated by the interlocutors themselves.

Furthermore, Tannen (1986), Vincent and Dubois (1996) and Mayes (1990) have shown that a large proportion of quotations never happened and are purely imagined. Moreover, there are large numbers of quotes which report original thought, viz. a sphere to which the reporting speaker does not have access (except, obviously in cases of self-report, cf. Chapter 3).

In sum, in this thesis, quotation is defined as the performance whereby speakers re-enact previous behaviour (speech/thought/sound/voice effect/gesture) while assuming the dramatic role of the original source of the enacted behaviour (cf. Yule and Mathis 1992:203). This re-enactment, I repeat, is not meant to be verbatim and is in an inherently approximative relationship to the initial event, if the initial event occurred at all.

After having defined the subject of my analysis, quotation, I will now discuss how I have delimited it against some related concepts. I will further justify why I have excluded certain constructions. I will also discuss my treatment of various borderline cases.

In my analysis, I have not counted as separate tokens cases of repetitions of quotes in sequences such as in example (2.6).

(2.6) Invented example - *Repetition after Repair*

01 A:     blah blah blah  
02 B:     what did you say?  
03 A:     blah blah blah

In this example, I have interpreted line three as a verbatim repetition of what was said in line one. This is because no re-enactment with a temporary suspension of belief (Wierzbicka 1974) is taking place, no demonstration, no role-play but only a reiteration of previously stated material because of auditory difficulties or other reasons (such as disbelief). Consequently, cases like the above, which feature a repeat sequence after an other-initiated repair prompt (Sacks, Schlegloff and Jefferson 1974) have been excluded from the analysis.

In the same vein, I do not count as two tokens but as two instances of the same token cases in which a speaker is running into communicative difficulties and the interlocutor proffers the element searched for. The following excerpt gives such a scenario. Here, A is doing a word search (clearly indicated as such by the lengthened sounds, ehms etc.).

(2.7) BrE - Derby 6: *Prompt after Communicative Difficulties*

A: you don't need any::: ehhm ( P A U S [ E ) eh ah overheads.  
B: [overheads,

In (2.7), speaker B offers the word A was searching for and A repeats it. This form of proffer-repetition does not count as quotation in my analysis. It is considered a case of echoic mention (Hymes 1987) and not role-play in the 'speaker as actor' framework of quotation. Also, I did not include cases of co-construction where one speaker jumps in and foreshadows the other speaker's quote such as in (2.8).

(2.8) BrE - Derby 2: *Foreshadowing*

D: yeah it was a case of "if you can['t beat them (...) jo(ha ha )in them" ah ha ha  
G: [beat them join them,  
D: yeah so I as asked to stand for the borough,

Here, G's speech *beat them join them* is a projection of D's utterance, a form of anticipatory overlap (due to the projectability of utterances, cf. Hutchby and Wooffitt 1998). Cases such as this are not included in the category quotation. It is not the case that the message is first produced by G and then quoted by D. Indeed, speaker D does not pick up G's utterance but only continues his own which is partly overlapping with G's. Hence, I interpret cases of this sort as instances of two voices one message, which serve interpersonal functions but do not count as cases where one speaker re-enacts the other.

The same also applies to cases of backchannel where one speaker repeats the other as in the following:

(2.8) BrE - Derby 6: *Backchannel*

A: I thought "brilliant,  
no kids".  
B: "no kids",

Again, in my interpretation, the three turn constructional unit (TCUs) in example (2.8) contain only one quote, A's *I thought "brilliant no kids"*. The concept of a TCU as the basic unit of talk has been established in Conversation Analysis by cf. Sacks, Schegloff and Jefferson (1974). A TCU can be sentential, clausal, phrasal, or lexical constructions. Every TCU contains a transition relevance place, which is a place of potential floor transition (where either the current speaker can select the next speaker, or the next speaker can self-select). Such TRPs are 'projectable' (viz. the conversational interlocutors can make predictions about when it is to end) and oriented to by the interlocutors (cf. also Hutchby and Wooffitt 1998). In example (2.8), B is not producing any quote but is merely echoing A's words. Cases like this where there is interspeaker alignment achieved through the sharing and repeating of the message, are not considered as re-enactments in the quotative sense.

I will now discuss at more length the case of partitioned quotes in the form *I said "quote" I said "quote"*. Generally excluded from the analysis were cases where the second instantiation is a verbatim repetition of a previous one. This can be seen in the next example:

(2.9) BrE - NC 14: *Full verbatim repetition*

- 01 X: I **says** "how much will you give me for give you to me for mine"?
- 02 Y: what did you get for it ?
- 03 X: he says "well" he **says** "well this price",
- 04 and you saw he sh- try put it off a bit you know,
- 05 Y: mh,

The first quote *he says "well"* in line 03 has not been included in the data. Rather, it is counted as the first instance of the second quote *he says "well this price"*, which is given in its entirety right afterwards. Consequently, in my analysis, line 03 only contains one quote, *he says "well this price"*.

The interpretation of two repetitive occurrences of a quotative frame + quote as instances of the same quote only applies to material which is exactly the same on a lexical basis and where no pronounced intonational difference occurs between instantiation one and two. This treatment applies to cases where (i) speakers clearly repair themselves and start a new sequence with exactly the same form as the

abandoned one or (ii) where speakers produce repetitions for stylistic reasons or for comprehension.

Generally, in the sequence *Pers.pron. V<sub>quot</sub>*<sup>10</sup> “quote 1” (*Pers.pron.*) *V<sub>quot</sub>* “quote 2”, a 1<sup>st</sup> quote is only counted when it fulfils the following prerequisites. A token is counted as an instance of quotation (i) if it is not repeated in the exact same form in the next quote as stated above (with respect to lexical, para- and non-linguistic factors as in example 2.8), and (ii) if the quoted material is a lexical item with clear semantic content (this excludes discourse markers) or an exclamation with a sound/voice effect. Otherwise I do not interpret the material following the first *personal pronoun V<sub>quot</sub>* sequence as a quote. This is shown in (2.10), where the first instance of *I said* is interpreted as a repair.

(2.10) BrE - Derby 4: *Repair sequence*

X: I said ehhrm I said “quote”

In (2.10), *I said ehhrm* is not counted as a quote in its own right if it does not have a pronounced intonation contour which would justify treating it as such. Bolden (to appear), Günthner (1998) and Klewitz and Couper-Kuhlen (1999) show that suprasegmentals such as pitch, intonation, register, lengthenings etc. are used to demarcate and interpret the beginning of quotations with no overt lexical signal (cf. also the discussion under unframed quotes later in this chapter). In prosodically unmarked cases, I interpret sequences such as *I said ehhrm* as repaired by the following sequence *V<sub>quot</sub>* “quote” *I said “quote”*. Hence, I consider (2.10) as two instances of the same quote instead of two independent quotes.

Only if the putative quotative frame contains lexical material or if the non-lexical material is produced with a clear intonation contour of its own are the tokens counted as quotation. As the next example shows, cases where the two potentially separate quotes have their own, non-identical semantic content and their own prosodic structure (as manifested by two complete intonation contours) are counted as two separate re-enactments.

---

<sup>10</sup> *V<sub>quot</sub>* here stands for any *verbum dicendi*.

(2.11) BrE - Derby 12: *Two separate quotes*

- 01 A: I says “shall we [ke-“  
→ 02 B: [“no”.  
03 he says “I want you inside so we can lock up”.  
04 so I’ve been back up this narrow little drive,  
05 he shuts the gate.

In line 02 and 03 there are two pieces of re-enactment which I interpret as two separate quotes, *no* and *I want you inside so we can lock up*. This decision is taken on the basis of their independent truth value and semantic content, and due to the fact that they constitute two separate intonation contours (two full falls, one after *no* and one after *up*, as is indicated by the full stops in the transcription, cf. Appendix 2).

In this example, the fall in line 02 and the new onset of an intonation contour in 03 signal that the quotative frame *he says* introduces the second quote rather than retroactively marks the first quote, *no*. In cases where the suprasegmental information does not disambiguate the direction of the scope, the question poses itself: to which quote does a quotation frame between two instances of reported speech belong?

The methodological assumption that I make here is that one quotative frame only has scope over one quote. If there is no clear intonational break either before or after *he says*, which breaks down the construction into coherent intonation units (Chafe 1988, 1994), we cannot disambiguate to which re-enactment *he says* belongs. In such cases, I go with statistical probability, which shows that the default scenario (over 99% of all clear tokens of quotation) is that the quotative frame occurs before the quote. Hence, if suprasegmental information does not indicate otherwise, I interpret cases like the above as two quotations, the first with a zero frame and the second framed by a graphic introducer such as *he says*.

Another problematic area is when quotes are broken up by insertion sequences. The question comes up: when do we consider the two parts to be one coherent entity and when do we have to assume that we have two separate quotes? Example (2.12) illustrates this phenomenon:

(2.12) BrE - NC 10: *Insertion sequence*

X: and I got to a half an hour later it just got a bit sick,  
and **said** “ahhh I’ll have to go”,  
(.) ha ha ha ha (.)  
“but ehh,  
call me back I’ll pay for that”.

Here, X’s self-quote “*ahhh I’ll have to go (...) but ehh call me back I’ll pay for that*” is broken up by a laughter sequence by the same speaker, represented as *ha ha ha ha* in the transcript. I interpret this laughter sequence, which is clearly set off suprasegmentally from the surrounding speech by micro pauses, as an interruption. Speaker X steps out of the quotation and adds the laughter sequence as a meta-comment to his narration. It does not form a part of the reported material. I consider cases like (2.12) as one quote with an insertion sequence when the interruption fulfils the following criteria<sup>11</sup>:

(1) It must not add anything to the propositional content of the quote. In the case of example (2.12), the laughter sequence does not have lexical status. Consequently, the continuation of the quote is an extension of the same message with only a meta-comment in-between. Any lexical item with clear semantic force would divide the quote in two. Hence, I allow for discourse markers and meta-comments such as voice, sounds, gestural effects and laughter in-between quotative sequences and still count the two parts as pertaining to the same quote.

(2) A non-continuous intonation contour needs to clearly set off the insertion sequence from the surrounding quote. Speakers in my corpus mark comments such as laughter sequences, but also discourse markers very clearly as either pertaining to the quote or as outside of the quote by the chosen intonation. (cf. Klewitz and Couper-Kuhlen 1999 for more a more detailed analysis of the prosodic marking of material with respect to its in-out status in quotative situations).

Having laid out the criteria for the delimitation of what I consider to be the subject of my analysis - reported speech and thought, framed or unframed, past or present, containing linguistic or extralinguistic material - it is important to

---

<sup>11</sup> Note that laughter sequences or other insertion sequences can still be part of the quote if they do not convey to the conditions mentioned in (1) and (2).



distinguish direct from indirect reported speech because I will restrict myself to the analysis of direct reported speech.<sup>12</sup>

### 3.3.1. Indirect versus direct quotes

Cameron (1998: 51, cf. also Maldonado 1991:20) contrasts description (indirect speech) and demonstration/re-enactment (direct speech). He adds that direct and indirect styles “fail to meet the basic criterion for definition as a sociolinguistic variable” as they perform different tasks in discourse (1991:48, cf. also Banfield 1982, Coulmas 1986, Leech and Short 1981, Lucy 1993, Maldonado 1991, Maynes 1990, Reyes 1993:41-42, Waugh 1995, Wierzbicka 1974). This fundamental breakdown seems to have been the common assumption in sociolinguistic treatment of quotatives; all studies I know of only consider direct reported discourse.

In this thesis, I will maintain the same subject of analysis in order to achieve general compatibility with previous work. I also agree with Cameron’s and Reyes’ claim about the general distinctness of direct and indirect speech. The following paragraphs sketch out how I define direct quotes and how I distinguish them from indirect ones. Clark and Gerrig (1990) demonstrate that direct and indirect reported speech can be differentiated in terms of them being exponents of different stylistic categories. They set up the opposition of more dramatic vs. more descriptive (cf. also Baumann 1986, Dubois 1989, Macaulay 1987, Short and Semino and Culpeper 1996, Sternberg 1982, Tannen 1986, 1989, Yule 1993). While this fundamental difference indeed holds in theory, the reality test in the form of a corpus of real occurring speech yields a surprisingly high number of borderline cases. I contend that the notion of what constitutes a direct quote has to be refined. While I acknowledge that the drawing of a line between direct and indirect quotes is a very slippery endeavour,

---

<sup>12</sup> The intermediate style, indirect free speech is often referred to in poetics as in Pascal (1977), Rabatel (2001), see also Fludernick (1993), Juillard (2000), Vuillaume (2000), Ehrlich (1990) and many others. I will not be concerned with this style of quoting and will restrict myself to the analysis of direct reported speech and thought. For a comprehensive treatment of direct and indirect styles consider Roncador (1988), Roeck (1994), from a cross-linguistic perspective refer to the edited volumes Coulmas (1986) and Lucy (1993).

I will now try to demarcate the related concepts of direct and indirect speech.<sup>13</sup> I will finish the definition by discussing my decisions in a few frequently recurrent ambiguous cases.

On the structural dimension, we can distinguish indirect from direct quotation because there is a formal prerequisite for indirect quotation. Indirect quotations require the quoted material to have the syntactic form of a coherent declarative sentence. Any utterance that does not syntactically fit the format required of the indirectly quoted material has to be restructured in order to be incorporated into the construction and to be embedded under the CP node. For example, direct questions and imperatives are structurally impossible in indirect quotes. Hence the sequence *he said that where is the bar* would not be interpreted as an indirect quote. The internal structure of the quote would need to be reorganized in order to be incorporated syntactically into an indirect construction. Equally, we do not find *no / yes /* or vocative items such as greetings and address terms directly following the *verbum dicendi* (+that) in indirect quotes. Furthermore, indirect quotations cannot start with discourse markers. Hence, *he said "sort of amazing that place"* is necessarily direct. But formal criteria alone do not fully delimit direct from indirect quotes. We will need to add other criteria based on functional principles.

Leech and Short (1981:32) claim that the effect produced by the use of indirect speech is one in which the reporter of the conversation intervenes as an interpreter between the person being spoken to and the words of the person being reported, instead of merely quoting verbatim the speech that occurred. Even though I do not accept the verbatim assumption for reported speech and even though it has been shown that direct reported speech also can contain interpretive elements from the reporting speaker, Leech and Short are right when they say that the point of view is a decisive criterion in the differentiation between direct and indirect speech. The fundamental difference between the two modes of speech representation lies in the perspective adopted by the person who does the reporting. I assume, like Clark and

---

<sup>13</sup> For other partially overlapping attempts to delineate direct and indirect speech consider Banfield (1982), Cameron (1998:48), Coulmas (1986), Leech and Short (1981), Lucy (1993b), Maldonado (1991), Maynes (1990), Waugh (1995), Wierzbicka (1974).

Gerrig (1990) and Cameron (1996), that direct quotations have a deictic orientation to the experiencer of the quote in spatial, temporal and personal deixis: *He said "I am leaving now"*, is a direct quote. This is in contrast to indirect quotes such as *he said (that) he was leaving now*, which are from the deictic viewpoint of the reporting speaker. Hence, the difference is the perspective from which the quote is reported (cf. Cameron 1996:51).

Note that I have not taken as a distinguishing characteristic for a direct quote two different criteria

(1) Whether or not there is a *that*. My data shows that many quotes feature the complementiser *that* followed by experiencer deictic orientation.<sup>14</sup> The following example shows that it is not the (non) existing *that* which makes a difference between indirect and direct quote but that it is the deictic orientation of the quote (*italics = voice effect*).

(2.13) BrE - NC 17: *Direct deictic orientation with that*

H: when they were teaching you to cook a potato.  
T: baked potato,  
and you think that "*well we're not stupid you know*",

In vein with Cameron's analysis for Puerto Rican Spanish (1998: note 9), where he concluded that *que* is not a decisive criterion for the delimitation of direct and indirect speech, I interpret quotes such as (2.13) as direct.

(2) Also, I didn't consider a decisive factor whether or not there are voice effects. Klewitz and Couper-Kuhlen (1999) show that voice effects are by no means restricted to direct quotes. Indeed, speakers use voice effects in indirect quotes quite frequently. The authors demonstrate that the difference between indirect and direct speech is much more fluid than currently claimed in the literature and that reporting speakers often move in and out of direct and indirect reported speech.

---

<sup>14</sup> Obviously, this interpretation opens up the question whether such quotes would have to be understood as repairs away from the indirect to the direct quotation. I will not dwell on this issue here and merely interpret such quotes as direct.

In sum, in this thesis I define indirect quotes as reproductions that have to follow structural requirements in order to be embedded and have a deictic orientation to the reporting speaker. They cannot include extralinguistic and pragmatic elements. By contrast, direct quotes are re-enactments of previous actions which are free in their syntactic form and have a deictic orientation to the experiencer. They can incorporate extralinguistic material and pragmatic markers.

I excluded quotations that were ambiguous between direct and indirect quotes. One reason for exclusion was ambiguity of temporal and personal orientation. This is most frequently the case when the quotative verb is in the present tense and consequently does not trigger any backshift in the subordinate if it is introducing an indirect quote. The example below is such a case:

(2.14) BrE - Derby 13: *Duplicating videotapes*

- M: marking all the videos,  
cause he's been taping them all,  
D: did he,  
M: I suppose [that] he's going to make em- almost taping all of them,

Here, the temporal deictic orientation in the quoted stretch lends itself to an interpretation as an indirect construction - a *that* can easily be substituted - giving the construction *I suppose [that] he's going to make em- almost taping all of them*, as well as a direct construction *I suppose he's going to make em... .* The same problem also occurs when the quotative verb is in the historical present (Johnstone 1987, Schiffrin 1981, Wolfson 1981, 1982). In the analysis, quotes of this sort, which could not be unambiguously identified as either direct or indirect have been disregarded. Also, all tokens where a possible embedding in an indirect quotative construction was merely a matter of clausal order have been excluded. The quotes have been excluded if syntactic incorporation into an indirect quote is possible if we swap matrix and subordinate. The syntactic non-possibility of incorporation in an indirect structure has to be a problem of general structural incompatibility rather than a mere problem of clausal order.

### 3.3.2 The problem of *think*

A frequent phenomenon in my corpus was the occurrence of the verb *think* as a modalising statement, as in the following example:

(2.15) BrE - NC 11: Think *in modalising function*

X: and I remember putting a pellet,  
Y: in,  
X: in the breech,  
Y: ah,  
→ X: and I think a cold button actually,  
got caught in the trigger guard.

Here, *I think* indicates that the speaker is telling the story from his perspective and from his substantially blurred memory given the fact that he is relating childhood memories. *I think* functions as a hedge and attenuates epistemic certainty (Fludernick 1993, Schneider 2002, Semino, Short, and Culpeper 1997). In this function, it has nothing of a quote and cases such as (2.15) have not been included in the analysis.

In some cases, though, it is slightly harder to judge whether or not to include *think* + potential quote in my analysis. This was especially the case where *I think* occurred in the present tense and was followed by material that could be considered reported as in the following example:

(2.16) BrE – Derby 13: Think *in modalising or quotative function*

X: so I was then eleven you,  
Y: you when we were a bit before then,  
**I think** we moved there when,  
we moved there when you were about 8 or nine,  
X: yeah I can't remember that time,

Like (2.16), many instances of *think* with first person in the present tense were hard to classify. They were somehow borderline cases between indirect and direct quotes. Also, they were ambiguous between *I think* as a hedge (Aijmer 2001, Holmes 1990, Schiffrin 1987, Simon-Vandenberg 2000) and as a lexical quotative verb. As the main focus of this analysis is not to understand the patterning of hedges, I have decided to leave out the tokens of *think* with ambiguity in key. This is because we would get too much noise in a statistical analysis if we left in instances which are

functionally hedges. As pointed out by Guy (1988), Young and Bayley (1996) and many others, when doing an analysis of variance such as a VABRUL or any other kind of a regression analysis, one should take out ‘knockout factor groups’, cases where no variation occurs. The normal procedure is to exclude categorical contexts (Blake 1997, Singler 2001). As *like* (and *go* in US English, cf. Butters 1980, Schourup 1982b) cannot occur with indirect speech, all tokens which are possibly occurring with indirect quotes (as well as ambiguous ones) had to be excluded from the analysis. This applies mostly to *think* in the present tense and the first person singular in environments such as in (2.16) where we cannot tell whether the quote is direct or indirect. But it also pertains to all other ambiguous cases as discussed above. As soon as any ambiguity in key arises, the tokens have to be eliminated.

Hence, my subject of analysis consists of all strategies speakers have at their disposition in order to signal that a re-enactment of previous behaviour is about to or has taken place. This is then demonstrated from the perspective of the reported speaker. I define the sum of enquoting strategies as the (say) variable. This is on the basis of the fact that the quotative devices have “one common function in discourse” (Dines 1980): they frame re-enactment of previous states or actions. As the (say) variable is functionally defined, we will have to admit any device that functionally falls within this category as a member of this category. When looking at a data-base of talk-in-interaction, we find that numerous quotes are not framed by a verbum dicendi (cf. examples (2.1)-(2.3) taken from Cameron 1998). I will now discuss this category in more detail.

### 3.3.3. Unframed quotes<sup>15</sup>

Güldemann (2001:28) claims that unframed quotes are strikingly rarely thematised in the literature. But while studies in the field of discourse analysis do tend to refer to this option, sociolinguistic research indeed includes this quotative

---

<sup>15</sup> Unframed quotes have been called “freestanding” by Clark and Gerrig (1990), “zero quotatives” by Mathis and Yule (1994), “unintroduced dialogue” by Tannen (1986) and “bald” and “unbracketed quotes” by Romaine and Lange (1991), “free direct speech” by Leech and Short (1981) “unidentified speech” by Wiesemann (1990), “null quotation formula” by Longacre (1994), “zero quotatives” by Yule (1993) and Mathis and Yule (1994).

option surprisingly rarely. Because such an omission of a variant “gives a distorted picture of the formal aspects of the entire domain” (2001:29), a balanced variationist study of the quotative system needs to include this quotative option.

With respect to the interpretation of unframed quotes, I follow Güldemann (2001:29) and consider them not as a case of structural deletion of the surface marker. Rather, I interpret unmarked quote as a variant in its own right, albeit with no overt lexical signal. Note that this definition considers marking/not marking a binary choice with no one member conceptually more basic than the other (see Ferrara and Bell 1995). Couper-Kuhlen and Klewitz (1999) and Günthner (1998) have shown that lexically unmarked quotes are zero on the surface but clearly demarcated from the surrounding material by their intonation, higher expressiveness of voice, pitch movements, pauses, lengthenings and other suprasegmental devices.<sup>16</sup> Other strategies for the disambiguation of the following quote as reported are (i) initial *oh*, which is, following Heritage (1984:299) “a change of state token which shows some kind of change in locally current state of knowledge, information, orientation, or awareness”, viz. a change in footing (Goffman 1979) (ii) and anaphor resolution according to the preceding and following context (Marslen-Wilson, Levy and Tylor 1983:361). Tannen (1986:323) set up a continuum of quotative options with graphic introducer verbs at the one pole and more expressiveness of voice at the other.

Consequently, in this analysis, quotatives without a frame are included in the (say) variable as well. By doing so, we can “close the set that defines the variable” (Labov 1996:78). Unframed quotes, as I will call them from now on, form one variant, albeit a zero variant, of the category.

As unframed quotes, I counted completely zero quotative frames (as in example 2.1) as well as quotatives that continue the experiencer subject<sup>17</sup> (+dummy verb *be*) as is illustrated below.<sup>18</sup>

---

<sup>16</sup> Some studies have shown that suprasegmental patterns have acquired a certain degree of conventionalisation (McGregor 1994:76), Du Feu (1996:14).

<sup>17</sup> The fact that I found many quotative frames of the form *and NP [be]* (cf. also Clark and Gerrig 1990:772-note 9) where the verb in brackets is a form of inflected *to be* runs counter to Cameron’s native speaker intuitions (1998: 58).

<sup>18</sup> See Cameron (1998:54) for the discussion of the same kind of construction in Spanish (with *ser* and *estar*). In his analysis, quotes of this structure count as embedded, though.

(2.17) USE – File 2464 *Unframed quote with experiencer and be*

- A: no I could smell an accent a mile away,  
 B: oh isn't that only people from there can tell.  
 → everyone else is "where did you get that weird accent"?  
 A: are you in Pennsylvania or Dallas right now?

In example (2.17), the narrator mentions the authors to whom the speech act is attributed (a generic *everyone*). No quotative verb frames the re-enactment. Instead, the narrator gives the source of the quote as well as the appropriately conjugated form of the verb to be. In my analysis, such cases also count as unframed. Hence, the category 'unframed' consists of quotes which are not framed by any item with inherent quotative function such as a the verba dicendi *say*, *think* etc. or other items which have grammaticalised or are en route to grammaticalization as quotatives such as *go*, *like* and *be all*.

In sum, my analysis covers the full range of means which speakers make use of in order to enquote reported speech, sound, gesture and thought. The (say) variable can be broken up as follows.

**Table 2.1: Frequencies of occurrence of variants of the (say) variable**

	British English		US English	
	N	%	N	%
<i>like</i>	93	4.5	121	8.8
<i>go</i>	264	12.8	80	5.8
<i>say</i>	967	46.9	562	41.0
<i>think</i>	175	8.5	121	8.8
<i>unframed</i>	462	22.8	298	21.7
<i>tell</i>	23	1.1	42	3.1
<i>other</i>	80	3.9	147	10.7
<b>Total</b>	<b>2064</b>		<b>1371</b>	

The data-driven analysis ensures that all variants, whether marked lexically or suprasegmentally, were included in my corpus and were counted as members of the (say) variable. This analysis covers the full spectrum of quotative devices starting from the graphic speech verbs, through the "new" quotatives *like* and *go* (there were only very few instances of *be all* in my corpus), to unframed quotes. I have thus closed the set that defines the variable and shown where quotation occurs. But note



that there is in my view no way to predict where quotation does not occur other than a genre-specific analysis.

Comparatively few examples of syntactic changes in progress have been located. (Labov 2001:12)

## Chapter 3

### Intralinguistic Factors

In this chapter, I conduct a quantitative analysis of the linguistic constraints of all the variants within the (say) variable. I especially endeavour to explain which functional niches the newest arrival, *like*, has appropriated for itself and will compare its behaviour with the patterning of the older non-canonical variant *go*. I have chosen to restrict my investigation to the US contingent of my corpus, the Switchboard corpus, because *like* was already considerably widespread in the US data at the point in time when the recordings were made (1988-1992). Given the fact that Butters (1982) was the first to note quotative *like* in Californian English, we can assume that the restructuring of the quotative system in this variety has had sufficient time to adapt to the intruding newcomer and settle down in its richer (+ *like*) form. In Chapter 5, I will compare the results reported here with the British English corpus in order to show how the British system, where *like* is a younger addition, has reacted to the more recent addition and whether we can postulate cross-varietal trends.

The quotative cohort had to react in some way to the intrusion of two new members into the set of available options in order to accommodate these new members. This chapter will give evidence of how the quotative cohort re-organised. It will become evident that the quotative system of US English has become fully restructured in order to accommodate additional items. It is not the case that quotative items follow simple patterns of probability matching (Labov 2001) in transmission where patterns of distribution are reproduced by later generations. Rather, because individual quotative variants have a specific function in the pool of quotative devices and are far from being semantically empty, exchangeable, variants such as [in] and [iŋ], the refunctionalisation of all its members proceeded in a systematic way, by reallocating the functional load of introducing different types of quotations amongst the larger set of possible variants. Later, I will give evidence that

this functional allocation is far from arbitrary; rather, it is motivated by semantic – pragmatic features (cf. Chapter 4).

The linguistic facts presented here support the claim made in Buchstaller (2001b) that the newcomer quotatives have linguistic as well as social significance (cf. Chapter 6 for their extralinguistic patterning). It will be shown that we witness an important division of labour within the paradigm of reporting devices. The “new” quotatives *like* and *go* have taken on quite specific functions within the pool of quotative devices while older quotative variants have been functionally reallocated to other functional niches. Evidence of their functional load underlines the new quotatives’ status as full members of the quotative cohort with their own functional niches. Hence, the new quotatives, *go* and *like*, are not just “picked up” from high status reference groups as has been suggested by much of the sociolinguistic literature to date. Rather, the task-sharing amongst the old and new quotative variants provides the newer members with a linguistic *raison d’être*. *Like* and *go* are far from pleonastic, intrusive items in the pool of quotatives.

This chapter is structured as follows: the first section explores the use of the quotative variants with hypotheticality levels. I will investigate their patterning with respect to the epistemic level of the quote and then compare this distribution with their occurrence with respect to contextual features (such as time reference, co-occurrence with a defined speaker and listener, etc.). It will be shown that epistemicity is closely mapped onto contextual factors. Then, I will explore the priming effects that hold amongst the quotative options, both on a structural level and as concrete lexical priming effects. And finally, I will shed some more light on the patterning of old and new quotative options with mimetic enactment. At the end of this section, it will have become evident that that *go* and *like* have a wealth of functions and are much more than mere intruders, or trendy fads.

### 3.1. Epistemicity and Quotation

Modern English is rich in verbs representing speech and thought. But as Romaine and Lange (1991) point out, the English system is not very good at distinguishing between and linguistically marking hearsay and factuality. This makes English unlike German, French and Turkish where epistemicity is marked via inflection of the verb. In earlier stages, English had different resources for representing *de dicto* domains in general, and more specifically for the reporting of speech and thought (see Dirven et al 1982). Vendler (1972, in Romaine and Lange 1991: 264) has shown the close connection between mental verbs and verbs of saying as well as of language and thought, claiming that this has to do with the close parallels that exist between forms of thought and forms of speaking (cf. also Frajzyngier 1991). Consequently, in English and many other languages, we notice a very close relationship between descriptions of mental states and the reporting of speech. Traugott (1986) claims that where non-speech act and speech act verbs co-exist, the former will emerge first. She postulates a cline mental verb > speech act verb.

Generally, a quotative situation pertains over two periods of time. Let us assume that  $t$  is the point in time of the initial mental/verbal activity and  $t_{+1}$  is its rendering as a quote between interlocutors. For example, imagine a situation where speaker A tells his brother at  $t$ , Christmas Day, *I forgot to buy you a present*. This speech act can be rendered at any given  $t_{+1}$ , say, when A is chatting to his buddy B on New Year's Eve as *I said "I forgot to buy you a present"*. Conversely, if speaker A thought at  $t$  *damn, I forgot to buy him a present*, this can be rendered at  $t_{+1}$  as *I thought "damn, I forgot to buy him a present"*. The difference between reported inner monologue and reported real occurring speech is their (non-) wording at  $t$ . Real occurring reported speech was realized at  $t$ . Other forms of reported utterances might or might not have been realized at  $t$ . For example, A might have mumbled *damn...* at  $t$  or even only screamed it inwardly with anger. But the important thing to notice is that, irrespective of their initial form in  $t$ , both hypothetical inner monologue and real occurring reported speech are uttered out aloud at  $t_{+1}$ , and this constitutes the actual quote.

My data shows that even when speakers use the most “committed” speech introducer, *say*, we do not always have certainty that the actual words have been spoken in t. This made me interested in how narrators, when quoting, index their relationship and attitude to the quote. In this section, I will explore how speakers index their relationship and attitudes to the quote and how they express the general probability of the occurrence of the quote. It will be shown that *like*, *go* and other quotative options are used to mark the degree of hypotheticality of the quote. This section is structured as follows: I will first explain my notion of hypotheticality and how it pertains to the study of quotations. Then, I will show how the quotative variants pattern with respect to epistemic levels.

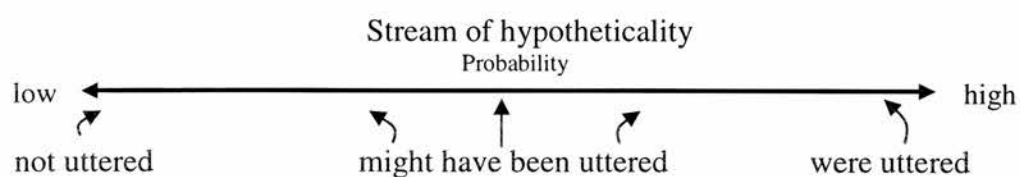
Hypotheticality is the probability with which a state of affairs or action is realised. For Standard English, the conditional sphere has traditionally been divided into a bi-partite (Quirk et al 1985) and tri-partite (Johnson-Laird 1986) system. A variety of criticisms have been levelled at this partitioning: Thumm (2000) shows that the criteria for determining which category to put a token into are often not based in the data but artificially determined (often with a glance to Latin grammar). Furthermore, Couper-Kuhlen (1999) argues that the boundaries between epistemic categories are fuzzy and that individual instances cannot be subsumed under 2 or 3 degrees only. Both, Thumm and Couper-Kuhlen demonstrate that the actual hypotheticality level is contextually determined, dependent more or less on the situation, and a matter of pragmatics. Comrie (1986:88) notes that different languages differentiate different degrees of hypotheticality. Buchstaller (1999) shows that in a Creole language various lexico-syntactic means can grammaticalise for the marking of hypotheticality and that those devices and the epistemicity levels they relate to are merely a matter of convention (cf. Hopper’s 1991 concept of emergent grammar).

Taken together, these arguments speak for a more fluid conceptualization of probability. In this chapter, I will follow Comrie (1986) and Akatsuka’s (1986) proposal of a factualis – realis continuum instead of arbitrary divisions of the epistemic sphere. One pole of this continuum is held by factualis actions or states and

the other by counterfactual or counter-to-expectation events or states. In-between those two endpoints, there are various intermittent steps of epistemic stance along the continuum. The sphere of probability/hypotheticality is finely differentiated into a multiplicity of epistemic stances. The whole range of epistemic relations ranges from realis over possible and hypothetical to counterfactual.

Now let us transpose the notion of hypotheticality, the probability of realisation of states of affairs, events or action, into the realm of quotation, which is the reiteration (in  $t_{+1}$ ) of words, sounds, and gestures that have already been produced (in  $t$ ). I noted that there is a probabilistic relationship between the quote and the original utterance. We can distinguish between utterances that could have been spoken out aloud in  $t$  (with various degrees of probability) and those that were definitely spoken out aloud in  $t$ . This means it is always possible and sensible to ask: How probable is it that the actual utterance was realised as an outward, overt speech act? The narrator in  $t_{+1}$  can present the quote as counterfactual or as realis or somewhere in-between. This builds again on Comrie's (1986) notion of an epistemic continuum where boundaries between categories are fuzzy. If, following Comrie (1986) and Akatsuka (1986), we conceptualise the sphere of hypotheticality as contingent and polar with a finely differentiated multitude of epistemic stances in-between, there is no reason not to believe that quotes might also be located along a cline of epistemicity (cf. Semino, Short and Culpeper 1997). The hypotheticality continuum for quotations can be conceptualised as follows:

**Figure 3.1: The Hypotheticality Continuum**



If there is a factualis relation between  $t$  and  $t_{+1}$ , the quote is located at the rightmost pole of the continuum. Quotes where the original utterance has not been produced at  $t$  but is nevertheless portrayed as a reproduction in  $t_{+1}$ , viz. where the relationship between  $t$  and  $t_{+1}$  is that of non-occurrence (counterfactual or counter-to-expectation), are positioned at the leftmost pole of the continuum. The body of the

continuum consists of quotes with various epistemicity levels from more to less hypothetical. Consequently, we can differentiate between quotes that were uttered aloud in t (on the high probability pole), quotes which definitely were not uttered out aloud (at the counterfactual side), and quotes which might or might not have been uttered (which span the continuum in-between).

The question then is whether taking this step sheds any light on the confusing field of different quotative variants. The following sections seek to answer this question.

### **3.2. Categories of Realisation**

I will now determine how individual (say) variants pattern with respect to epistemicity levels. In order to do this, I have set up three rough categories of quotations, depending on their epistemic stance, which I termed realis, hypothetical and situational. I will exemplify these categories at length below. For the analysis, all quotative tokens were categorized as one of the three categories of use. I will now explain the procedure that was adopted for the determining and coding for epistemic stance. Please note that the exact parameters that were used for the classification of quotes into the categories of use, as well as the contextual factors involved, will be described at length in the discussion in sections 3.2.1-3.2.3. and 3.3.

The coding of all quotative tokens in the data was done in a three-way process.

- (i) Every instance of quotation was identified in the data.
- (ii) As a next step, all instances were coded as realis, hypothetical or situational. The coding procedure was as follows: First, the quotative frame was deleted and substituted by a dummy, QUOTE. This was done in order to ascertain that the epistemic stance of the singular token was determined on the basis of the quote and its context and not on the basis of the framing quotative. The tokens were given with as much context as was needed in the singular instance for the determination of the contextual variables (see below). Every quote, which was presented in a large chunk of speech, was listened to and judged for its epistemic level. The actual coding for

category of use was done on an intuitive basis. Later in this chapter, I will give evidence that there is independent validity to this categorization. An independent coder and I coded several instances together.<sup>1</sup> After several initial sessions of joint coding, the coding was done by myself with regular crosschecks by an independent coder, checking problematic cases as well as random tokens from my data. In order to verify my coding system with other native speakers' intuitions, I counterchecked the intuitiveness of the categories as well as the allocation of several random tokens with two native speakers of British English, to whom I explained the mechanisms. I also discussed several instances with those two informants along the way. For consistency purposes, regular counter-checks with older coding were done. Finally, 39 ambiguous cases (2.9% of the overall corpus) were eliminated from the analysis. (iii) In a third step, the individual quotative tokens were coded for the presence or absence of various contextual variables (such as a defined speaker, hearer, situation etc). I shall explain this process in detail below. Coding every token for contextual rootedness (as defined by the aforementioned contextual variables) enabled me to correlate the category of use of the quote (realis, hypothetical, and situation) with the contextual factors in which it occurred. Chapter 3.3 will give evidence of the significant correlation between probabilistic occurrence of overt contextual factors and intuitively determined categories of use. We find that the epistemic stance of a quote very nicely maps into its degree of contextual rootedness.

I will now exemplify and discuss the three categories of use for quotations, realis, hypothetical, and situational.

### **3.2.1 Realis**

Realis quotes are re-productions of past utterances that did occur in t. They are at the high probability pole of the continuum. Example (3.1) shows a quote of this type:

---

<sup>1</sup> Thanks to Miriam Meyerhoff for help in setting up the procedures and subsequent checks on problem cases.



### (3.1) *Being mistaken for a woman*

- 01 A: the other day I went into a bar and this guy asked me to dance,  
02 B: ha ha ha [ha ha ha ha  
03 A: [and all he saw was my hair,  
04 and **he goes** “do you wanna dance”?  
05 I turn around and **go** “what”?  
06 B ha ha ha .hhh  
07 A: and **he goes** “do you wanna dance”?  
08 **I go** “no no”.  
09 **he goes** “oh oh I’m sorry”.  
10 **I go** “yeah you better be”,  
11 **I go** “you better be”.

Here, structural properties ratify the quotes introduced with *go* as real occurring speech. The two exchanges in lines 04-05 and 07-08 are question and answer scenarios. Sacks, Schegloff and Jefferson (1974) note that a first pair part such as a question (lines 04 and 07) structurally requires an answer (lines 05 and 08). If no answer had occurred, it would have been noticeably absent (Hutchby and Wooffitt 1998). Hence, the quotes can be classified as ‘realis’ speech. The “second turn proof procedure” (Sacks, Schegloff and Jefferson 1974) ratifies their interpretation as real occurring speech.

Note that the classification of quotes as realis does not rely on the verbatim postulate (Leech 1974, Genette 1980). Rather, the interpretation of quotes as realis is in accordance with Tannen’s (1986) concept of “constructed dialogue”, whereby quotes are considered approximative depictions of an earlier action (cf. also Clark and Gerrig 1990). The approximate nature of the depicted utterance is due to personal constraints on the speaker and because quoting always also involves interpretation and often contains “internal evaluation” (cf. Labov 1972c). Speakers may enhance speech acts in order to create involvement and make the situation more vivid (for the concept of “staging”, see Clark and Gerrig 1990, Wierzbicka 1974).

### 3.2.2. Hypothetical

Ferrara and Bell (1995:279) observe that a clear boundary between reported speech and thought is hard to draw (cf. also Tagliamonte and Hudson 1999, Vincent and Dubois 1996, Chafe 1994). This is because speakers often express a past attitude

or opinion in the form of reported speech in order to make inner states available to hearers (Ferrara and Bell 1995, Nordberg 1984). Furthermore, because such packaging of thoughts into the format of a quote makes the narrative more vivid and engaging, a narrator can heighten the dramatic impact of the story and thereby secure a claim to the floor.

In such cases, it does not make a difference whether the original speech act was actually uttered aloud or not. We are dealing with quotes whose function is comparable to Goffman's (1978, 1981:114-116) "response cries". Goffman defined these as used to "show or index the mental state of the transmitters [...] to clarify the drama of their circumstances". This is exemplified by the next stretch of speech.

### (3.2) *Plastic bags*

- B: yeah in fact I have one today,  
A: ri[ght.  
B: [the only problem with those is sometimes they got holes in the bottom.  
A: yeah [they  
B: [and ha ha **it's like** "whoops there goes my chips,  
A: [yeah  
B: [okay fine".  
A: uh huh

The larger conversation makes it clear that we are to imagine the following situation. There is a woman (B) with her chips in a bag. She may be alone. The bag breaks and the chips fall out. Now imagine what one would say in a situation like this: one might just be angry and swear inwardly. One might even swear out loud. But in the absence of an interlocutor, to whom is this talk addressed? The chips?

Such verbalizations of mental states are exactly what hypothetical quotes are about. Because hypothetical quotes are often attitudes or opinions expressed through and in the form of a quote in order to make them more vivid, it does not make any difference whether they are uttered aloud or not. Quotes in this category fall into what Chafe (1994) has called "verbally uncommitted thought". Their status as verbal or non-verbal, or even a combination of both, is left completely open. If uttered out loud, they have the function of putting into words (at least in the speaker's and hearer's now) what was going on in the mind of the person presented as the animator of the quote. But as Goffman (1981:97) points out, with no one present the quote is likely to be omitted altogether (cf. also Tannen 1986, Yule and Mathis 1992). In

cases of self talk, it is impossible to know whether the words were actually spoken and it is even less likely that they were heard (cf. Vincent and Dubois (1996)).

The category hypothetical is set up so as to span the whole spectrum of hypotheticality and not to break it up into arbitrarily chosen epistemic stances. Quotes in this category can be situated anywhere on the continuum.

### 3.2.3. Situational

As I have explained above, quotes have an inherent probabilistic relation between the utterance in  $t$  and the report of it in  $t_{+1}$ . This epistemic stance can either be very high, as for realis quotes, or somewhere in-between, as for hypothetical quotes. For situational quotes, there is a counterfactual relation between  $t$  and  $t_{+1}$ . Hence, there is no communicative situation in the past - because there never was any original utterance - but only in the present (in  $t_{+1}$ ), when the quotation occurs (cf. Vincent and Dubois' (1996) concept of "assertion"). This means that the original speech act is only portrayed as such by the reporting speaker. Example (3.3) provides a nice case in point:

#### (3.3) *Cooking*

- B: so I enjoy you know cooking things to take over to her hou[se or-  
A: [oh that is nice,  
B: yeah and it is fun for me to do that,  
it's something I enjoy doing,  
it's funny though **it's like** "I don't really want to cook for us" ha [ ha  
A: [ yeah ha ha

In this stretch of speech, it is very unlikely that B had ever said "*I don't really want to cook for us*", which she is now reporting to A, to anyone else before. It is also possible that she had not thought it before either. The situation is more of the kind where A says to B *I do not really want to cook for us* in their interactive present. There are no past events, no reproduction. This means that the only interlocutors involved are the current speaker (B) and the current hearer (A); there are no past speaker and hearer. The quote *I don't really want to cook for us* can be understood as a comment on the past situation clad in the format of a quote. Consider also the following example:

### (3.4) *Playing golf*

B: well I watch my husband swing,  
and his swing you know compared to the pros is so,  
it's ridged,  
**it's like** "he's trying so hard.  
it's not that flowing movement".

Here, again, it is very unlikely that B had ever said *he's trying to so hard* at  $t$  and is reporting this utterance to her interlocutor in  $t_{+1}$ . This is because there was no real interactively communicative situation in the past. It also does not seem that she thought *he's trying so hard* at  $t$ . Rather, like (3.3), her utterance is more a comment on the situation at  $t$  clad in a quote. It is not a re-telling of an original quote. *It's like* here can be glossed as "*my perception of the situation is such that it [=the situation] is like x*"

Because this category of quotes somehow straddles the border between a description of a situation and a quote, I have termed it situational. And indeed, it is often only the added or not added voice effect (Klewitz and Couper-Kuhlen 1999) that helps to distinguish between a description of a situation and a reported quote.<sup>2</sup> Vincent and Dubois (1996), while discussing this category of quotes (called "assertion" in their terminology), point out that presenting information as reported speech instead of as mere description gives it emphasis and makes it more vivid. They also argue that by packaging a depiction of a situation into the format of a quote, speakers assume full responsibility for what they are saying as they openly state its subjective, agent-based status. Hence, for assertions, aka situational quotes, the introductory verb can be replaced by the propositional attitude "*it seems to me that*". A situational quote indexes a speaker-infused rendering of the situation, it shows that the perspective is that of the current speaker (cf. Sanders and Redecker 1996). Quotes in the situational category can be understood as comments or assertions (Pomerantz 1979) clad in the form of a quote. The moment of speaking is in the now and becomes the quote.

---

<sup>2</sup> While I have not considered suprasegmental features for the delimitation of the three categories of use, intonation and prosody did inform my choices of whether or not to count something as a quote in the first place (cf. Chapter 2).

In conclusion, a relationship holds between the original utterance and its reproduction. This relationship is one of more or less probability of occurrence of the original utterance in *t*. The epistemic modality of the quote refers to the different ways in which speakers can show their commitment or their attitudes towards what they are reporting (Karkkainen 1999). They can portray quotes as situated somewhere on the hypotheticality continuum for quotatives, which I have split into three rough categories for this investigation.

The next few paragraphs discuss the importance of contextualization (Auer 1992:4) in the determination of the epistemic stance of the quote. I will show that the probabilistic co-occurrence between epistemicity based categories of use and contextual factors gives independent validity to the intuitive categorization. It will become clear that the epistemic continuum matches very nicely into a contextualization continuum.

### **3.3. Contextualization and Hypotheticality**

Fillmore (1992) maintains that speakers have to open up an alternative world if they want to express potential states and that they use different means to position speech along a continuum of hypotheticality. But, as Thumm (2000) has cogently noted, high hypotheticality need not be marked on the verb or even be signalled by overt markers as long as the broader context provides the necessary conceptual features (cf. also Snitzer-Reilly 1986). He argues that the epistemic stance of a singular instance relies heavily on the context and does not have its locus in one feature only. Couper-Kuhlen (1999) displays the variety of means developed by speakers of Standard English for expressing epistemic stance in conditionals. Moreover, Buchstaller (1999) and Couper-Kuhlen (1999) have shown that while grammatical constructions get conventionalized as a means for the marking of epistemic notions, speech communities develop methods for more subtle differentiations within the epistemicity continuum. Some are more grammaticalised and some less. The above mentioned research underlines Johnson-Laird's (1986) claim that in English, content and context signal (i) that a conditional notion is involved in talk and (ii) what degree of hypotheticality is involved.

Transporting these insights into the realm of quotation, I will now show that there is a direct statistical correlation between contextual factors and the epistemicity level of the quote. In a data-driven analysis, I have isolated contextual elements which can serve as contextualization cues (Auer 1996, Gumperz 1989, 1992, 1997) for interpreting the epistemic level of the quote (cf. also Karkkainen 1999). I will now demonstrate their patterning with respect to the hypotheticality level of the quote. On a frequency-based scale, we can show that quotes with higher probability are statistically more likely to occur in situations where contextual features are specified.

As I have laid out in the methodology section above, all instances of quotes were tagged for contextual factors in their environment. I will now describe in more detail how this tagging was done:

There was no pre-determined frame of analysis. Contextualization cues can have scope over hours and hours of conversation. For example, once the location of a conversation has been mentioned, it can be considered as given and need not be restated in later talk any more. For this thesis, I have restricted the contextual factors to be included in the analysis to five: a defined time and situation in which the quote is embedded, a known and given speaker and addressee, and the conversational necessity of the quote.

Obviously, the fact that someone is present does not make him or her automatically an addressee (see Goffman 1981, Kang 1998, Levinson 1983). In the definition of an addressee, I follow a deixis model of conversation, whereby participants are classified according to their status in the exchange. As conversational necessity, I interpret quotes which are structurally required by their sequential position in talk-in-interaction (cf. the notion of adjacency pair in Conversation Analysis i.e. Sacks, Schegloff and Jefferson 1974). Consider for example the exchange in (3.1), where a question-answer relationship makes the answer conversationally required. Also included in the category of conversational necessity were contexts in which the interlocutors uttered meta-comments (referring to overtly produced speech) on what/when/how something was said by the same interlocutor who quoted. For example, some interlocutors commented that that they found an utterance annoying or wonderful or too loud. Cases which were referred to as printed

sources were also included in this category. Finally, Vincent and Dubois (1996) conclude that a case of conversational necessity is given when an action causally follows out of the utterance of a quote. Importantly, such actions would **not** have followed out of inner speech (cf. Ferrara and Bell 1995:279, Tagliamonte and Hudson 1999:156). In this category, I include cases where a quote leads to the drawing of a conclusion or some knowledge or an action, which would not have taken place, had the quote consisted of inner speech or mental action only.

I will now illustrate the coding procedure: Consider example (3.5):

### (3.5) *Getting Married*

S:     you came down to the Chronicle offices to seek me.  
          and you waltzed me into the registry office,  
          because I wasn't 21.  
          and you got us into Sattle Road,  
          before you really said what you were gonna do.  
          ((11 lines deleted))  
          you never ever asked us,  
→     you just waltzed us up the registry office,  
          and I **said** "where are we going"?  
          "you'll see you'll see".

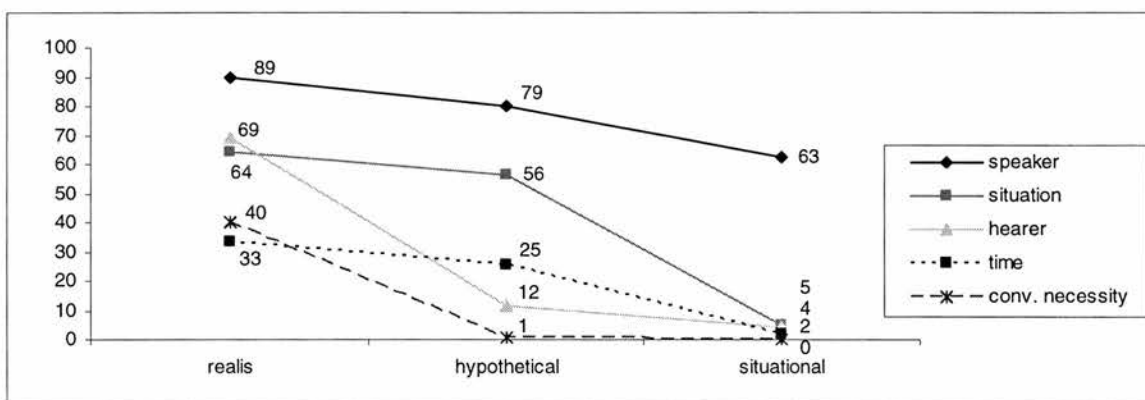
The quote framed by *I said* in (3.5) is embedded in a clearly defined interactive communicative situation. Both the speaker and the hearer in *t* are specified as the husband and wife who are present at the moment their original speech is re-enacted ( $t_{+1}$ ). The original utterance occurs at a real and given time, the day they got married. Furthermore, a real location (on the way to) *the registry office* is given. Also, the answer *you'll see you'll see* is contingent on the question *where are we going?*, which in turn conversationally requires an answer (as a second pair part). I have adopted a binary coding method whereby every single instance of a quote is marked for the presence or absence of all the respective contextual features. The quote framed by *I said* in the line marked by the arrow was consequently coded as [+speaker], [+addressee], [+time], [+situation], [+conversational necessity].

It has to be borne in mind that we are looking at a discourse variable and this makes it tricky if not impossible to quantify exhaustively or to specify an envelope of variation (cf. Holmes 1989:297, Meyerhoff 1994, Rickford and McNair 1994). But in

an attempt to capture the fact that speakers index and, conversely, hearers recognise and interpret the intended degree of hypotheticality of a quote, I have set up a contextualization matrix (as depicted in Figure 3.2 below) with the five factors I coded for. Such a matrix enables us to investigate the occurrence of categories of use (determined by degrees of epistemicity) with contextual variables. A frequency-based representation of the categories of use by contextual rootedness has the advantage that it can depict them as fluid along the continuum of hypotheticality.

The following figure shows the correlation between hypotheticality levels (the x-axis) and degree of contextualization (as a percentage plotted on the y-axis) for every contextual factor involved. There is a general trend in that slowly falling values for contextual grounding correlate with lower levels of epistemic stance.

**Figure 3.2: Co-occurrence of contextual factors with degrees of hypotheticality**



We see that realis quotes achieve the highest values on all 4 parameters. Out of all realis quotes (N=372), 89% are marked for speaker, 69% for a specific hearer, 64% for a specific situation, etc. We can now set up an overall index-scale which gives the ratio  $\frac{\text{actual contextualization}}{\text{possible contextualization}}$ , whereby actual contextualization is the number of quotes in the respective category (here realis) that are marked for a contextual factor, and possible contextualization is the overall number of tokens in this category. Realis quotes achieve the overall contextualization ratio of .59.<sup>3</sup> In a comparison with

<sup>3</sup> This index is an attempt to statistically represent the patterning of discourse features. Such a probabilistic representation enables us to draw comparisons between categories. For a discussion of other indexes see Meyerhoff (1994) and the articles mentioned therein.



quotes of other degrees of hypotheticality, realis quotes scored highest of all. Hence, realis quotes occur in a more plausible and concrete utterance situation, which can be numerically estimated by a frequency estimation of the contextual factors of time, situation, speaker, hearer, and contextual necessity.

Hypothetical quotes (N=287) have intermediate values. Vincent and Dubois claim that for this category of quotes (which they call “pseudo-reproduction”) there are “insufficient indications to clearly place the words in a precise context of being uttered” (1996:366). The contextualization matrix above shows that Vincent and Dubois’s intuitive claim is borne out by the facts: hypothetical quotes are less contextualized than realis ones. They only achieve a contextualization index of .35. We notice that while contextualisation values for hypothetical quotes are below the ones for realis quotes with respect to all factors, it is especially in the category ‘hearer’ and ‘conversational necessity’ that hypothetical quotes lag behind. The main difference between realis and hypothetical quotes seems to be the presence or absence of a hearer and of conversational necessity. While the realis score of .59 is not an indication of overwhelming determination, there seems to be an important step in epistemic level attributed to quotes which have a determined hearer and are conversationally required. This step manifests itself in the classification of the quotes as [+/- likely to be uttered].

It needs to be stated that the investigation of how speakers index and hearers interpret the epistemicity of more or less probable quotes within the overall category hypothetical is a topic that deserves further research. In order to do so, we would need to fine-tune our analysis of contextual situatedness via contextualization factors and include more (quantitative and qualitative) contextual indications. For the present analysis, however, it will suffice to classify quotes with undetermined epistemic level as pertaining to the category ‘hypothetical’; I will not delve into a further sub-classification at this point.

Situational quotes have not been uttered before. Hence, there has never been a past illocutionary event in *t*. There is no other interactive context where speech or thought really occurred other than that of the interactive event in progress between the interlocutors in their current now. This is manifest in the very low contextualization ratio of only .15 for situational quotes (N=204). Notice that on the right hand side of

the continuum, the values in all categories except the feature speaker (63%) have been reduced to values under ten percent (cf. Vincent and Dubois 1996). But even with respect to the contextual feature speaker, the values are lower than for the other two quotative categories, realis and hypothetical.

Figure 3.2 shows that as hypotheticality levels rise, the values for contextual factors decline. Hence, there is a clear overall correlation between contextual factors and hypotheticality levels. The hypotheticality continuum is neatly mapped onto a contextualization continuum. This shows that assigning something to one or another level of hypotheticality, if intuitive, is not entirely subjective. The above matrix gives evidence that choice of category can be operationalised in fairly reliable, independent terms (such as the presence or absence of contextual features). The differentiation between epistemicity levels of quotation can now be more than just intuitively fixed but firmly rooted in their patterning with respect to contextual variables.

We will have to assume a circularity of interpretation: if realis quotes usually occur in defined and plausible communicative situations (Vincent and Dubois 1996:366) and hypothetical and situational quotes with less and less contextualization respectively, it is reasonable to assume that speakers exploit this fact. In order to express probability levels, they can present quotes with different levels of contextual rootedness (different degrees of contextual factors). Hence, we have a fluid contextualization continuum, which is in fact indicative of the intended marking of the degree of probability. The continuum of contextualization is used by speakers as an outward marker of the underlying epistemic continuum. As Johnson-Laird (1975) rightly claimed, people have rules of inference or natural decision systems stemming from features they find in the context. It is consistent with a situation in which speakers design their talk for recipients and set contextualization cues in order to trigger such inferencing processes. They provide the context in which the following utterance is to be interpreted with respect to its epistemic stance. Hence, with respect to contextualization, we can affirm that the communicative situation can be fixed onto at least the axes of time, situation, speaker, hearer and communicative necessity (Clark and Gerrig 1990, Vincent and Dubois 1994). This

leads to the conclusion that contexts are not just given but they are emergent, enacted by participants.

The following table shows the results of an ANOVA and a regression analysis in which all factors were entered simultaneously. While the overall result was highly significant with  $F(3,1366)=114.474$ ,  $p<.001$ , an array of post hoc tests can indicate where the significant differences are located.

The last line in Table 3.1 shows that all contextual variables prove to be statistically significant. All factors used in this analysis are important for the contextual characterization of hypotheticality levels.

**Table 3.1: ANOVA Post Hoc table (LSD) for the hypothetical categories**

	time	situation	speaker	hearer	conv.necessity
realis-hypothetical	0.005	0.000	0.000	0.000	0.000
realis - situational	0.000	0.000	0.000	0.000	0.000
hypothetical - situational	0.000	0.000	0.175	0.014	0.838
r <sup>2</sup> change per factor added	0.002	0.003	0.036	0.133	0.304
sig. per factor <sup>4</sup>	0.019	0.000	0.017	0.000	0.000

The  $r^2$  values show how much each factor contributes to the explanation of the variation of the hypotheticality levels. Hence, conversational necessity contributes most to the explanation of the variation of the quotative categories. It accounts for 30.4% of the variance (but note that it contributes little to the difference between situational and hypothetical, cf. line 3). The factor hearer also is selected as a very important contributor with 13.3%. As discussed earlier, it is largely used to differentiate realis from hypothetical quotes. But the factors time, situation and speaker do not have very high  $r^2$  values.

The rows above the  $r^2$  show that all hypotheticality levels are behaving significantly differently in their patterning with all contextual factors (this excludes the aforementioned hypothetical-situational contrast with respect to the factors speaker and conversational requirement). We see this in the line-graph where the plotted values for the category communicative necessity and hearer are pretty much stable between the categories hypothetical and situational.

In a linear regression, a model with the all predictors (speaker, hearer, addressee, situation, time and conversational necessity) was chosen as best fit with  $p < .001$ . Also, the regression analysis yielded that the contextualization factors discussed in this section account altogether for 47.5% of the variance. The rest must be explained by other factors with orthogonal influences. Obviously, not the whole variety of all possible contextual cues for epistemicity can be taken into account in a large scale quantitative analysis. There are certainly many more variables which play a part in the contextualization of quotatives and which can serve as contextual clues for the hearer (and for the researcher) as to how to interpret the quote. Also, in many cases, speakers do not need lexical contextual cues. Rather, we have to assume that “knowledge of how the world works” or background assumptions guide speakers in their interpretation.

An example where world knowledge can be a sufficient indicator to interpret a quote that features many contextual values (which would point in the direction that the quote was actually spoken out aloud) as not having been uttered is when the utterance of this quote would constitute a severe breach of face. An outward utterance would trigger a noteworthy reaction. In the absence of any report of such a reaction, the listener assumes the quote not to have occurred as outwardly realised speech. For example, any gross insults towards persons of higher social standing and in most formal situations would necessarily have reportable consequences. These conversational reactions are important contextual elements which are discourse factors and not quantifiable. Another important factor for any later investigation that seeks to continue this line of research is certainly intonation (cf. research done by Bolinger 1976, 1977, 1989).

I suggest that only a qualitative analysis which aims at identifying the contextual elements in the singular case could do justice to the wealth of inferences that can be drawn from contextual cues. Furthermore, these factors are generally culture-specific. Their interpretation depends on the general cultural context, the narrower interactive context between interlocutors in the here and now of their interaction and on how a communicative situation is evaluated. The choice of what counts as a defining factor at all and how they are consequently weighted and

---

<sup>4</sup> These significances are given for a model that includes all factor variables as predictors.

defined is an emic concept which has to be considered from a more culturally-sensitive perspective such as critical discourse analysis.

But given the aims of this thesis and the above-mentioned limitations, I had to restrict myself to lexical contextualization cues that gave important contextual information, viz. information that is more salient and quantifiable. Notice also the discussion of speech, thought or ambiguous context later on in this chapter.

Hence, in the above analysis, I have classified the epistemic level of a quote first on an intuitive basis [+/- likely to be uttered]. I have then set up a number of contextual elements which occur in the surround of each quotative token. Note that these contextualization cues do not define the categories. Rather, a correlation of the frequency of discorsal contextual elements and the intuitive category hypotheticality level has shown a direct association between contextual rootedness and epistemic level. The post hoc exercise (Table 3.1) has given statistical evidence that the standard contextualization cues for any real world utterance are more or less likely to occur in the three (intuitive) divisions I have initially set up. Hence, the presence or absence of the contextualization cues and conversational necessity are very robust means of differentiating realis, hypotheticality and situational. This result underlines that fact that there is no experimental bias in the analysis. Indeed, my research shows that two separate approaches arrive at the same conclusion: both the intuitive and the data analytic investigation point to the fact that quotes pattern along a continuum. This continuum represents contextual rootedness as well as epistemic level. As will become evident in the discussion to follow, a combination of methodologies, which takes into account discourse analytic and more experimental approaches gives more validity to the results presented in this section.

### **3.4. Comparison between the Quotatives**

Having set up a differentiated set of categories of use and having shown them to be closely correlated to their contextual grounding, I will now compare how individual quotatives are used with respect to those epistemic categories. The table below compares the frequency with which the new quotatives *go* and *like* occur with respect to the three categories of epistemic stance.

**Table 3.2: Correlation between *like* and *go* and degrees of hypotheticality**

	realis		hypothetical		situational		SUM
	%	N	%	N	%	N	
<i>like</i>	27	29	39	41	34	36	106
<i>go</i>	63	46	33	24	4	3	73

$\chi^2(2) = 35.397, p < 0.001$

Table 3.2 shows that *go* is frequently used for higher probability levels, especially for realis quotes with 63%. It is slightly less frequently employed in the hypothetical category than *like* (33% versus 39%). Also, *go* generally does not introduce situational quotes (only 4%, N=3). It does not have an equative function between a quote and a situation. This suggests that *go* needs a real communicative situation, if only a hypothetical one.

*Like*, on the other hand, functions more like a wildcard. It can be used for all probability levels with quite substantial frequency turnouts (27% for realis, 39% for hypothetical, 34% for situational quotes). When using *like*, speakers do not commit themselves to any epistemic stance at all.

By means of comparison, consider the co-occurrence of *say* and *think* and unframed quotes with hypotheticality levels.

**Table 3.3: Distribution of degrees of hypotheticality as a percentage of all quotatives**

	realis	hypothetical	situational
<i>say</i>	70	24	6
<i>go</i>	63	33	4
<i>unframed</i>	38	42	19
<i>like</i>	27	39	34
<i>think</i>	1	83	16

$\chi^2(8): 280.694, p < .001$

As expected, *say* is used most frequently with the realis category (70%). The semantics of *say* at least theoretically pin down the quote as to its realization. *Say* conventionally implies that the quote was actually physically uttered out loud, even though, in real life, it need not have been uttered as *say* can also be used with hypothetical quotes with 24% (but much less with situational quotes, 6% of the time). The next most frequent quotative to be employed with the realis category is *go*, then *like* and unframed quotes, then *think*. As is also shown in Table 3.2, of the new quotatives, *go* is used more for the higher epistemic stances, i.e. for quotes in the

realis category (with 63%). We have a neat range of falling frequencies from *say* through *go*, unframed and *like* through to *think*. As a quotative that encodes mental states, it comes as no surprise to see *think* used least frequently with real occurring quotes (only 1%).

Concerning the hypothetical category, the respective frequencies run in the other direction *think*, *like*/unframed, *go*, *say*. *Think* usually refers to attitudes, opinions, points of view, which are rendered as inner monologue but not spoken out loud. It is the most frequent introductory item for the hypothetical category (83%). *Like*, *go* and unframed quotes are in the middle field. Because they can be used for ‘verbally uncommitted thought’ as well as for real talk, they can function as a hedge. They do not commit the speaker to the actual occurrence of the speech act.

As for situational quotes, we see that *say* and *go* do not frequently occur with quotes of this category. *Think* and unframed quotes can be used for this category. But Table 3.3 shows that the main introductory item for descriptive quotes is *like*.

Hence, all quotatives, including the newcomers *like* and *go*, pattern in their own way with respect to hypotheticality levels. They all have their own functional niche. By association with this particular function, the quotative variants fulfil their specific task, such as the introduction of real occurring quotes for *say* and *go*, the introduction of hypothetical quotes for *think*, unframed and *like* and the introduction of situational quotes for *like* and to a lesser degree also *think* and unframed quotes. *Like* is most indeterminate in this respect. It most frequently introduces hypothetical quotes but can be found with realis and situational quotes. In this functionally flexible patterning it parallels the non-lexical variant, unframed quotes. I assume that one of the reasons why *like* has become so frequent in the lect of many young speakers is its functional versatility: it functions like a wildcard, has no selectional restrictions and can be used with every hypotheticality level. Hence, we can assume that it is *like*’s shallow semantics which make it functionally close to an item which is zero on the surface, and which is by definition not restricted in its functional load by its semantics.

### 3.5. Priming

I will use the term priming in the sense used in the sociolinguistics literature by Cameron (1998), Meyerhoff (2000), Scherre and Naro (1991, 1992), namely making the occurrence of a linguistic item more probable, being aware that there are other understandings of what it means for something to be primed.

Tannen (1987) has shown that speakers are more likely to use a word that has already occurred in a conversation than a completely 'new' one. Cameron (1998) showed priming effects in the quotative frame on a larger scale for Puerto Rican Spanish. This raises the question whether we also find priming effects with respect to quotation strategies in English. The following few paragraphs explore priming effects within the quotative system from a wider and from a narrower perspective. First, on a semantic level, I investigate the co-occurrence of *like*, *go*, *say*, and *think* with lexical indications of contexts of speech and thought. Second, on a structural and lexical level, I explore the correlation between quotative strategies.

#### 3.5.1 Contextual Semantic Priming Effects

Earlier in this chapter, I showed that *go* and *like* and other verbs of quotation occur with quotes of different probability levels. This raises the question of whether they are also associated with different 'surrounds'. The concept 'surround' is defined with reference to the mental space model as described by Fauconnier and Sweester (1996) and is meant to refer to the 'base space' (cf. Chapter 2).

I will now investigate whether the hypotheticality category of the quotes correlates with their surround in terms of speech or thought. The notion of 'epistemicity of the surround' is in accord with the discussion of the hypotheticality continuum presented earlier in this chapter. In this section, I transfer this continuum to the context of the quote and assume that the surrounding utterances can be classified with respect to the probability of outwardly realised speech as well. Speakers can signal the general key of a chunk of speech via linguistic choices. Expressions relating to outwardly realised speech indicate that vocal activity takes place, whereas expressions relating to thought processes are indications of mental,



inner activity. Hence, the context can be classified as pertaining to real occurring speech, as ambiguous with respect to whether speech occurs or not and as pertaining to inner activity.

In order to investigate whether the hypotheticality level of the quote and the mode of the surround are in concord, we first have to determine the key of the context. For this analysis, all instances of quotation were coded with respect to the occurrence of expressions of speech or mental behaviour in their surround. We can then compare which quotatives occur more in a context of speech and which ones occur in a context of mental activity. Obviously, the notion of context or surround is problematic. Other studies have defined it with respect to a certain number of turns or information units. For example, Scherre and Naro (1991:24) operationalise the context to 10 clauses, while Cameron (1998:66) only gives a context of 2 preceding clauses, which he terms a “minimal sequence”. Note that whereas the coding for the contextual variables took the whole preceding narrative context that was available to me (viz. which was on tape) into consideration, for the purposes of exploring priming effects, I imposed a narrower domain. I restrict the scope of the context to 5 TCUs each way (cf. page 28).

I am aware that there is an element of arbitrariness to this measure. But I have chosen to restrict the context in this way for two reasons. Firstly, a large scale quantitative analysis with big token numbers and a delimited time frame has to restrict the frame of analysis of the individual token. And secondly, because of the generally ambiguous nature of discourse features with respect to their scope. It is impossible to determine generally the scope of a discourse item (such as an expression of speech) and to give overall valid delimitation of where it stops. I do not think that we can generally quantify scope (at least not in numbers of words, syllables, lexical items or turns) due to its highly variable and contextually determined nature. Moreover, it may be dependent on intonation as well. It is possible that only a qualitative in-depth case-by-case analysis which aims at delimiting the scope in the individual case and which pays attention to the in-sequence patterning of every single token would provide a full account of the phenomenon.

Hence, for this analysis, I adopted a frame of 5 Turn Constructional Units before (01-05) and after the quote (07-11) as indicated by the brackets in example (3.6).

(3.6) *Tied down by belongings*

01 A: { because I didn't have a bed or a couch or anything,  
 02 I have too much stuff.  
 03 B: { yeah,  
 04 A: { but now I could not move back home,  
 05 and I just *realised* this the other day,  
 06 **I'm going** "wait a minute,  
 07 I can't go home for summer or anything.  
 08 I have too much stuff".  
 09 B: { yeah,  
 10 A: { so it was just a weird transition out of home.  
 11 B: { yeah mmhhh.

Within this context, all linguistic expressions that refer to verbal interaction (such as *to/a phone, to explain, explanation, stories,...*), mental action (such as *to remember, to know, to a dream,...*), or ambiguous ones were coded. 'Ambiguous' included expressions which may index either speech or thought. In this category, we mainly find quotatives *like* and *go*, but also verbs like *reflected* that can signal inner as well as outward monologue. The non-occurrence of any such expressions was marked as 'not'. The 'not' category also includes unframed quotations. Hence the constraint 'surround' can have four values 'speech', 'ambig', 'mental', 'not'. For example, in (3.6) the context contains an expression which pertains to mental activity *realised* in the preceding context (in line 05), the token *go* in line 06 was consequently coded for 'mental'.

I will first examine *like* and *go*'s correlations with the hypotheticality levels of the context and will then compare them with the rest of the quotative pool.

**Table 3.4: Co-occurrence of *like* and *go* with mental/speech contexts (in %)**

	speech	ambig	mental	not
	%	%	%	%
<i>like</i>	21	22	28	29
<i>go</i>	36	21	16	27

$$\chi^2(3) = 7.856, p < .05$$

Table 3.4 shows that *like* occurs with all environments, speech, mental activity, ambiguous ones as well environments that are undetermined with respect to mental and verbal action, coded as ‘not’. The frequencies show that it does do so with roughly equal probabilities (21%, 22%, 28% and 29%). Hence, *like* is uncommitted concerning the epistemicity of its context and can occur in a surround of speech, thought, ambiguous terms or no index of speech or thought. This finding reinforces *like*’s emerging function as an anything-goes item. As was seen in the discussion in 3.4, it acts as a wildcard and can be used with various hypotheticality levels. Table 3.4 now gives evidence that it does not have any preference with respect to the patterns of co-occurrence in its surround. But notice that if there is a preference, *like* tends to occur in the contexts marked for mental activity. This distribution parallels its main functional patterning with hypothetical quotes.

Section 3.4. has provided evidence that *go* most frequently frames real occurring speech. It is associated with the realis category (cf. Table 3.3). The direct association with *go*’s high realis levels (63%) is here mirrored in independent factors: *go* mainly occurs in a context indexed as involving speech. Hence, the correlation of a quotative verb with hypotheticality levels is expressed on a different level as a distributional preference: we notice a clear peak in the correlation between the occurrence of *go* and speech environments (36%). Also, *go* often co-occurs with ambiguous or non-determined contexts, which underlines the fact that *go* can occur in realis contexts as well as in hypothetical contexts. But note its much lower frequency than *like* in strictly mental contexts.

Overall, these findings support the earlier conclusion that *go* is frequently used as a quotative verb for real occurring speech whereas *like* functions more like a wildcard. The convergence of results, that is, the fact that different methods arrive at the same conclusions, strengthens the claim about the functional distribution of *go* and *like* and their association with different hypotheticality levels.

I will now discuss all relevant quotative variants with respect to the epistemic level of their context. If we include the patterning of *say* and *think* with mental and speech environments across varieties we get the following distribution:

**Table 3.5: Co-occurrence of (say) variants with mental/speech contexts (in %)**

	speech	ambig	mental	not
	%	%	%	%
<i>like</i>	21	22	28	29
<i>go</i>	36	21	16	27
<i>say</i>	53	3	38	6
<i>think</i>	21	2	68	9
<i>unframed</i>	44	6	44	6

$\chi^2(12) = 146.400, p < .001$

The first thing to point out is that all quotatives except *like* and *go* have very low levels in ambiguous and ‘not’ contexts. In these two categories, *say* and *think* and unframed quotes have values of less than 10% whereas *like* and *go*’s frequencies are closer to 30%. Such a finding is again consistent with the quotations’ patterning with respect to hypotheticality levels as shown in Table 3.3. *Say* and *think* had much more clear-cut contexts of occurrence and consequently much lower frequency values in others.

While we can certainly not directly map the hypotheticality continuum onto the contextual values, looking at the distributional patterning of the members of the quotative pool gives an astonishingly consistent pattern: the quotative variants are all surrounded by their counterparts. *Say*, prevalently used for framing realis quotes (cf. Table 3.3), most frequently occurs in a context of real occurring speech (53%). *Think*, relating to mental activity (Table 3.3), usually occurs surrounded by mental activity (68%). The newer quotatives *like* and *go*, while less defined in their functional load, also show a more amorphous pattern with respect to their context. Both occur much more frequently with ambiguous or non-defined, ‘not’, contexts. But even in their patterning, we find correlations with the epistemic level with which they have the highest probability of correlating: *Go*, according to Table 3.3 most frequently associated with realis quotes, occurs with the highest statistical probability with speech surrounds. *Like*, already shown to be a wildcard quotative but with a higher frequency for hypothetical quotes than for any other category, patterns exactly this way with respect to its contextual values.

I suggest that this distribution is due to a semantic priming effect in a broad, more abstract sense. This finding underlines the results noted earlier. Table 3.3 demonstrated that all quotative variants have their functional niches on the hypotheticality continuum. Now we see that they are also contextually distributed according to their co-occurrence with speech or thought. The contextual patterning of quotative options confirms earlier results concerning the distribution of quotatives with respect to epistemic stances.

### 3.5.2. Birds of a Feather Effect<sup>5</sup>

The ‘birds of a feather effect’ is based on the colloquial saying “birds of a feather flock together” (Scherre and Naro 1991, 1992) and is a more specific form of the priming discussed in the last section. Scherre and Naro, working on the presence versus absence of plural –s in Brazilian Portuguese NPs, have provided evidence that marking leads to more marking and no marking leads to lack of marking. Applied to the field of quotatives, Cameron (1998:66) has shown for his Puerto Rican Spanish data that there is a ‘birds of a feather effect’ in the sense that quotative frames that contain verbs of direct report trigger more marking with verbs of direct report.


In the next few paragraphs, I am looking into the question of whether the preceding quotative construction or even just a quotative verb has a priming effect on subsequent quotatives. I will show that there are clusters of reciprocal attraction of quotative strategies.

Again, I set my frame of analysis to 5 TCUs. Each quotative strategy was analyzed with respect to the preceding instance of a reported speech strategy within the previous five TCUs. This is explicated in the following excerpt:

---

<sup>5</sup> Note that the statistical analysis in this section is calculated over higher token numbers of *like* and *go*. This is for the following reason: The statistics of earlier sections of this chapter were all calculated with the quotative variants attested in the sociolinguistically balanced subset of US English (cf. Table 2.1). This results in 121 occurrence of *like* and 80 occurrences of *go*. But as I will subdivide these token numbers twice in this section, which would have led to very small numbers per cell, I have decided to include all tokens of *like* (N=238) and *go* (N=186) attested in the US data into this analysis. As the frequency distribution is calculated over a stable denominator (the overall number of tokens per variant), this larger token number of *like* and *go* does not skew results.

(3.7) *Adjusting a computer*

B:  and **they would say** “how did you do that”?  
 and **I'm like** “well it's really easy”.  
 and **they were like** “well I never knew that”,  
 so obviously they'd never considered looking it up you know,

When coding *I'm like* “quote” for its birds of a feather effect, *they would say* “quote” was coded as the preceding strategy because it occurs within the preceding 5 TCUs of the token in question. Conversely, *I'm like* “quote” is coded as the preceding strategy for *they were like* “quote”. Hence, in my coding, in any sequence of quotations, the analysed strategy is the second instance (in this case *I'm like*) and the strategy for which it is coded is the respectively preceding instance (*they would say*).

Following Cameron (1998), I coded first for broad quotative strategy. Here, the alternative quotative strategies are the following: the token in question can be preceded by a quote framed by a quotative verb, it can be preceded by an unframed quotation, or there can be no quotation in the frame of analysis. The results of this analysis are displayed in Table 3.6 below (note again the larger token numbers for *like* and *go*):

**Table 3.6: Frequency of occurrence of quotative strategy per quotative**

	quotative verb		unframed		no quote		SUM
	%	N	%	N	%	N	
<i>like</i>	<b>64</b>	153	3	8	32	77	238
<i>go</i>	<b>66</b>	123	5	10	28	53	186
<i>say</i>	30	165	5	27	<b>66</b>	366	558
<i>think</i>	19	22	4	5	<b>76</b>	87	114

$\chi^2(6): 203.084 p < .001$

Both new quotatives co-occur quite frequently in contexts of repetitively occurring quoted speech with lexical framing verbs: 64% of the quotations before *like*-quotations and 66% before *go*-framed quotations are framed with a quotative verb. Note while neither of the new quotatives often occurs in contexts where no quote has previously occurred, the context ‘no quotes’ is the primary locus of occurrence for the traditional quotative verbs (with 66% and 76%). While we do find the canonical quotatives following any other quotative verb 30% and 19% of the time, the main

distributive slot for the canonical verbs *say* and *think* is where the quotative realm has not yet been opened by previously occurring quotes, either framed or unframed. Also, generally, we note that quotations do not usually co-occur with unframed quotes.

These more abstract, higher level effects suggest that there are priming effects with respect to quotative strategies. The following paragraph explores whether one level down, on the lexical level, there are also patterns of priming. When we split up the category ‘quotative verbs’ from Table 3.6 into lexical items we find the following concrete lexical priming effects:

**Table 3.7: Frequency of occurrence of previous lexical quotative for *like* and *go***

	<i>like</i>	<i>go</i>	<i>say</i>	<i>think</i>
	%	%	%	%
<i>like</i>	37	10	31	22
<i>go</i>	3	39	50	8

$\chi^2(3): 50.944, p < 0.001$

Table 3.7 reports from a subset of the full numbers of token and only takes into account those instances of *like* and *go* that are preceded by a lexical quotative frame within the preceding 5 TCUs. We can see that there are very strong lexical priming effects for *go* and *like*. *Like* follows previous quotative *like* with 37% out of all lexical quotative frames. *Go* directly succeeds *go* in 39% of all cases. Hence, there seems to be much mutual attraction between quotes framed by *go* and *like*. The new quotatives tend to come in clusters. But notice that, in contrast to the reciprocal attraction of *go* and *like* amongst themselves, they seem to strongly disfavour one another. *Like* and *go* do not occur frequently next to each other (with only 3% frequency of occurrence for the sequence *like-go* and 10% for the sequence *go-like*). This shows that the priming effect is lexical and not categorical.

In addition, the frequency for successive *go*-frames is higher than for successive *like*-frames (37% versus 39%) Looking at the contingencies and occurrences of reciprocal attraction that stretch over 3 or more items, we find that *go* globally has a slightly stronger lexical priming effect than *like* for itself. The table below shows the frequency of occurrence of 3 directly subsequent lexically identically framed quotes.

**Table 3.8: Sequential triples of quotative verbs**

Sequence	%	N
<i>go-go-go</i>	11	21 / 186
<i>like-like-like</i>	3	8 / 238

The frequency of 11% for the sequence *go-go-go* (compared to only 3% for the respective *like-chain*) shows that *go* has a stronger tendency to cluster. It seems to prime itself more than *like* does.

Notice also that *like* occurs much less frequently than *go* after *say* (32% versus 50%). As has been shown above, *like* occurs most often after no quoted speech at all. We have already seen that *go*'s primary locus of occurrence is where real occurring quoted speech (as opposed to thought) is already present (Table 3.5), and that *go* is found most frequently in situations where quotative verbs are in the immediate preceding context (Table 3.6). *Go*'s distribution points to the fact that it is to be considered an item which pertains to the sphere of real occurring speech - while it certainly has a very vivid secondary function of introducing hypothetical, inward verbalizations. This has led to the conclusion that *go*, overall, is an item associated with speech. Table 3.7 indeed shows that *say* and *go* frequently co-occur.

There is another point to mention concerning the frequency of the sequence *say-go*: Speakers in my corpus use the alternation of *go* and *say* in order to demarcate speaker roles. This is illustrated in (3.8):

(3.8) *Picking lemons*

- M: and **I said** "hi can I help you"?
- P: ha ha [ha ha ha ]
- M: [you know]?  
and **she goes**,  
and I- you know,  
of course it's this long drive,  
so I - I probably look like ... total hell,
- P: right.
- M: and **she goes** "oh,  
um I was just getting ... some lemons".
- P: ha ha ha [ha ha ha ]
- M: [and **I said**] "oh yeah?  
who are you"?
- P: ha ha [ha ha ]
- M: [and **she goes**] "oh,  
I'm your next door neighbour".



Example (3.8) shows that the roles of 1<sup>st</sup> person and 3<sup>rd</sup> person singular can be differentiated not only by the pronouns, but also by the tense and the verb of quotation. As Romaine and Lange (1991:237) suggested, alternation between quotative frames, while adding to the point-counterpoint nature of conversational exchanges, also helps speaker identification if one quotative is consistently allocated to one speaker. Table 3.9 demonstrates the frequency of the sequences *like*(1s)-*say*(3s)-*like*(1s) and *go*(1s)-*say*(3s)-*go*(1s). Counted in this number were only cases of uninterrupted sequences, which means without any other intermittent quotative strategy such as unframed quotes, or other quotative frames. The frequency count shows that the co-occurrence of *go* and *say* in the pattern of example (3.8), *go* used in the present tense and for 3<sup>rd</sup> person singular and *say* in the past tense used for speaker is quite frequent in my data. This effect does not occur as regularly with *like*.

**Table 3.9: Co-occurrence of 1s *go* and 1s *like* with preceding and following 3s *say***

sequence	%	N
<i>say-go-say</i>	12	22 / 186
<i>say-like-say</i>	1	3 / 238

Here we see that clusters of the sequence *say-go-say* occur with a frequency of 12% in the data. The demarcation of speaker role is a function that *go* has taken up via the alternation with *say*.

In sum, the above paragraphs have provided evidence that priming and birds of a feather effects are quite important with respect to quotative verbs in US English. The effect has been shown to hold on three levels: on a semantic level, with respect to the contextual epistemic stance, on a general structural level and on a lexical level.

### 3.6. The Co-occurrence with Mimetic Performances

While it can be traced back to Plato (Book III of the Republic), the notion of mimesis has been taken up by Goffman (1981), Wierzbicka (1974) in her “quotations as performance” approach, and more recently by Clark and Gerrig (1990). In this approach, quotes are regarded as demonstrations; quoting is “playing someone’s

part”. The enquoting person “does not say what the content of the quote is (i.e. what was said), instead he does something that enables the hearer to SEE for himself what it is, that is to say, in a way, he *shows* this content” (Clark and Gerrig 1990:802). The literature lists several reasons for the incorporation of mimetic performances: (i) to convey a more emotion-based rather than factual rendering in order to reveal how the speakers felt in and perceived the situation (ii) to add more vividness, which is supposed to lead to audience involvement (Blyth et al. 1991) (iii) and to superimpose internal evaluation without having to step outside the quotation frame (Labov 1972c).

Mimesis is understood as direct representation, the total imitation of the event. In contrast, diegesis is summarized representation, a mere synthesis of the original event.<sup>6</sup> For our purposes, may it suffice to say that the difference between mimesis and diegesis is between showing and describing, dramatic and descriptive, between reporting the ‘how’ and the ‘what’ of the original speech event.

But even though the claim that these modes of representation are to be fundamentally kept apart holds in theory, in every day talk-in-interaction the boundaries between them are fluid and creatively exploited by speakers. Because direct speech can incorporate “delivery aspects” (Clark and Gerrig 1990), such as voice effects, gestures, inarticulate sounds etc., or even consist entirely of them, pure direct reported discourse is a hybrid means of rendering past speech events. The two modes of quoting can thus be considered as two endpoints on a continuum (Güldemann 2001, Yule 1993:236).<sup>7</sup>

I will now outline my coding conventions: as quotes hardly ever occur in a pure form as sounds only or speech only, a choice had to be made as to how to delimit the category voice or sound and no voice or sound against one another.

---

<sup>6</sup> These extremes are claimed to exist in their purest form in direct and indirect discourse, respectively. But Klewitz and Couper-Kuhlen (1999) show that the borderline between direct and indirect discourse is fuzzy even with regard to this point. They find that the age-old claim that mimesis defines direct quotes can be violated. Mimesis can and does occur both in indirect and direct quotes. In this study, I am only concerned with direct quotes.

<sup>7</sup> As an anonymous reviewer of Buchstaller (2002b) commented, if mimesis is scalar, it should be possible to quantify a quote’s mimetic value, or at least to position a quote with respect to its placement on such a scale. Quotes with only voice effect should then count as less mimetic than quotes with voice and gestural effect. While this is an important fact to mention and does deserve investigation, I have not included any measurement of a mimesis coefficient in this study. The main point here is that mimesis is a scalar, multi-value phenomenon and that its mere non-/occurrence does not constitute an *a priori* criterion for categorizing a quote as direct or indirect.

Cameron (1998:62) codes quotes as [+mimesis] when they only consist of mimetic enactment. I find that this gives undue supremacy to verbal matter and drastically reduces the frequency of mimetic quotes. In practice, while we do find quotes which consist of only sound effect, the bulk of my corpus consists of quoted speech interspersed with sound tokens and/or stretches of voice effect. Hence, in my data, a stretch of quoted report is coded as 'sound' if it contains any token of non-lexical sound effect. It is coded as voice if it contains any stretch of speech that is impressionistically a different voice from the reporting speaker's one. The remaining quotations, which do not feature any voice or sound effects speech, count as 'speech'.

The counting of gesture is more complicated (see Cameron 1998). As my corpus consists of audio recordings, I could only rely on the following cues to detect gesture: when bodily movement caused noise of some sort or when the context clearly indicated the presence of mimesis. These contextual cues could be twofold: either the interlocutors commented on the mimetic re-enactment or an auditorily empty quote triggered some sort of response from the interlocutor, such as laughter. We would then have to assume an extralinguistic, inaudible content of the quote. Obviously, many gestures will still be undetected. Hence, overall, my corpus only contains eight instances of gesture. The impoverished information obtained through audio-recording is one of the many facts which speaks for the use of video cameras in research on naturally occurring speech.

For the present investigation, I have lumped the three mimetic categories, voice effects, sounds and gesture together in the superordinate category 'mimesis'. Coded as mimesis were all aspects of mimetic enactment to be revealed on an auditory or contextual basis. This includes voice and sound effects of all sorts, and gestures, where they could be retrieved from audience reactions.

I will now discuss the co-occurrence of quotative verbs with mimetic enactment. The following table gives an overview of the co-occurrence of mimesis with the most important verbs of quotation in US American English.

**Table 3.10: Co-occurrence of quotatives with mimesis**

	%	N
<i>go</i>	73	58
<i>like</i>	75	91
<i>say</i>	60	337
<i>think</i>	75	91
<i>unframed</i>	59	176

$\chi^2(4): 47.373, p < 0.001$

Table 3.10 suggests firstly that many quotes contain mimetic re-enactment, no matter which introducer they have. But there is still a significant difference ( $p < .001$ ) between the occurrences of mimesis with quotative verbs. Amongst the pool of quotatives, *go* and *like* (76% and 73%) are most frequently used to frame mimetic enactment, a finding that is consistent with much of the previous literature (Blyth et al 1990, Butters 1980, Cukor-Avila 2002, Singler 2001, Schourup 1982a, Tagliamonte and Hudson 1999, Tannen 1986, Yule and Mathis 1992). In contrast, only 59% of the tokens of the most frequent dialogue introducer *say* as well as of unframed quotes co-occur with mimetic performances.

Note that with 75% occurrence with mimesis, the quotative *think* patterns with the new quotatives. This is not very surprising given that *think* enquotes inner monologue, opinion, attitude and point-of-view. Table 3.3 gave evidence that *think*-framed quotes are often inward, not uttered out aloud, not interactively realized. Also, it shows that *like* and *go* often frame hypothetical quotes. The common patterning that the new quotatives *like* and *go* have with *think* is the cue to their common behaviour with mimesis. Hypothetical speech such as evaluation, attitudes, etc. is often high in emotion (cf. Chafe 1994) and very frequently embodies voice or sound effects. Because mimetic enactment has concentrated semantic reference, speakers can represent inner mental activity in a more expressive form than only speech. According to Goffmann (1981), these categories should be expressed via “response cries”. Such quotations leave the question of the original speech event’s production entirely open.

But in contrast with *think*, *like* and *go* can also be used with realis quotes (cf. Table 3.3). My claim is that speakers using *go* and *like* play with this indeterminacy between speech and thought. They exploit the fact that the new quotatives operate in

the grey area between real occurring and hypothetical speech. Using the new quotatives, speakers quote as if they were reproducing a real speech act but package it in a more expressive form, in sound and voice effects. This suggests that speakers take advantage of the full creative possibilities the language offers them in the new quotatives: a stream of consciousness-like display of inner states and attitudes realized in vivid, immediate speech. *Like* and *go* have introduced this quotative style into the spoken language. It now fills a space within the spectrum of poetic formulae of the spoken register, where indirect free speech, commonly used in writing, is not an option (Chafe 1994, Romaine and Lange 1991) and where the theatrical topos of soliloquy did not catch on (Ferrara and Bell 1995). Whereas *think* theatricalises inner speech by outwardly displaying it in vivid, emotionally heightened output, *like* and *go* are indeterminate with respect to whether the quote is hypothetical or real.

Table 3.11 shows that the newly grammaticalised quotatives *go* and *like* are distinguished from the old quotatives *say* and unframed by their function as mimesis markers. But they have a non-significantly different pattern to *think*.

**Table 3.11: ANOVA post hoc (Bonferroni) results for significances patterning with mimesis**

	say	think	unframed	like	go
<i>say</i>	x	<b>.000</b>	.937	<b>.000</b>	<b>.002</b>
<i>think</i>	<b>.000</b>	x	<b>.000</b>	.567	.366
<i>like</i>	<b>.000</b>	.567	<b>.000</b>	x	.628
<i>go</i>	<b>.002</b>	.366	<b>.004</b>	.628	x

We see that the difference between *like* and *go* is not statistically significant, both can be used to enquote sounds. In the spirit of Gldemann (2001) and Yule and Mathis (1992), we can claim that in US English, the canonical quotative *say* (and its non-canonical, non-lexical variant unframed) foreground the semantics, the propositional content of the quote. *Like* and *go*, the newcomers in the quotative complex and still more marked constructions, highlight the ‘how’, the demonstrative-enacted side of the material. This is in line with *think*, the other quotative used for the hypothetical category. Ferrara and Bell’s (1995:282) claim that *like* is a “theatrical, highly conventionalized utterance” fits well with the above finding that it frames the emotionally charged content of a sound effect or theatrical display of an effect quote.

### 3.7. Conclusion

This chapter presented an investigation of the functional patterning of the quotative system in US English. It has demonstrated how *like* and *go* are used synchronically as quotative items in US English in comparison with the older quotative variants. I have shown that the fact that *like* and *go* convey linguistic information is best represented and understood by recourse to the notion of a hypotheticality continuum. We need to be able to understand how quotes can be used with different epistemic stances and how quotative items function with respect to probability levels. This is the basis we need to show how speakers index how they feel towards the quote, how they want to re-enact and represent the enquoted material, and how they index their commitment to the epistemic stance of the quote. Hence, I have argued that while both new quotatives are heavily used for the expression of hypothetical talk, *go* has a stronger affiliation with higher probability levels. *Like* can be considered a wildcard quotative as it does not have any selectional restrictions with regard to the quote it frames.

Note Johnstone's (1987: 33) claim that "quantitative analyses of discourse must be supplemented with qualitative microanalyses of what individual speakers do in particular situations" and that we have to "show how rhetorical microanalyses of some of the data can explain aspects (of tense choice) which quantitative analysis leaves unexplained" (1987:35). While a quantitative investigation of the quotative system would exceed the scope of this thesis, I am exploring these issues in on-going work. I intend to push further our understanding of the association of quotative variants, hypotheticality levels and items which can serve as contextualization cues in the surround. Contextual elements that have been linked with the marking of epistemicity are, amongst others (for a more comprehensive list see Thumm 2000 and in preparation), *would/was gonna*, negative polarity items, modal adverbs such as *perhaps/maybe* (Buchstaller 1999) and questions or question intonation (Bolinger 1989). Furthermore, Snitzer-Reilley (1986) has shown that, as hypotheticality is in closely associated with beliefs and mental action, verbs of imagination can be used to mark epistemicity as well as verbs of volition, enablement and deontic stance. Finally, unspecified number (*any*), gnomic statements, and the position of a quote in

a row are contenders for features that have an impact on the interpretation of the epistemic stance of a quote.

The distributional effects that I have discussed in detail in this chapter show a clear division of work amongst the (say) variants. This is further underlined by the quotatives' patterning with respect to the speech and thought mode of their environment. Again, *like* patterns as an 'anything goes' item whereas *go* is more affiliated with its speech surrounds. Overall, again, we witness a task-sharing within the quotative pool.

Looking at priming and 'birds of a feather' effects, it has been shown that there are strong priming effects both with respect to the general domain of the quote (realis versus irrealis), as well as concrete lexical priming effects. Also, speakers use the alternation between *say* and *go* to demarcate speaker roles.

With respect to the enquoting of sounds and other mimetic performances, my findings underline the use of the new quotatives as mimesis markers. In this use they pattern in line with the quotative for mental states and actions, *think*, and in contrast with the older speech-related quotative, *say*.

The overall picture that emerges from this investigation into the intralinguistic patterning of all quotative variants is that the quotative pool has reacted to the intrusion of two incoming items. The functional load is neatly distributed between the members of the quotative cohort. A functional reallocation must have taken place within the system with older members giving way to the newcomers and then settling in new niches. The fact that the new quotatives pattern quite independently of the rest of the pool of quotative devices underlines their functional load. Hence, *like* and *go* are not vacuous, taken over for purely social and stylistic reasons and because they are cool, as has been suggested in much of the variationist literature to date. The enlargement within the pool of quotative constructions is not simply that of two intrusive pleonastic items edging their way into a stable paradigm. But, the new quotatives have a justified place amongst quotative devices. They have taken on quite novel functions with respect to mimetic enactments, the marking of epistemicity, speaker role demarcation etc. This is an important finding, as it shows the division of labour amongst quotative devices. Hence, this chapter demonstrates

that go and *like* have functional significance. In Chapter 6 we will see that they also have social significance.



Each and every step [...] is an innovation, not only the initial act, through which a new linguistic entity comes into being. It is through innumerable individual acts of innovation – of acceptance, adoption, and acquisition – that any new entity gains currency and enters into competition with traditional entities in the usage of a linguistic community. (Andersen 1989:14)

## Chapter 4

### The Grammaticalisation of *like* and *go*

This chapter is a synchronic account of the grammaticalisation of *like* and *go*, two items which behave exactly as pointed out by Güldemann (2001): they are functionally versatile elements which build up a whole network of uses, they are notoriously polyfunctional outside the quotative domain and their utterance-introduction meaning only surfaces in their quotative function. The classic definition of grammaticalisation is Kuryłowicz's (1965 [1975]:52) statement that it is a phenomenon whereby we witness "...the increase of the range of a morpheme advancing from lexical to a grammatical or from a less grammatical to a more grammatical status".<sup>1</sup> But amongst the scholars working in this field, there is disagreement about whether grammaticalisation is to be studied "from the point of view of patterns of language use across time or at a synchronically segmented moment in time" (Traugott and Heine 1991:1, cf. also Hopper and Traugott 1993); whether grammaticalisation is the diachronic evolving of grammatical structure or whether it is rather to be considered a discourse pragmatic fact. In this chapter, I will take grammaticalisation very broadly to mean a phenomenon whereby discourse patterns lead to the expansion of functions. More explicitly, this is the case when new, more grammatical functions develop out of already existing material.

The new quotatives, *like* and *go* have quite a number of functions outside the quotative frame. To date, all the models that have been proposed to explain their polyfunctionality rely on the postulate of unidirectionality in grammaticalisation

---

<sup>1</sup> For a review of this and newer definitions of grammaticalisation. cf. Traugott (2001).

(Bybee, Pagliuca and Perkins 1991, Bybee, Perkins, and Pagliuca 1994, Fleischman and Yaguello to appear, Heine, Claudi and Hünemeyer 1991, Meehan 1991, Romaine and Lange 1991, Traugott and Heine 1993 and many others). In a large overview of quotative structures in African languages, Güldemann (2001:2) points out that the approaches to the grammaticalisation of quotative items are very homogeneous. This manifests itself in the strong tendency previous studies have displayed to derive the functional versatility of any quotative item from its previous use as a speech verb, which is invariably seen as the diachronic starting point for their development. The traditionally assumed speech-verb channel (Lord 1976, 1993, Saxena 1995) postulates a unidirectional chain linking generic speech verb > quote orienter > quotative > complementiser > other, where the *de dicto* function is considered a precondition for the development of the other functions.

At this point, a brief clarification of terminology seems in order: a **channel** (Givon 1979, Heine and Reh 1984, Lehmann 1982,) refers to a particular way or path of grammaticalisation. Channels constitute alternative options which have been shown cross-linguistically to be the lines along which the same source concepts typically develop on their way towards the introduction of a new grammatical category. They are usually defined with reference to their endpoint.<sup>2</sup> **Chains**, on the other hand, concern the internal structure of the channels. They imply a development wherewith one function leads to the other in a chain-like fashion (see Figure 4.1 below).

Frajzyngier (1996:99) claims that “one of the most common sources for the *de dicto* complementisers are verbs of saying, confirmed time and time again in many languages of the world”. But Güldemann (2001) has shown the inadequacy of the speech-verb channel, which has long been considered the default scenario for the creation of quotative introductory items. Giving a rich array of data, Güldemann

---

<sup>2</sup> It is true that knowledge of what constitutes a natural development - based on what we know and what has been established cross-linguistically - gives grammaticalisation mildly predictive value. Nevertheless, I would argue that viewing it as necessarily goal-directed with the endpoint as the crucial element diverts attention from the centre of investigation: the developmental process. This is because presupposing an endpoint has teleological implications, which can cloud one's view of the linguistic reality. By looking at developments from the perspective of a postulated endpoint, we might not notice developments which take different paths from the same source concept or developments which have not (yet) arrived at the goal. It seems to me that it makes more sense to conceptualise channels in a non-telic manner, as development from source concepts onwards (Hopper and Traugott 2003). See in this vein Bybee, Perkins and Pagliuca's (1994:282) inquiry into the processes underlying creative language use.

argues that unidirectional postulates often blur our understanding of grammaticalisation processes and lead to self-fulfilling prophecies in the understanding and analysis of the data along pre-postulated channels. He shows that the synchronic polyfunctionality which quotative items display cross-linguistically is largely determined by ultimate origins other than the quotative frame.

Having shown that there is no single unproblematic case in the grammaticalisation of quotatives in African languages, he voices strong reservations regarding overgeneralizing about the speech-verb channel and calls for more rigour in reconstructing the facts (see also Naden 1989:165 and, in a cross-linguistic study, Ransom 1988:364). He issues a call for a thorough cross-linguistic investigation of quotative introductory items and their functional networks.

Sweetser (1991) claims that we cannot rigidly separate synchronic from diachronic analysis because a synchronic slice is, by definition, nothing but a stage in the history of a development (cf. also Heine, Claudi and Hünemeyer 1991:172, Lehmann 1982/1995:26, Schiffrin 1987).<sup>3</sup> Language structure does not provide a static semantic organization and conceptual structure is changing continuously. Therefore, a synchronic functional account cannot be fully disengaged from diachrony because the meaning of an item is in constant flux. Any description at any given time is only a synchronic snap-shot of the developments so far and as such simultaneously the starting point for future ones. Sweetser argues for their combination in a unified account. Hence, the proposed radial structure has a temporal element to it in that it captures what *go* and *like* have boiled down to at the time of data collection. Heine et al. interpret grammaticalisation chains as the “frozen” result of conceptual manipulation (1991:99). Frozen seems to me to imply a certain element of immobility, so I prefer to refer to the situation as being one the linguistic system has boiled down to, because this seems to capture the fact that it is a synchronic snapshot of an ever-developing mechanism. Diachronic data informs our knowledge of how linear or non-linear processes led to the present scenario. It also gives information about the sequential order of events. I argue that as a linguist,

---

it is only on the basis of such knowledge that we can make informed statements about the present. Hence, the proposed radial structure and especially the core meaning are built on the diachronic information available from historical sources such as the OED. But to language users, as long as all tokens in question are equally alive in their linguistic system, any diachronic information beyond the point of their personal recollection is irrelevant. It does not make any difference for speakers whether e.g. the temporal use of *go* was an earlier function than its modal function. What matters is the synchronic link that they perceive between the synchronically existing functions.

I will, therefore, allude to the diachronic grammaticalisation history which informs or motivates *like* and *go*'s "core" meaning and its functional extensions, but this will not be the centre of attention here (for diachronic accounts of their development consider Meehan 1991 and Romaine and Lange 1991 for *like*, Bybee and Pagliuca 1987, Bybee, Perkins and Pagliuca 1994, Pérez 1990, Royster and Steadman 1923/1968, for *go*). In what follows, I will explicate the functional relations between the uses of *go* and *like* and will show that they are semantically-pragmatically motivated. Basing my comments on the findings of Lakoff (1987), I will show how the synchronic functions that *like* and *go* have assumed can be conceptualised via a radial structure model, a variation model.

The synchronic uses of *go* and *like* can be conceptualised as an interrelated network of functional extensions stemming from a common core meaning. The underlying processes that triggered the transfers will be discussed as will the common semantic core that I argue still pertains in each of the extended functions. I will show that, at least with respect to the *go* and *like* scenario, a radial structure model can show how the superficially messy facts are linked in an orderly way.

I will concentrate on motivating their synchronically interconnected functional extensions in semantic-pragmatic space rather than through time. Hence, this chapter is a synchronic attempt to describe the grammar of *go* and *like* as it is realised by speakers in present day English. I will show that synchronically, a unidirectional model, which would propose a linear development along a cline of

---

<sup>3</sup> Variationist sociolinguistics also incorporates this insight: the idea that the synchronic study of variation at one point in time can give evidence about larger, diachronically occurring processes, is known as the "apparent time" construct.

grammaticalisation, would not hold much explanatory power in cases where multiply ambiguous and overlapping meanings co-exist synchronically. I will come back to this point later on in this chapter.

#### 4.1. Conceptualising grammaticalisation

Lichtenberk (1991) asks the following question: do meanings have structure or are they unanalysed wholes? As, by definition, all the meanings of a polysemous word are related, it has been argued that they can be considered as linked through a complex network of partially shared commonalities (Lichtenberk 1991), family resemblances (Jackendoff 1983, Wittgenstein 1978), or gradual relatedness (Lakoff 1987). In fact, Brugman (1988) and Lakoff (1987) have suggested that word meaning is structured and unified. This view implies that meanings are not unanalysed wholes but that they are motivated and explicable. In the same vein, Sweetser (1991:9) claims that words do not randomly acquire new senses but that derived functions are acquired by cognitive re-structuring and that consequently, the multiple synchronic functions of any given word are interrelated in a motivated fashion. Hence, we should be able to discover the systematic structure underlying such a semantic field and it ought not come as a surprise to find that certain semantic changes are frequently cross-linguistically attested. Later on in this section, I will show that there is evidence for a universal cognitive / perceptually salient pragmatic-semantic field lying behind *like*'s attested functions. However, the same is not true of *go*. Neither Güldemann's (2002) extensive study of African languages nor the findings reported in the present study (cf. page 131) on typologically unrelated languages has found any compelling evidence for an underlying field. It remains to be fully investigated whether *go*'s functional extension to quotation is a largely idiosyncratic phenomenon.

Sweetser (1991) and Lakoff and Johnson (1980) have proposed that linguistic usage reflects the inherently metaphorical understanding of many basic areas of our lives. The extension of a linguistic item is motivated by the relation that speakers perceive between the old and the new use of the item. Given this explanation,

linguistic structure can be considered a mirror image of cognitive structure. Heine, Claudi and Hünemeyer (1991: 168) claim that there is iconicity between cognitive and linguistic patterning in the sense that language is the reflection of conceptual manipulation. Consequently, language, as the iconic representation of human cognition, is no more and no less logical or objective than any other aspect of human cognition. The functions of a linguistic item are sedimentations of the linguistic modeling of the cognitive processes involved in grasping connections and contingencies in the real world (cf. also Blank 1999). Overlapping, ambiguity, and polyfunctionality are a function of the underlying cognitive processes of metaphor, context induced interpretation, metonymy etc. and are to be considered the outcome of human agents' creative language use rather than autonomous processes.

When concepts are manipulated, they are subject to contextually induced changes. Due to the inferencing, sense-making mechanism of language users, these reinterpretations come about via the infusion of meaning from their local occurrence within the discourse. Real-time production pressure forces speakers to comply with Horn's principle of least effort "say no more than you must" (1984)<sup>4</sup> and requires the hearer to extract all possible meanings from the message, including all implications as long as they are not controversial. Those inferences are abductive, based on world knowledge (of which linguistic knowledge is a part) and lead to an "enriched interpretation" (Blakemore 1987) or "context-induced reinterpretation" (Sperber and Wilson 1986:1). This second meaning, which first arises in a specific context, can gradually become conventionalized (Heine, Claudi and Hünemeyer 1991). Also, when an implicature commonly arises out of the meaning of a form, it can become part of its meaning (Heine, Claudi, and Hünemeyer 1991:150/1, Traugott 1989, Traugott and König 1991) and can even go so far as replacing the original meaning. The common underlying idea of most grammaticalisation accounts is that, as a linguistic item loses more and more of its original inherent meaning, it is more susceptible to changes brought about by the syntactic-semantic context in which it occurs.

---

<sup>4</sup> See in this vein Grice's (1975:45-6) Second Maxim of Quantity "Do not make your contribution more informative than is required".

## 4.2. The role of ambiguity

A diachronic chain model, which is employed in most accounts of multi-meaning entities, is able to show overlappings and ambiguities between the diachronically adjacent stages within the channel. But such a model fails to account for the fact that meanings overlap and even reinforce one another at different ends of the “channel”. To get around this problem, Craig (1991) and Bybee, Perkins, and Pagliuca (1994) allow for multiple grammaticalised constructions developing from the same source and claim that they are often at different stages along a grammaticalisation path. This amends the problem inherent in a purely sequential step by step model. But it still fails to account for the interrelation of meanings amongst themselves. Craig (1991:486) puts it this way, it “raises more questions than it answers about the nature of grammaticalisation chains and how to conceive of them in detail”. As we will see, in cases like *go* and *like*, where an analysis of changes resulting from grammaticalisation reveals a complex network of ambiguous and even overlapping meanings, it seems inappropriate to try to force the multiplex development into a linear sequential model.

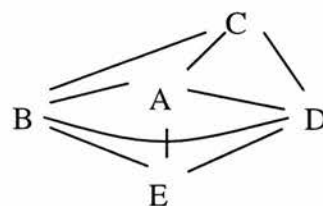
This discussion agrees with the traditional accounts of grammaticalisation that during a transition from A to B there is a stage (A,B), where the preceding and the succeeding entities exist side by side, causing ambiguity and variation. It will be argued that this stage is not necessarily ephemeral (Hopper and Traugott 2003). Heine, Claudi and Hünnemeyer (1991) rightly claim that the transition in meaning is not a sudden replacement but a stage where the former co-exists with the latter. A semantic field model considers this overlap, the functional in-between-ness being an important factor. Ambiguity between two or more functions is not necessarily an indication of a flux moment of language change – it can stabilize as a radial structure model. The difference between a unidimensional and a multidimensional model is shown in the following diagrams.

**Figure 4.1: Different conceptualisations of the grammaticalisation process**

Uni-dimensional chain model  
Heine, Claudi and Hünemeyer (1991)

A → A,B → B

Radial structure model  
(Lakoff 1987)



A unidirectional model implies that the next steps in the chain are  $\rightarrow B, C \rightarrow C$ . A radial structure model, however, does not constrain the sequential development to a linear order. Developmental sequences such as A, C, D pose a particular challenge to chain models in general because they cannot be explained by a unidirectional model. A radial structure model, while in no way incompatible with developments that progress linearly, does not restrict them to unidirectional clines. It is capable of conceptualizing the functional developments of all sequential orders as they occur during language use. Because a radial structure model is completely a-teleological and does not make any assumptions about the temporal order of events but only depicts functions as related in their semantic-pragmatic field, it is flexible enough to cope with multidimensional developments such as *like* and *go*.<sup>5</sup>

Also, a radial structure model such as the above is especially useful when we look at cases of polygrammaticalisation, where one single form develops different grammatical functions in different constructions, a complex form of split where “various chains of grammaticalization interlock” (Craig 1991:486). Such processes are not just a matter of the original meaning remaining and developing in a step-like fashion where one extension informs the other. Rather, synchronically, items such as *like* or *go* are conglomerations of highly ambiguous and context-dependent meanings which are interrelated amongst one another. Lakoff (1987), Meyerhoff and Niedzielski (1995, 1998), Mosegaard Hansen (1997, 1998) and Buchstaller (2001a,b) argue that cases like the above are best explained by a common core

---

<sup>5</sup> As one anonymous reviewer commented, such synchronic grammaticalisation accounts can only be understood as ways of organizing the data, NOT as actual paths of change. While this is certainly true, as is underlined by my use of the term ‘model’, I argue that a radial structure model still has the advantage of graphically displaying the complexity and multidimensionality of the functional extensions, a fact that a linear model can capture but fails to depict.



model. In such a model, all functions of an item centre around a semantic-pragmatic core. They retain a strong link to this core and have overlapping functions, which are more or less closely tied to each other.

The claim that the synchronically occurring functions of *go* and *like* are structured as a field around one focal member, the semantic-pragmatic origo, is backed by work on the structure of concepts, which has shown us that cognitive structures can be explained in terms of family resemblances (Wittgenstein 1978) and prototypes (Berlin 1992 and Rosch 1975, 1978). The manipulation of such concepts can be understood as creative with one meaning giving rise to one or more others. Semantic and pragmatic shifts do not take place between discrete categories but involve a continuum without any clear-cut boundaries (cf. Bath-Weingarten and Couper-Kuhlen 2002). Overlapping, ambiguous meanings are not coincidental. Rather, they are part and parcel of an associative, creative language use that leads to the extension of meaning. Consequently, grammaticalisation processes are gradual and continuous rather than discrete. They result in complex networks of associations with nodes which are interrelated amongst themselves via cognitive links. A semantic field provides the flexibility to depict such gradualness because it encompasses various partially overlapping semantic dimensions. Functions can be conceptualised as clines with more or less prototypical instances aligned on them. This has the advantage that we do not have to postulate any rigid category differences into which we have to force the occurrences of the lexeme when we find it in the data. Rather, individual instances can be conceptualised as positioned simultaneously on several clines because a polydimensional field permits multiple ambiguities and not just two consecutive ones as a unilinear cline does.

Also, as concepts are in flux and creatively used, the structures resulting in the linguistic system can be, and often are, just as complex. A core structure model has the advantage that it can explain already existing meaning-transfers while being flexible enough to provide for more motivated extensions *in every direction*. It embodies the history of the cognitive-semantic developments to date which have boiled down to meaning-transfers in the linguistic system but every single one of its dimensions can be enriched by associative language use. Hence, a radial structure

encompasses the polydimensional semantic-pragmatic field of a linguistic item without constraining its development in the way a unidimensional chain does.

I propose a radial structure that, I will argue, underlies the many senses of *go* and *like*. This structure highlights the multifunctionality and the overlap of the various functions associated with *like* and *go* that are evidenced in my data. I will explain the transfers in the linguistic system where newer functions have arisen out of the original linguistic item and are considered by speakers to be conceptually connected. I will start with *like*, then discuss *go*.

### 4.3. The *Like*-Scenario

*Like*'s repetitive occurrence and the reciprocal attraction in its multiple functions have been interpreted as a sign of grammaticalisation underway (Dailey-O'Cain 2000, Romaine and Lange 1991, Singler 2001, Tagliamonte and Hudson 1999). Furthermore, in its non-standard uses as a discourse marker<sup>6</sup> and as a quote introductory item, *like* has commonly been associated with the colloquial, every-day discourse of adolescents and young adults, a distribution which points to its evaluation of an item with newly grammaticalised functions. (As we will see in Chapter 6, there is some truth to these stereotypes, but my focus here will remain on the semantic-pragmatic dimension of its use).

The grammaticalisation of *like* in all its functions, as a preposition, a conjunction, a discourse marker, and a quotative (cf. Webster's New World Dictionary 1994:783, The American Heritage College Dictionary 1993:786, The Random House Webster 1999:768) is yet to be fully resolved. Ever since Romaine

---

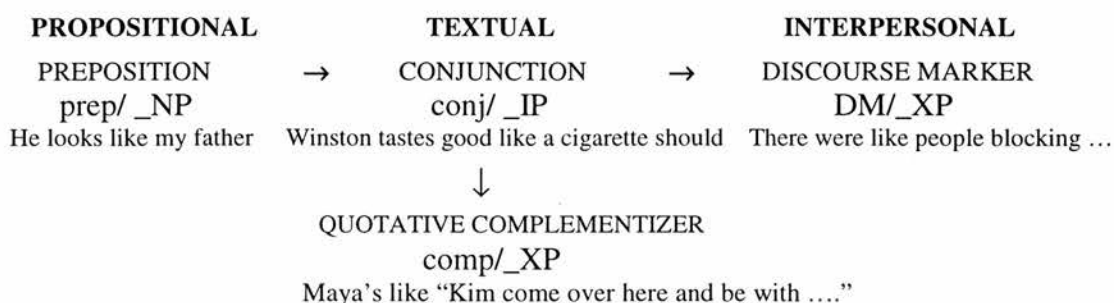
<sup>6</sup> I consider discourse markers to be items that have much syntactic freedom and possess great mobility inside the utterance in the sense that they can precede or follow a clause or any phrase. Following Schiffrin (1987), discourse markers are "sequentially dependent elements, which bracket units of talk and which are *independent of sentential structure*" [emphasis mine]. They have been commonly classified as particles drawn from a heterogeneous group of functional classes that are stylistically stigmatized, short, unstressed, optional items that do not affect the truth value of a sentence and do not contribute to propositional content. Their occurrence is a typical feature of oral style where they have to be interpreted on a global level as they have textual and interpersonal function, see Kroon (1995), Schiffrin (1987). Cf. also Sankoff et al. (1997:195) "[discourse markers] do not enter into constructions syntactically with other elements of the sentence".

and Lange started off the investigation into *like*'s grammaticalisation with their brilliant and most insightful 1991 article, follow-up corpus-based studies featuring a grammaticalisation model which is able to link all of *like*'s occurrences in the data have been notably absent. The literature usually recites Romaine and Lange's (1991) model and keeps reporting on *like*'s multifunctionality without suggesting how to incorporate its manifest ambiguity and fuzziness into a coherent account (Dailey-O'Cain 2000, Macaulay 2001, Tagliamonte and Hudson 1999). Sociolinguistic quantitative analyses, when having to incorporate cases of overlap and "deviant patterns", speak vaguely about imperfect introduction from other varieties (Macaulay 2001:13, 16), or small numbers of tokens (Tagliamonte and Hudson 1999). These tokens are then set aside from further investigation or incorporation into a comprehensive account. Singler (2001:274) mentions several cases of what he calls "transitional forms" but does not attempt to incorporate them into a coherent model. In short, so far, no model has been able to capture the whole picture of *like*. What we need is a) an account of *like*'s grammaticalisation that is robust enough to cover its entire distribution and b) a conceptual structure that is able to account for the inter-relatedness and ambiguity of all grammaticalised uses of *like* (such as the hedging and the quotative function) *amongst* themselves.

In this chapter, I will show that a corpus-based analysis of *like* reveals that the development of its quotative function is actually not as smooth and simple as a chain or as multiple extensions from one original core. I will first trace the links between *like*'s use as a quotative and its older grammaticalised functions. This discussion will bring out the highly interrelated nature of those functional extensions. Then I will sketch the situation as it presented itself in the two corpora I was using (discussed in Chapter 2). I will also give several examples from the web. My claim is that when we take into account **all** instances of *like*, it becomes clear that while the canonical grammaticalisation accounts certainly have some explanatory power, there is more to *like*'s patterning than is captured in those models. I will show that a polydimensional field with multiple paths can fully explain *like*'s patterning in all its functions as we know it today.

I will now briefly discuss the most influential of all models concerning the grammaticalisation of *like*. Romaine and Lange (1991) propose a grammaticalisation chain for *like*, as depicted in Figure 4.2, which is based on Traugott's (1982) model. They account for the co-occurrence of *like*'s uses and for the fact that its development is not strictly sequential by postulating a branching model. (For other attempts to conceptualise *like*'s current uses, cf. Meehan 1991 and Fleischman and Yaguello to appear, who restructure the Romaine and Lange (1991) model to a multiple pathway model).

**Figure 4.2: Romaine and Lange's (1991) Model**



Romaine and Lange's model shows that when *like* precedes a noun phrase, it is used as a preposition and when it takes a sentence as a complement, its function is that of a conjunction. Note that, in accordance with its cognate suffix (-ly), it can also follow whole chunks of discourse. As *like* has a considerable amount of syntactic freedom and variable scope, it is at the point where it can be reinterpreted as a discourse marker, which is syntactically entirely free. Corpus-based analyses have given statistical evidence that the placement of *like* as a discourse marker is usually before the linguistic material it qualifies in US English and after it in British English (cf. Miller and Weinert 1995). If the material following *like* is a direct quote, *like* can be interpreted as serving the function of an introductory item to that quote. The question of the directionality of this associative process (viz. of cause and effect) is one which is hard if not impossible to answer. Suppose an item which is en route to grammaticalisation occurs in a "new" construction, e.g. quotation. It can metonymically be re-interpreted by language users as being constitutive of this context, that is it is interpreted as a quotative. It can then be used productively with this new function. This, in turn, strengthens its association with that context. In its

turn, this will lead to a higher frequency of use, etc. Hence, the process, as postulated here, is both reflexive and constitutive.

Note that *like* can also co-occur with a verb of saying, a fact that led to its classification as a quotative complementiser. Mostly, though, nowadays it accompanies the semantically empty ‘dummy verb’ *to be*.

The model given by Romaine and Lange (Figure 4.2) is typical for unidirectional approaches: it traces the diachronic development of *like*, concentrating on the syntactic development of the marker and trying to link it up with the semantic-pragmatic facts. In this chapter, I will sketch a semantic-pragmatic model, which I find is more able to conceptualise the synchronic facts. This is in tune with Romaine and Lange themselves, who propose (1991:262) that a “network of related meanings” captures best what we find synchronically. In the same line, Ferrara and Bell (1995: 281-2) suggest that the “present plasticity of *be+like* - that is, its ability to introduce constructed dialogue, thought, and quotable gestures of self and others - arises from systematic variation and expansion from a core, paradigmatic case. Individuals can range along a continuum of usage levels”.

But even when we look at it from a diachronic perspective, *like*’s development does not seem so straightforwardly linear as many might wish to believe. In a short excursion on *like*’s diachronic development, I will trace the developmental history of its various functions according to the OED.

First, I will sketch the history of suffix *-ly*, a cognate of *like*. Originally a free form *lich*, meaning ‘body’<sup>7</sup>, its development is typical for grammaticalisation. It follows exactly the course as predicted by Meillet (1912, cf. also Kuryłowicz 1965:52). The cline is independent word > suffix with lexical properties > suffix with grammatical properties.

---

<sup>7</sup> Consider the OED entry for *-like*<sub>1</sub>: “The original Teut. adjs. in *-liko-* were compounds of the n. *\*likom* appearance, form, body (...). Thus *\*mannliko-* (‘manly’) means etymologically ‘having the appearance or form of a man’; (...) The primitive force of the suffix may therefore be rendered by ‘having the appearance or form indicated by the first element of the word’. When appended to ns., the most general senses of the suffix in all Teut. langs. are ‘having the qualities appropriate to’, ‘characteristic of’, ‘befitting’. In English of all periods it has been a prolific formative....”.

In present day English, *-ly* is a grammatical morpheme with the function of a denominal adjectivising suffix (compare with French *-ment*). Suffix *-ly* looks like a school-book example of grammaticalisation and indeed, in the course of its development, we notice phonological attrition and semantic bleaching to the point where *-ly* attains the status of a dummy marker. It signals grammatical meaning ‘adverb-forming suffix’ and change of word class. Also, the process of substitution has taken place; *lich* no longer exists.

The history of *-ly* stands in sharp contrast to the free form *like*, which is a direct descendant of the adjectival form *gelic* ‘having the form of’. The following sketch gives every new attestation of the form according to the century in which it first occurred (the symbol ‘>’ reads, according to the convention in historical linguistics, as ‘leads to’). Attested uses of *like* are as follows:

- |       |  |
|-------|--|
| C14th | adj. ‘in the same manner, extent’, ‘similar’ |
|       | > - approximately                            |
|       | - as if                                      |
| C19th | > - for example                              |
|       | - as such                                    |
| C20th | > - discourse marker                         |
|       | - discourse introductory item (quotative)    |

Hence, the adjective *gelic* with the semantics ‘in the same manner, extent, similar’ gave rise to the meanings ‘approximately’ and ‘as if’ in the 14<sup>th</sup> century. It further gave rise to two other meanings, ‘for example’ and ‘as such’ in the 19<sup>th</sup> century. And it further branched off twice in the 20<sup>th</sup> century, giving rise to the discourse marker and the quotative function. This scenario does not correspond to the one painted by the common unidirectional models where one item develops into another one in a chain-like fashion. And it certainly does not look at all like the developmental history of *-ly*. Romaine and Lange comment on this situation as follows: “certainly, historically, the nondiscourse functions of *like* appear to precede the discourse uses but that may simply reflect the nature of our evidence. Both coexist synchronically and the steps we are postulating here are almost *certainly not strictly sequential* since in Middle English, even where *like* was used in clearly prepositional function, it had some positional mobility ...” (1991:262, *emphasis mine*). It seems that even diachronically, a unilinear model cannot account for *like*’s development. There is no chain-like development but more a situation where

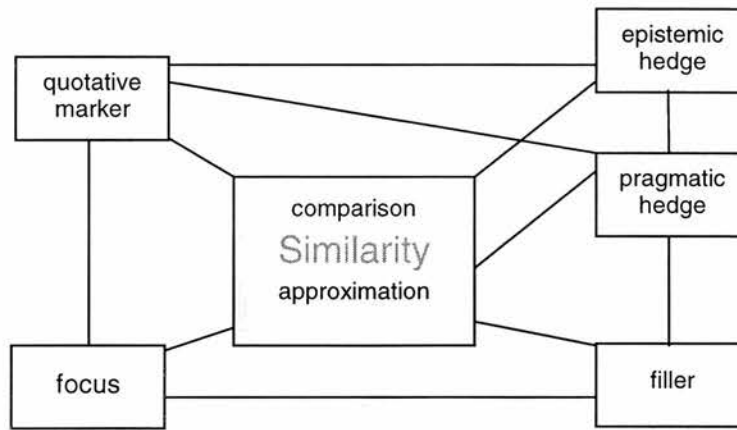
functions bifurcate and spread out further and further. Furthermore, a channel suggests a suppletive development whereas, synchronically, we find persistence of meaning. The newer uses of *like* have not supplanted the older ones and all uses alike seem to persist with a great deal of vivacity as there are no indications of suppletion or of earlier items decreasing in frequency. Furthermore, we witness no cases of phonological or morphological attrition.

This diachronic excursion has shown that unilinear models are useful in explaining chain developments such as the one of suffix *-ly*. But in cases where multiple meanings develop simultaneously out of one original core, unidirectional models are not able to capture reality as a whole, neither synchronically nor diachronically. This is because there are more dimensions to the grammaticalisation process than only one, linear chain. We need an explanatory model that can account for lateral connections between functions developing simultaneously. The evidence presented above calls for a more complex model.

#### 4.3.1. The radial structure

In the case of overlapping and ambiguity of meaning between functions at opposite ends of a grammaticalisation “channel”, a chain model, or even a branching one, seems to me a less appropriate account of the synchronic linguistic reality. So far, I have found no model in the literature that is able to account for the inter-relatedness and ambiguity of the grammaticalised uses of *like* amongst themselves. But Ericsson (1995), Klamer (2002), Meyerhoff and Niedzielski (1995, 1998), and Mosegaard Hansen (1997) have proposed field models for the conceptualization of multi-meaning entities in languages as diverse as Bislama, French, Swedish, and Kambara. In this section, I will show that it is possible to conceptualise *like*'s often highly ambiguous and overlapping uses, which depend heavily on the intra- and extralinguistic context, in a coherent model without having recourse to a chain. Figure 4.3 shows the semantic field of all uses of *like* configured as a radial structure containing its core meanings and all its synchronic functions centering around it.

Figure 4.3: The Radial Structure Model<sup>8</sup> of *LIKE*



Note that this chart shows a synchronic semantic field. It does not make any strong diachronic claims concerning the temporal sequence of *like*'s functional extensions. While it is true that a radial structure has a temporal element to it as it shows the stabilization of the diachronic development in a synsemantic field, this diachronic component is nevertheless not the main feature. The structure aims at displaying the synchronically existing functions of a linguistic item and the way they are related to each other. These functional extensions are not predictable but motivated by connections that the language users perceive or make between the respective phenomena. By this feature it has also mildly 'predictive' value; it does not make further functions of the item (or of the same lexeme in another language) necessary but somehow "less surprising" than others (Lakoff 1987). The ensuing model is thus both a description of the synchronic status quo of the development of the lexeme, and a motivation for the potential development of further functions.

The above model depicts the semantic functions that *like* synchronically exhibits as spaces (boxes) in semantic space. The semantic-pragmatic links that speakers perceive, and which are connecting those functions, are represented as lines between such functional categories. In the following discussion, I will show that *like*'s synchronic functions have a common underlying denominator, depicted in this

---

<sup>8</sup> The term 'model' has been used in the literature with varying meaning and without any clear definition. It seems in order to justify my choice of terminology. I have chosen the term 'model' in order to keep within the legacy of papers on the grammaticalisation of discourse phenomena, a literature which tends to refer to explanatory schematic representations as 'models'. In this article, I will take the term 'model' to mean 'empirically validated graphical representation with mildly predictive value'.



model as the shaded central functional space. Bybee, Perkins and Pagliuca (1994) point out that the source meaning restrains the grammaticalisation path an item will travel in its semantic development. The functions that arise from a linguistic item are related to the source construction, a phenomenon termed “persistence” by Hopper (1991) and “split” by Heine and Reh (1984:57-59).<sup>9</sup> Hence, if semantic-pragmatic extensions from one original core are not arbitrary but motivated, this implies that the output of a meaning shift still necessarily contains a part of its input.<sup>10</sup>

I agree with Romaine and Lange (1991), Miller and Weinert (1998) and Fleischman and Yaguello (to appear) that the newer functions of *like* have their origin in the semantics of comparison. I assume as the basic core meaning of *like* the notion of similarity which I take to be the basic underlying notion of both comparison and approximation (cf. Romaine and Lange 1991). This is for two reasons. The first is historical: the OED gives as the first entry for *like* “Having the same characteristics or qualities as some other person or thing; of approximately identical shape, size, colour, character, etc., with something else; *similar*; *resembling*; *analogous*. c1200 ORMIN 7931 *þe 33re sang iss lic wi 33 wop.*”<sup>11</sup> (*italics mine*). The second is that, as I will show in the discussion to follow, similarity is the persistent semantic trait which can be found throughout the extended functions. The connection between comparison and similarity has been motivated by Bredin (1998:69) who describes comparison as “the process in which I inspect the things being compared in order to discover whether they are alike or not”. I conclude that comparing is looking for +/- similarity. Under approximation, I understand the notion which holds when two concepts are “similar (...) but not exactly the same” (Collins Cobuild online, <http://www.linguistics.ruhr-uni-bochum.de/ccsd>, March 1<sup>st</sup>, 2003, search for ‘approximate’). Hence, similarity is the semantic primitive inherent

---

<sup>9</sup> The literature seems to agree on the fact that outcomes of grammaticalisation are related. But there is much less consensus on the way in which the various functions of grammaticalised constructions are linked. Whereas Sweetser (1988) and Perez (1990) claim that metaphor is the major creative impulse for grammaticalisation, Hopper and Traugott (1993) and Traugott and Heine (1991) see metonymy or association in linguistic context as the driving force. Bybee and Pagliuca (1985) and Bybee (1997) hold generalization to be the underlying force, whereas Traugott (1995) postulates that realignment lies behind the transfer.

<sup>10</sup> Lichtenberk (1991) claims that a common core meaning exists only historically but that synchronically, there is not necessarily a common denominator exclusively shared by all functions.

<sup>11</sup> Lit. Their song is *like* with weeping.

in both comparison and approximation. The similarity notion then gives rise to various other meanings, which I will show to be partly interrelated and which still contain this core semantic meaning to a greater or lesser degree

I will now motivate this model by illustrating in what way *like*'s multiple uses are interrelated and how they can be tied to one core meaning. The following paragraphs present a categorization of the synchronic functions of *like* as found in my data: first its use as a discourse marker in its various functions, then in its use as a quotative complementiser. I will show that while the existing functions are still more or less closely linked to this core meaning, one can nonetheless not postulate one single grammatical channel joining uses that are progressively more remote from this core, even synchronically. Rather, my claim is that these functions are semantically and structurally linked in a field around this one core. I will support this claim with cross-linguistic evidence. Later, I will argue that the same kind of model can also be used to explain *go*'s synchronic functions and display how *go*'s development can be understood as a non-suppletive development of multi-layered meanings via functional extensions.

#### 4.3.2. The hedging function

*Like* as a preposition or as a conjunction has comparative function. It signals near-identity between the compared and the comparator. This can be seen in examples such as *Paul looks like my father*, where it is assumed that *Paul* has a close resemblance to the speaker's father. Consider in this respect function 1 given by the OED: "Having the same characteristics or qualities as some other person or thing; of approximately identical shape, size, colour, character, etc., with something else; similar; resembling; analogous."

In talk-in-interaction, speakers can compare entities using *like* even when there is a substantial difference between the two entities compared.<sup>12</sup> But the question poses itself: when speaking of real-world similarity, how different can two entities be and still be compared using the word *like*? While this question cannot be solved

---

<sup>12</sup> This and the statements below obviously do not apply to metaphorical comparisons such as the ones in poetics (cf. Bredin 1998:69).

off-hand and will remain a matter for further investigation (Buchstaller, in preparation), for the purposes of the present discussion, I will accept Schourup's (1982a: 30) proposal of a comparative "more like" reading for *like*, which can still be subsumed under its standard meaning "somewhat resembling" (Webster's New World Dictionary 1994). Consider example (4.1)

(4.1) BrE - *The barbecue*

- V: but ehmm she's ahmm,  
so we just thought- and the other ones were sort of standing outside,  
and we were just so,  
like you know where the toilet were and everything,  
and then we just went down the bridle path.  
(...)
- K: we were sitting here.  
there was **like** ah::: panel then.  
and they were on that side,

Speakers V and K describe their and their friends' position with respect to the placement of some other acquaintances. The geographical location is given by V. Then, K describes where they *were sitting*: in a place where they were shielded from their friends' because they were on the other side of a panel. V uses *panel* as a representation of a dividing item of some sort. Knowing *panel* is probably not the exact term, she employs it as a substitute for the lexical item she does not know or she does not have in her immediate active word stock (if there is any).<sup>13</sup> A short word search for a better and perhaps happier lexical choice is also indicated by the prolonged *ah:::* before the word *panel*. Using *like* before the lexical item she eventually picks, V implies that the item they had been shielded behind somewhat resembled a *panel*. *Like* in this function is still related to its old core use of similarity: the comparative notion pertains between what was said and what was conceptually intended. *Like* thus signals that the listener should not take the utterance too literally and should be aware of a potential discrepancy between what the speakers have in mind and what they actually say (Siegel 2002). The approximative value of the verbal representation is acknowledged and highlighted (cf. Fleischman

---

<sup>13</sup> I am grateful to Keira Ballantyne for pointing out to me that, in some cases, speakers may be reluctant to put in exact technical terms even if they know them because the assertion of in-field knowledge would put them in the position of claiming knowledge of a field that they might be unwilling to take on (cf. Channell's (1994) concept of vague language).

and Yaguello to appear). Schourup (1985) argues that by using *like*, speakers alert their interlocutors to the “loose fit” between what was intended and the overt expression; Andersen (1998) speaks of “loose talk”. *Like* can thus be interpreted as a signal of imperfect rendering of what the speaker actually intended to express, an epistemic hedge (on the notion of hedge, see e.g. Brown and Levinson 1987, Holmes 1984, House and Kasper 1981, G. Lakoff 1972). This is also underlined by findings of Sweetser (1991:28), who points out the frequent link between items that signal physical likeness and epistemic stances.

*Like* as an epistemic hedge does not fully commit the speaker to the content of what she says. In Underhill’s words (1988:241) speaker V “leaves the statement slightly open”. She chooses a more general term which she hedges with *like* because openly stating her uncertainty or using a more exact but possibly inaccurate lexical choice would (if she were taken to task for it by her interlocutors) constitute a threat to her positive face.<sup>14</sup> Thus, taken at an interpersonal level, *like* functions as a pragmatic hedge. By marking the lexical choice as approximate and giving the speaker reduced responsibility, *like* can be a face-saving device used to forestall criticism.

The literature mentions the use of *like* with numbers and the usefulness of an item meaning “approximately” with numerical material has been pointed out. Consider the following exchange:

(4.2) US English - *Cold weather*

- A: I mean, one day it was **like** sixty below zero.  
B: um are you serious?  
A: and uh- uh they said “uh, you know, stay inside”.

Speaker A has not measured the temperature herself (and cannot remember it exactly either it seems). Especially with rough estimates of numerical values such as this statement, *like* shields the speaker from being taken to task for non-precise information by the interlocutor (Precht 2003). Schourup (1982a:37) paraphrases this

---

<sup>14</sup> For the notion of face see e.g. Brown and Levinson (1987), Holmes (1992).

as “What I say is *like* what I mean”. *Like* “invites the hearer to infer a comparison, either actual or hypothetical” between the actual number and the nearest round number (Romaine and Lange 1991:246). If we are not in a situation where exact numbers matter (such as a physicist researching the exact boiling point of Radon at a certain altitude), Horn’s statement “say no more than you must” and Grice’s (1975) 2<sup>nd</sup> Maxim of Quantity hold. According to Wittgenstein (1978), definitions are exact if they are good enough for whatever purpose they have been introduced. Hence, numerical exactness is often not only unnecessary but would even come across as odd. Unless precision is required, speakers give approximate numerical values and have the choice of marking them as such with *like*.

Berlin (1992) argues that some entities are more representative, more focal than others, especially when categories are fuzzy. This is in accordance with Rosch’s (1975, 1978) prototype theory, which holds that human categorization is not arbitrary or accidental but rather the result of psychological principles of categorization. In (4.2), the round number *sixty* is not arbitrarily chosen over the possibly more accurate *fifty-seven* or *sixty-four* but rather because it is a focal member of the category. It is more cognitively salient than other possibilities and is thus more prone to be chosen given the time constraints of talk-in-interaction (the same obviously also holds for non-numerical concepts such as *panel* in (4.1)). *Like*’s effect as a linguistic hedge before the numerical expression is to signal the slight deviance of the real-world phenomenon from the reference point, *sixty*. Arguing with Rosch (1975:533) that natural categories such as numbers have “anchoring points”, in relation to which other stimuli are perceived and classified, *like*, with its comparative semantics, marks the relation of a prototype and its less prototypical, specific instantiation of the category. The hedging via *like* in situations where prototypicality is involved can be very nicely seen in (4.3).

(4.3) US English - *The discovery*

- A. and what did he discover ?
- B. well,  
     he’s still in the process of being discovered.  
     what he was discovering,  
     he discovered (.) higher- **like** millions of [[indistinct]]

*Millions* here can be understood as a prototype for ‘a lot’. As it is a multiple of 1000, a salient number in the decimal system, it provides “a cognitive focus of [the] human-processing mechanism[s]” (Berlin 1992). The choice of the prototype in place of a non-focal member is not random but because it is a conceptually easily accessible classificatory item.

### 4.3.3. The focus function

In the literature, *like* is often interpreted as a focus marker (Blyth et al. 1988, Dailey-O’Cain 2000, Romaine and Lange 1991, Schourup 1982a, Underhill 1988). Haiman (1989:310) states that “the comparative construction is one which contrasts, and hence focuses the elements which are compared... the element compared being more highlighted” (see also Nølke 1983). Hence, the semantic link between comparing and focusing seems to be a fairly salient one (see also Rooth 1996). In this chapter, following Underhill (1988) I interpret the most significant information in a sentence, or, by extension, in a chunk of speech as focus (cf. also Kuno 1980).<sup>15</sup> If *like* marks such salient information, we should expect to find it adjacent to linguistic material that has such status. Consider example (4.4):

(4.4) BrE - *Food cravings*

- Y: she was even **like** into meatballs,  
X: phh gosh your sister is straining you.  
ha ha,  
Y: we all know that.

In (4.4) speaker Y has been telling her interlocutor about her ogreish sister. She is enumerating the things her sister has cravings for and concludes her list with stating the fact that her sister *even* had a yearning for meatballs. Note in this respect the OED entry for *even* “9. Intimating that the sentence expresses an *extreme case* of a more general proposition implied (= Fr. *même*). Prefixed (in later use often

---

<sup>15</sup> For other definitions of focus consider Beheny (1998), Jackendoff (1972), Lambrecht (1994), Sperber and Wilson (1986:202ff), Vallduvi and Engdahl (1996), from a typological perspective see Lambrecht and Polinsky (1997). Consider also individual papers in the edited volumes Downing and Noonan (1995) and Payne (1992).

parenthetically postfixed) to the particular word, phrase, or clause, on which the *extreme character of the statement or supposition depends*” [emphasis mine]. Thus, *meatballs*, preceded by *even*, is the extreme last member of a climactic list of staples (cf. also Berckmans 1993). It is double marked as such with *like*, which acts as a focus marker just like *even*.

Note that focused elements need not necessarily be new. Older, known elements can be focused on when speakers tap into old common knowledge and reintroduce it into the currently activated information state of the conversation (cf. Gundel, Hedberg and Zacharski 1993). Alternatively, currently activated knowledge can stay in focus over a stretch of conversation as can be seen below:

#### (4.5) BrE - Calculator breaking down in the middle of a maths exam

- X: I was “ohh God,  
I have to go through couple of maths exams higher math without a calculator”.
- Y: I can imagine Nicola because you know how her old one,  
where you used to [bang it,
- X: [bang it,
- Y: I’d see her **like** banging this ehhm calculator to get it on during the exam,  
but she got a new one so,
- X: ahhh so.

Here, the concepts *banging*, *calculator* and *exam* have already been introduced into the current state of awareness during the TCUs before the one containing *like*. What the TCU with the arrow does, though, is bring them all together in one coherent picture. Y paints a vivid picture of *Nicola* trying to get the calculator to work. It is this scene which contains a whole scenario which is in focus and marked as such with *like* even though it contains nothing conceptually ‘new’.

Note the fact that the focusing function overlaps with the approximative function especially in numerical contexts such as the following:

#### (4.6) US English - Cattle farming

- A: they’ve got to have them there to milk them,  
and they- and I read in this article,  
I couldn’t believe it where you know like one cow produces- like a day,  
produces **like** a hundred or so pounds of manure,
- B: o-oh

When speaker A states how much manure a cow produces, this measure is obviously an approximation. As estimates of this kind are averages that do not apply to all members of the species cow at all times, any number will necessarily be approximate. The rough and ready numerical indication *hundred*, again a focal member, is marked with *like* as well as with *or so*. But note that, as in this case, measures are often the focused information in a piece of talk as they answer the important question “how much?”. In (4.6), the phrase *a hundred or so pounds of manure* is in focus. *Like*, which is immediately preceding it, can be interpreted as a focus marker as well as a hedge. Indeed, in all of my examples *like*'s focus meaning seems to co-occur with another of *like*'s synchronic functions as can be seen in the next example. *Like* here precedes a stretch of speech, the content of which is a typical situation, the story of someone who suddenly becomes rich and famous.

(4.7) US English - *Talking about a man coming from India to England:*

- X.      and he was **like** put up in a house,  
          Cambridge and everything,  
          was just amazing.

X relates a story about a poor immigrant from India who comes to England, finds his way into the highest intellectual circles and ascends the social ladder in no time. *Like* in this context could be interpreted as hesitant or hedging. This would in turn be underlined by the fact that the speaker does not attempt an elaborate expression of her thoughts - marked by *and everything*. But intonation suggests otherwise: Because the speaker utters what follows *like* in a monotonous, dragging voice, I interpret the material following *like* as stereotypical, commonly known information that only has to be hinted at and needs no elaboration. Because the story is a well-known motif, X marks it as that: the prototypical success story. *Like* precedes this material and serves as a focus marker (Schourup 1982a). Note that speaker X only gives an approximate rendering of the story, because she assumes that it is so well known that everyone can fill in the rest. The use of the general extender *and everything* (Overstreet 1999) signals the speaker's tapping into the common knowledge of her interlocutors. Here, *like* retains some of its core comparative function in the sense that this illustration is compared to and seen as one instance of



the prototypical success story. In accordance with Hopper's principle of persistence (1991), *like*'s similative meanings persists in its focusing function.

Note also that *like* in this use is very close to its quotation-introduction function discussed later in this chapter. The speaker uses prefabricated concepts or even parts of speech she has heard or read before, embedding them into personal narration from her perspective. *Like* then sets off parts of speech the speaker cannot claim responsibility for, a sort of indirect quote of a stereotype (Meyerhoff 1992). This patchwork of chunks of speech of various authors and "voices" (Bakhtin 1981) creates a speech mosaic (Buchstaller 2001a). Goffman (1981) talks of "lamination". The problem of speaker commitment to quotes that have become stereotypes or somehow 'common goods' has been commented on by many scholars, see especially Bakhtin (1981, 1986), and Vološinov (1973). Flaubert thematises the second-hand nature of every-day discourse which builds on knowledge in the form of pastiches which are reiterated over and over again in his 'Dictionnaire des idées reçues' (1966) and its fictional counterpart 'Bouvard et Pécuchet' (1966). As I will exemplify in section (4.3.5), with *like* used as a quotative complementiser, the speaker can mark the borrowed part as second-hand, as reported speech.

#### 4.3.4. *Like* as a filler

Current discourse analytic research (Fox and Jespersen 1995, Schegloff 1996) shows that speakers plan in advance while it is not their turn and then jump in and claim the floor without necessarily having properly planned their whole turn ahead. As a result, problems can arise when the utterance is in the middle of production. Faced with the threat of losing the right to the floor, speakers who have run into problems of formulation and who have to re-structure their lines of thought and their possible wording are forced to use strategies which allow them to avoid pauses where their role as current speaker is at risk. Fillers (*ehm, well, you know*) lengthenings (*a ma:::n, we:::ll*) and repetitions serve as time-buying mechanisms where speakers are heard saying something while actually not saying anything new or contentful. *Like* is one of the possible fillers used in such situations. As an

extension of its ‘more like’ approximate sense, *Like* can assume this function, which has been most criticized by normativists as it is identified with slang or very casual speech (Dailey-O’Cain 2000, Schourup 1982a:39, Underhill 1988:234). It usually precedes afterthought modifications by speakers who want to continue their utterances but have difficulty in formulating them. In accordance with this, in my data, *like* in this use often precedes a restart or an anacoluthon.<sup>16</sup> Consider, for example, (4.8), where B starts off with *we*, breaks off and reformulates her choice of pronoun as *I* (*want a part time job*).

(4.8) BrE - *Part time job*

- A: and “I’m paying you such an such on full time,  
no we can’t afford you not to work”.
- B: I know **like** ahh we ehheh (.) - I want a part time job as well,  
dad doesn’t seem to agree but.  
I don’t know I want a part time job,

In (4.8), in the TCU marked with an arrow, B has obvious trouble formulating her thoughts. She tries to buy time by using the items *I know*, *like*, *ahhhh* and *ehhh*. In spite of these fillers, a little pause ensues which threatens her right to the floor.<sup>17</sup> I argue that *like* here has about the same function as other discourse markers, it is a “sound shadow” (Goffman 1981:109) to hold the floor. *Like* fills the pause, or part of it, and thus enables the speaker to keep the floor by filling in the silence. Consider also the next example

(4.9) US English - *Who are your friends?*<sup>18</sup>

- S: not only am I like one out of very few white people with blacks, but I’m also not=  
=exactly too wealthy,  
so **like** I don't know that's affected me a lot when it comes to- **like** being able to do=  
=things **like** - um there was one thing I wanted to do,

---

<sup>16</sup> This finding underlines Andersen et al.’s (1999:1339) claim that discourse markers signal production problems.

<sup>17</sup> Note that there is an alternative interpretation, *like* as a quotative in the sense *and I know like* “*ahh we ehheh (.) - I want a part time job as well*”. But intonation suggests otherwise: there is no noticeable jump in the intonation contour (Klewitz and Couper-Kuhlen 1999) nor is there any indication of any of the other suprasegmentals, which have been shown to mark quotes, such as interjections, exclamations etc. (cf. Güldemann 2001:28).

<sup>18</sup> Source: The Smart Museum of Art, Dawoud Bey: The Chicago Project > Simone ([http://smartmuseum.uchicago.edu/chicagoproject/portrait\\_simone.shtml](http://smartmuseum.uchicago.edu/chicagoproject/portrait_simone.shtml)).

*Like* again occurs at the points where the speaker hesitates several times (*so like I don't know... it comes to- like being able .... things like um there was one thing*). This results in clusters of *like*. It co-occurs with other items that have been classified as discourse markers with similar functions (*so, um*) in earlier accounts (Brinton 1996, Östmann 1995, Schiffrin 1987). The phenomenon that several discourse markers, which are syntactically redundant and semantically shallow, co-occur and are “multiply reinforced” has been considered one of the distinctive features of discourse markers in general (Romaine and Lange 1991:251 cf. also Schiffrin 1987:66). It can thus be counted as supportive evidence for the view that *like* functions as a discourse marker in this particular sequence. Note that, while letting the speaker gain thinking time, *like* is also suspending the hearer's attention. It is signalling that there is more to come because the material following *like* is portrayed as standing in a relation of comparison or approximation to what has already been said.<sup>19</sup> By opening up a comparative notion, the construction is suspended as it needs a comparator in order to be complete. Thus, by using *like* speakers put their interlocutors on hold.

As speakers are forced to make complex choices in real time during conversation (Ferrara 1992), the choice of an item such as *like* is motivated by the fact that not much conversational harm is done if the interlocutor interprets *like* at face value as a simulative item. Potential similarities are everywhere and their salience is often very subjective. Hence, if interlocutors interpret *like* as a comparative and even if they themselves do not perceive any obvious comparison, they may put this down to a speaker-specific interpretation of the situation which they do not share. *Like* is the perfect particle to fill a pause, and to hold the floor. By indicating that what follows is only an approximative rendering, the pause is “detoxified” (Schourup 1982a: 46). This links up to the discussion under 4.3.3: the presentation of new or newly focused on information can trigger problems of formulation. Using *like*, a marker with an approximative function, speakers can introduce focused material - marking it as such while giving the speakers time to mentally prepare their following speech.

---

### 4.3.5. The quotative function

Romaine and Lange (1991) have shown that syntactically, *like* can occupy a slot before a clause or a sentence. In US English and even more frequently in British English, it can also follow the material it has scope over (cf. earlier in this chapter). A wide scope is the syntactic prerequisite for the framing of quotes, which are mostly clausal. Hence, *like*'s increase in scope – from preposition (+NP) to conjunction (+IP) – has set the stage for its quote-introductory function. If *like* occurs within the context of a quotation and starts co-occurring with it more frequently, it can become associated with its environment. Note Gumperz's comment in a paper held at a conference in 1988 (1992:47) that *like* “foreshadows a quote”.

Indeed, once *like* has grammaticalised as a focus marker and focuses on the most significant information in a stretch of speech, it is not surprising to find it co-occurring with quoted material. Discourse analytic research has shown that speakers incorporate reported speech into narratives in order to simulate the feelings and the setting at the time of the quote and make the narration more dramatic or expressive (Blyth et al. 1991:222, Ochs 1997, Yule and Mathis 1992).<sup>20</sup> Also, because storytelling with incorporated voices is more captivating (Labov 1972c), it makes it easier for the speakers to hold the floor (Nordberg 1984:20). This is because speakers who give the protagonists narrative voices do not take on the function of an omniscient narrator who covers all roles. By presenting a narration as a series of interchanges, like in a radio-play, a whole auditory scenario is built up in order to involve the interlocutor. Speech with these characteristics has been called “involved style” by Wolfson (1982), “performed narrative” by Tannen (1989) and “replaying” by Goffman (1981). Quoted speech usually occurs most heavily around the climax or the coda of the story and qualifies thus as the most focused part of the narrative (Fleischman and Yaguello to appear, Labov 1972c). The quotation can then be

---

<sup>19</sup> In its filler function, *like* still retains traces of its comparative approximative meaning, even if its propositional meaning has already become bleached (Lehmann 1985, Romaine and Lange 1991, Sankoff et al. 1997).

<sup>20</sup> Labov (1972c) and Chafe (1982) state that narratives are more vivid when direct speech is used to report dialogue.

interpreted as a variant of that focus. The semantic overlap between focusing and introducing quotations is thus motivated.

The observation that *like* is an item heavily used for introducing interjections, sound and voice effects and other mimetic enactment of previous events (cf. Chapter 3) further underlines this claim. When speakers include such non- or paralinguistic elements, they perform the reported event rather than merely tell it. The aim of such performances is to create listener involvement and to increase the dramatic impact of the story. Güldemann (2001) claims that *say* and other common speech verbs focus on the semantics, the propositional impact of the quote, which is then not particularly salient. By contrast however, other, more marked frame elements such as *like* and *go* focus on the presentation. This underlines Blyth et al.'s (1990) statement that "*be like* may be viewed as a focus quotative, that is a quotative which introduces a particularly salient piece of information packaged in the form of reported speech".

Note the obvious link to Goffman's (1981:99) "response cries", which are used to "show or index the mental state of the transmitters", to "clarify the drama of their circumstances". Consider the example below:

(4.10) US English - *Bottle deposit*

- A: and it was they come in like uh eight and and eight and ten packs you know,  
B: right,  
A: instead of six-packs and uh and they were like,  
→ it was like two dollars and something for the bottles,  
you know **I was like** "Go:::::d almighty it costs more for the bottles than it did =  
=for the cokes".  
B: that was my brother's first job in a grocery store,  
he was in the bottle area,

Speaker A has been telling a story about an experience she had in her local supermarket during the first stumbling attempts to introduce bottle deposits in the USA. The story stretches well beyond the excerpt represented here but has been trimmed down for space reasons. In the line marked with the arrow, A represents her reaction to paying enormous deposits. She re-enacts herself as producing the hypothetical speech act *Go:::::d almighty it costs more for the bottles than...* Note the lengthenings and the stretches of loud speech. This quote incorporates her feelings in the situation. The emotionally heightened event is expressed through

marked suprasegmentals. This quote is at the climax of the narrative and concludes it; B starts another narrative on the same topic after it. The use of focusing *like* before the emotionally and narratively salient quote ratifies its interpretation as a quotative as well as a focus marker. *Like* frames quotes containing sound and voice effect that are salient in the narrative line. The use of *like* with mimesis goes along with the focusing effect.

In addition, the filler function of *like* makes it an item well-suited to use before quotations. Quoting means inserting a new mental space (Fauconnier and Sweetser 1996) which hosts re-enactment of other or self-quotation at a time removed from the speaker's (here and) now. In order to quote, the narrator needs to remember the original speech, sound or voice, s/he intends to represent (which is then re-enacted). Furthermore, as the linguistic host material of the quote has a different deictic orientation, personal and temporal orientation needs to be shifted corresponding to the original utterance. The cognitive processes that underlie such restructuring need time. As a consequence, we typically find large numbers of discourse markers with filler functions around quotes. In my corpus, roughly one fourth of all quotes co-occur with discourse markers (often doubly marked), the most frequent of which is *you know*, followed by generalized continuers (Overstreet 2000) such as *or whatever*, *and all that* etc., and the markers *just* and *like*. The following excerpt shows *like* co-occurring with other filler items in a quotative situation.<sup>21</sup>

(4.11) US English - *Reading the classics*

- 01 A: uh huh  
02 B: I'd read some Dickens before but I hadn't read that one,  
03 and I- I was **like** (.) I thought to myself **you know** "what other things are out there=  
04 = that are that are classics"?  
05 A: uh huh,  
06 B: **you know** "that have just withstood time and that are just excellent,  
07 something's excellent about them,  
08 whether it's the way they are written or whether it was the material they were=  
09 = written" **you know**,

Speaker B's reports of her thoughts are framed by *I thought to myself* in line 03. The on-line pressures of reproducing the original inner quote lead to phenomena

---

<sup>21</sup> As I will show in the next paragraph, another meaning of *like*, namely its hedging function, plays an important role in the re-analysis of *like* as a quotative introductory marker.

commonly associated with repair (Sacks, Schegloff and Jefferson 1974, Fox and Jespersen 1995) such as re-starts (*I-I*), a pause (.), and several instances of the filler *you know* (in bold in lines 03, 06 and 09). Also, we notice an instance of *like* before the quotative frame (in line 03). Sperber and Wilson (1986) have shown that in utterance interpretation, speakers interpret concepts in accordance with their context, a phenomenon called “context-induced reinterpretation” (cf. the related concepts of “pragmatic strengthening” (Traugott 1989), and “strengthening of informativeness” (Traugott and König 1991). As concepts are subject to contextual factors in utterance interpretation, depending on their small-scale loci of occurrence, we can assume that if *like* occurs in the vicinity of reported utterances, its context of use - quotation - can encroach upon its interpretation. *Like* can take on the function of an introductory item for quotations.

A quantitative analysis of my corpus yields the finding that *like* is the third (in USE) / fourth (in BrE) most frequent lexical discourse marker occurring in and around quotes.<sup>22</sup> Several papers in Bybee and Hopper (2001) show that frequency has an important effect on the perception and cognitive treatment of linguistic material as related. Thus, via the process of metonymic transfer, the string *filler like* + *quote* can be perceived as a quotation frame. Through association with its environment, *like* can change category membership (Krug 2001).

Thus we have a situation where the same surface form belongs to different but related functional categories, even more so as *like* is often ambiguous in its function. In example (4.11), *like* could be interpreted as having a filler function, regularly co-occurring with quotations and buying the speaker time to mentally prepare the incorporation of the quote, just as *you know* does a few lexical items later. At the same time, once *like* has already acquired its grammatical function as a quotative, it could also be interpreted as a quotative item from which speaker B repairs away (the repair would then be *I was like* → *I thought to myself*). In this account, B had initially planned to use the quotative frame *I was like* “quote” but repairs it with the

---

<sup>22</sup> The sequential positions under which I have counted a discourse marker as directly associated with quotes are the following (slots marked with X): immediately before, in or after the frame as in X *pers. pronoun* X *V<sub>quot</sub>* X “quote”, as insertion sequences (analysed as such on the basis of intonation) as in “quote” X “quote” or immediately after the quote as in “quote” X. Obviously, discourse markers can occur in any slot at any time, and they frequently co-occur.

strategy *I was thinking to myself* “quote”. Tannen (1986) has shown that if one lexical item occurs in a given context, it is more likely to be chosen by the speaker(s) over and over again, a case of lexical priming. Scherre and Naro underline this claim on a larger scale (1991, 1992). Hence, if the discourse marker *like* regularly and frequently occurs in the vicinity of a quote, it will not be surprising to find that the same surface item, once it has acquired the new quotative function, is chosen for quote introductory functions as well.

When quoting, speakers report the utterance, but its form and content can only be rendered approximately because of the idiosyncrasy of expression in terms of suprasegmentals such as accent, style, prosody of the original speaker. Due to her imperfect memory of the original utterance (Chafe 1980, Stafford and Daly 1984) and due to her personal restrictions concerning voice quality, pitch etc., the reporting speaker cannot give an exact rendering of the features of the original speech act. Tannen (1986) takes this into account when she claims that every attempt to quote is actually “constructed dialogue” (cf. Fleischman and Yaguello’s “interpretive quotative” (to appear: 9), see also Chafe 1994, Ducrot 1984, Tannen 1989, Mayes 1990, and many others). *Like*, an item which bears approximative function and has already been grammaticalised to function as a semantic and pragmatic hedge, seems the perfect item to mark the approximateness of reported speech. It explicitly mitigates the claims to verbatim reconstruction (Clark and Gerrig 1991). Example (4.12) gives an instance with *like* directly before the quote:

(4.12) BrE - *Take away food*

- B: well we are gonna go for the pizza but,  
well the first the first thing well that he **says like** “well me mate’s got a=  
=he owns he owns a taxi”.  
uh his mate,  
A: uh huh,

B reports what his acquaintance said when he and his friends were going to get a pizza. He represents the words *he* said as a direct quote. This quote is necessarily a “selective depiction” (Clark and Gerrig 1991:767) of the speech act as it was uttered by the original speaker in terms of voice, prosody, lexis, facial expression, etc. By adding *like* to the quotative frame, speaker B acknowledges the fact that he does not



render the speech act as a gestalt but only certain aspects of it, in this case only the linguistic material, without any attempt of replaying the original speaker's voice. The content of the quotation can only be an approximative rendering of the whole emotional and contextual situation. Using *like* with its approximative-comparative semantics signals the possible non-equivalence between what is reported and the actual utterance. The speaker retains a reduced responsibility with respect to what was said and how, as a *like* in quote introductory position does not commit them to the form and the content of the quote. The quotative frame thus consists of two elements: a verbum dicendi *say* and *like*, which functions as a hedge, both on the referential-epistemic, and on the interpersonal-pragmatic level.

Due to its occurrence in this and other positions in the vicinity of quotations *like* is thus en route to being grammaticalised as a quotative. Example (4.13) shows *like* with reported thought:

(4.13) US English - *Feeling stuck*

- B: like you don't have any money to- to get out,  
A: oh,  
B: so uh and uh I was **thinking you know like** “↑oh my God you know,  
I go back there and then and I get married there and now I have kids there,  
and then I'll never leave the place” you know,

In (4.13), *like* precedes a quote consisting of inner speech framed by *I was thinking*. Speaker B produces a quotation frame with the quotative, *you know* as well as *like*. *Like* in its discourse marker function can here be interpreted as a filler as well as a focusing marker and also as a hedging item. Structurally, because *like* co-occurs with *think* in the quotation frame in (4.13) and with *say* in (4.12), it can be interpreted as a quotative complementiser. While intonation often disambiguates between the functional possibilities, the point I want to make in this discussion is that it is sometimes not possible, nor even intended to be possible, to pin down the item *like* to one function only (cf. Eckert and McConnell-Ginet 2003). It is *like*'s ambiguous and overlapping functions that are the constitutive element of its grammaticalisation. Speakers can exploit *like*'s functional vagueness and use it in contexts where it could have multiple meanings without giving any cues that help disambiguate it.

Note that especially in BrE, the quite frequent association of *like* with quotation via the equally non-canonical quotative *go* as in *and I go like "what did you say?"* (4% of all *go*-quotative frames co-occur with *like* as compared to only 1% in US English). We can assume that speakers using the relatively new quotative *go*, which we will see is still a marked option amongst the (say) variants, might tend to be more unsure about their quotative choice and might accordingly use hedging *like*. This is especially the case as *go* very often occurs with quotes the content of which is inherently approximate, such as sounds, voice effects etc. Hence, the advent of the new variant *go* might have also opened the door for the grammaticalisation of hedging *like* in quotative complementiser function. While this is a hypothesis that needs further research, it lets us imagine a scenario where one new quotative option has led to the introduction of another new one, which is of high explanatory value given the historical facts.

In US English, where *go* is much less frequent overall, other channels seem to have played a bigger role in *like*'s grammaticalisation (as will be discussed below). But in British English, where *go* is the second most frequent lexical quotative option overall, the association of strongly stereotyped *go* with hedging *like* cannot be ruled out as a contributing factor in the development of *like*'s new function.

Once *like* collocates frequently with verba dicendi before quoted speech and thought, it can become associated with its environment. Reanalysis occurs and *like* acquires the function as a quotative complementiser. The mental salience of the link COMPARATIVE MARKER - QUOTATIVE COMPLEMENTIZER has been sustained cross-linguistically by Meyerhoff and Niedzielski (1995, 1998), Romaine and Lange (1991) and Schourup (1982a: 33-34). In other words, if in a number of languages the cognate equivalents of *like* have become discourse introductory items, we have cross-linguistic evidence for a correspondence between the functions of this linguistic item.

Once *like* has been sufficiently associated with the notion 'quotative', it can bear the functional load of introducing reported speech alone, without a verbum dicendi (consider Gumperz' (1992) statement that it foreshadows a quote). The next stretch of speech contains a particularly nice example, which typifies *like*'s

grammaticalisation history from a quotative complementiser to a full quotative item with dummy *be*:

(4.14) BrE - *Scary situation*

Y: pfh,  
he thought I was mad,  
→ he **said like** “what are you going on about”,  
→ and I **was like** “oh I know that was scary” .hhhhhh

Through the frequent association of *like* with the quotation (i.e. here as a complementiser with quotative verb *say*), *like* with dummy *be* can now take on the function of introducing quotes. This use is represented in the second line with an arrow. The grammaticalisation path goes from *say/think like “quote”* to *be like “quote”*. Consequently, *like* with dummy verb to *be* (Romaine and Lange 1991) can function as a quotative in its own right even without any preceding verbum dicendi which opens the “quotative realm”. This can be seen in the next example.

(4.15) US English - *Teaching English*<sup>23</sup>

S: I tried to get her to say hello,  
→ and **she'd be like** “(CHOKE)=  
C: [ha ha ha ha  
K: [ha ha ha ha] ha ha ha ha [ha ha ha ha ha ha ha ha ha]ha ha ha  
S: [no puedo no puedo”].  
→ **I'd be like** “yes you can,  
just [say ‘hello].  
K: [ha ha ha ha ha  
S: [hello’ Amaria”].  
K: [ha ha ha ha ha ha ha ha ha  
S: “(CHOK[E)”  
K: [ha ha ha ha]

In (4.15), *like* precedes quotations twice, once with 3<sup>rd</sup> person singular and once with 1<sup>st</sup> person singular. Here, *like*'s hedging function becomes particularly obvious as the content of the quotation can only be an approximative rendering of the whole emotional and contextual situation. I have shown in Chapter 3 that *like* is especially frequent in the framing of quotes containing sounds, prosodic and paralinguistic devices, gestures or mimicry (cf. Kendon's 1996, 1994 “quotable gestures”), the

---

<sup>23</sup> For the transcription conventions of this and the following examples, see again Appendix 2.

approximate character of which is evident.<sup>24</sup> Gldemann (2001) traces a scenario whereby mimesis markers are first introducers of onomatopoeic elements or gestures and only then occur as quotative verbs for the whole spectrum of reportable events including reported speech (cf. also Buck 1915). They only encroach upon this function later and can then become routinised for the whole category. This statement parallels other claims in the literature according to which *like* is only or prevalently used with sounds or gestures (Ferrara and Bell 1995, Tagliamonte and Hudson 2001 and many others).<sup>25</sup> By the time of the collection of my data, roughly 72% of all quotative *like*-use in US English and British English was with some form of mimesis. Hence, 28% of all quotes framed by *like* contained lexical elements **only**. The one third incidence of *like* with purely lexical material (and much higher frequency with lexical as well as mimetic material) justifies the claim that in my corpus *like* has generalized its function to a quotative item that can take all quotative complements.

Irrespective of the form of the quote, due to *like*'s approximative-comparative semantics, speakers can signal the possible non-equivalence of what is reported and the original utterance. Fleischman and Yaguello (to appear) contend that the reduced speaker liability brings quotations with *like* into the realm of indirect speech. A quotative introduction with *like* signals approximateness and is completely non-committal with respect to its quote. *Be like* functions as a hedging quotative, epistemically as well as pragmatically. Thus, when Tannen (1986:321) says that "*be + like* functions like a formulaic introducer, not by its literal meaning but simply by convention", I interpret this to mean that *like*'s function of introducing quotation, which is approximative by nature and focused on, is a direct extension of its literal meaning of similarity and is in the process of being conventionalised.

---

<sup>24</sup> This use is also typical of quotative *go* as I will discuss below (Butters 1980). Consider also the new use of *all* in American English as in *She was all "..."*, which seems to parallel the use of *like* as a quotative complementiser for token mimicry, sound effects etc (Waksler 2001).

<sup>25</sup> The fact that it is still associated with mimetic enactments is expressed by statements such as the following "there are several environments where quotative choice is constrained. ... a gesture, facial expression, or non-speech sounds *are all, are like or go* but do not *say*" (Singler 2001:260). While in my corpus, quotative *say* does occur with sounds and voice effects (especially with the older speakers who have smaller repertoire without *like* and with much fewer *go* tokens), it is certainly a fact that for speakers who have them in their repertoire, overall, the noncanonical quotatives unframed, *go* and *like* tend to take over the function of mimesis introduction.

Furthermore, as Ferrara and Bell (1995:279) point out, a clear boundary between speech and thought is hard to draw. Especially for first person, it is often impossible to distinguish thought from actual speech. Because of *like*'s still more or less inherent semantic comparative-approximative property, it can be given a 'for example' reading. It is the ideal item to frame direct speech and inner monologue and can be used to present imaginary discourse as if it took place. Consider the following example:

(4.16) US English - *Using plastic grocery bags as lunch-bags*

- B: yeah in fact I have one today,  
A: rig[ht.  
B: [the only problem with those is sometimes they got holes in the bottom.  
A: yeah [they-  
B: [and ha ha **it's like** "whoops there goes my chips,  
A: [yeah  
B: [okay fine"  
A: uh huh

At the point in time when B's chips fell into her lap she was sitting alone in her car, there was no interlocutor to whom she could have directed the words "*whoops there goes my chips...*". It does not really make any difference if the quoted material was actually uttered or if it was inner monologue. This quote is a hypothetical quote, as explained in Chapter 3 (cf. Chafe's 1994 "verbally uncommitted thought"). Now it becomes clear why *like* as a quotative is semantically and pragmatically a good and frequent choice to introduce such hypothetical quotes: Romaine and Lange (1991:227) claim that by using *like* "the speaker invites the listener to infer that this is what the speaker was thinking OR saying at this very moment". Rather than the exact words, the quote is the expressive content of the speech act or the original speaker's thoughts packaged in the more vivid form of reported speech. *Like* as a quotative, due to its simulative semantics, can frame direct reported speech and inner monologue. In using *like* as a quotative introductory item, speakers sidestep the problem of where thought begins and where speech ends by presenting the quote as if it had taken place without committing themselves to its actual utterance. "Discourse introduced by *like* blurs the boundaries between direct and indirect representation of both speech and thought report" (Romaine and Lange 1991:234). In this case, the link to *like*'s hedging function is obvious. Using *like*, speakers take a

non-committal stance towards the quote they are representing. They forestall possible threats to their face by counteracting the possibility of being taken to task for misrepresenting thought as speech and the other way round.

In example (4.16), B's speech is inward, a verbalization of what she thought at that moment. *Like*, as in the above example, precedes internal comments on the situation that can be given in this short form without having to step outside the quotation frame and give external evaluation (Labov 1972c). In this vein, Ferrara and Bell (1995) claim that the function of *like* in contexts such as the above has the function of the soliloquy, which was previously used conventionally on stage in order to convey characters' inner emotions to the audience, as it allows speakers to open up their internal worlds to the public.

Now that I have motivated the semantic/pragmatic ties between quotation and comparison, approximation, focus and hedge, I will try to describe the full messiness of *like*'s patterning in synchrony. Any model that aspires to do justice to reality will need to be complex enough to account for all of *like*'s attested distributive instantiations. It will also have to be flexible and simple enough to be of explanatory value with respect to all of *like*'s case-wise ambiguities.<sup>26</sup> After a discussion of its distribution in my corpus, I will suggest a way in which we can incorporate the status quo into a coherent model.

When looking at *like*'s occurrence with verbs of saying in my data, the first thing that stands out is the fact that the slot between quotative verb and reported material is not the only one it can occur in. Its position in the quotative frame is by no means fixed, and in my data, *like* quite often occurs before the verb, either in the sequence *NP like V<sub>quot</sub> "quote"* as well as in the sequence *like NP V<sub>quot</sub> "quote"* (cf. the examples below).

---

<sup>26</sup> Case-wise means the following here and the following paragraphs: *Like* is an item that displays multiple ambiguities. But between which functions these ambiguities hold can only be determined in the single, particular case. Hence, there are situations where *like*'s functions are ambiguous between the filler and the quotative function. In other cases, *like* might be ambiguous between a hedge and a focus marker etc. Only a case-by-case consideration can give evidence about which ambiguities are involved.

(4.17) BrE - *NP like V<sub>quot</sub> "quote"*

A: I mean it doesn't bother me if something doesn't have this or that,  
but **she like says** "ahhhh I got this and I got that",

(4.18) BrE - *like NP V<sub>quot</sub> "quote"*

T: and I'm thinking that's a bit stupid,  
because it's - it's what's in your head that counts and exercise in your brain,  
→ **like** she used to **say** " (...) do come to the church this week",  
and I'd say "no",

Also, as is shown in example (4.19), *like* can come before or after the quote, especially in British English as in this variety it is more frequently found with retrospective scope.

(4.19) BrE - *The driving licence*

B: I think I mean my dad asked us what I was getting what I wanted for me birthday,  
I mean he **said** "you I bought you a provis- provisional" **like**,

A: ay,

B: and I - I never really asked for driving lessons,

Here, hedging or focusing *like* has scope over the quotation preceding it. Notice that in cases like this, where *like* is not occurring directly adjacent to the traditional quotative verb, its association with the quotation is nevertheless given. Not only does it occur within the close vicinity of a quotation itself, but, also, while the slot it occurs in is not the one overwhelmingly used by quotative frames (90% precede the quotation in my corpus), it is nevertheless one of the possible slots for a framing item. *Like* in this slot is predictably much rarer in US English because in this variety, the occurrence of discourse markers with backwards scope is less frequent as discussed above.

Also, *like* can occur in an insertion sequence between a quote that is split into two parts, resulting in the pattern *V<sub>quot</sub> "quote" like "quote"*. This is exemplified in the excerpt below:

(4.20) BrE - *The not so nice guy*

A: I mean everybody says "he",  
**like** "him up there",

---

Speaker A represents a quote in which everybody is portrayed as saying “*he*”. This first quote is framed by the canonical quotative *say*. She then rephrases this with a second quote “*him up there*” which is preceded by *like*. Paralinguistic information alone tells us that *like*’s function here is not that of a discourse marker with retrospective scope, following the quote *he*. Due to an intonation break after *he* and a new intonation starting curve with *like*, I interpret it as preceding and associated with *him up there*. Speaker A can be interpreted as repairing her initial choice “*he*” with a longer quote “*him up there*”.<sup>27</sup> Fleischman and Yaguello (to appear, footnote 10) say that *like* “can signal a rewording, expansion, or justification of a preceding utterance. It introduces a metalinguistic comment by the speaker in his/her own words. In this capacity it overlaps functionally and often commutes with *I mean (...)* and/ or *you know*”. This is exactly *like*’s function in this piece of talk. As the discussion above has shown, *like*’s comparative semantics allow speakers to use it in situations where the repaired item is compared to the material it is repaired with. By using items with comparative semantics, the speaker presents them as having something in common (cf. Meyerhoff and Niedzielski 1995). This way, the speaker saves his face and does not present himself as having made “wrong” lexical choices (cf. Meyerhoff 2002).

This mitigation task on an interpersonal level taps into *like*’s already established hedging function. Obviously, another possible interpretation of this excerpt is that speaker A represents what the speakers had really said at the time, “*he*” and compares this to her own interpretation which is “*him up there*”. As we do not know what the original speech act was, both readings are open. *Like*’s simulative semantics make both interpretations possible. The most important thing to note, though, is that when *like* occurs between the parts of a two-part quote, it is exactly in the slot where we expect a verbum dicendi to occur. The frequency of *like* in such positions, especially in British English but also in US English calls into question the

---

<sup>27</sup> Tagliamonte and Hudson (1999) mention the following case: Here, the first quote consists of indirect speech framed by *said like* followed by a direct quote framed by *like*.

“So my friends said like, to go away, but not very politely ,  
*like* “Fuck off, go away” (Uk/x)”

In their interpretation, *like* is not a full quotative but a false start. Given a certain intonation pattern, this is a possible interpretation. However, I claim that if *like* regularly and repeatedly occurs in the position directly in front of the quote, whatever its function, its co-occurrence can give rise to its reinterpretation as a quotative.



importance of the sequence  $V_{quot}$  **like** “quote” for the genesis of *like*’s quotative functions.

Consequently, *like* can occur alone before a quote which is not a direct continuation of a previous one when it occurs in a context in which the de dicto realm has already been established. Consider example (4.21).

(4.21) BrE - *Discussion on religious and ethnic tolerance.*

- B: so me mum and dad left the choice to make me,  
A: uh huh,  
→ B: I saw people **say like** “what nationality are you all,  
and what [ ((religion and)),  
→ A: [**like**“why shall I let you know what national I am”],

In this example, speaker B creates a hypothetical scenario in which intolerant peoples’ attitudes are represented via what they say. In the line marked with the first arrow, speakers enquire about other peoples’ backgrounds (*what nationality are you all.....*). B quotes those people using the canonical quotative *say* plus *like*. Her interlocutor A joins in and gives an equally hypothetical answer to the question formed by B, “*why shall I let you know what national I am*”. Thus, B collaboratively creates a hypothetical interactive situation with speaker A. This follow-up quote, the answer to the question framed by her interlocutor, is preceded by *like* only. The quotative situation has been established through the use of a canonical quotative; *like* occurs in a bath of speech. According to findings in the field of conversation analysis (Pomerantz 1979, Sacks, Schegloff and Jefferson 1974), an answer to a previously asked question is structurally required, due to its structural position as a second pair part. In (4.21) it need not be explicitly marked as the realm of quotation has already been opened (cf. on a similar note Güldemann 2001:31). Here, *like* occurs in exactly such a position and is framing a second pair part. Also, we can assume more generally that once *like* has repeatedly occurred in the vicinity of quotation, it is at the stage where it has become sufficiently associated with its context to serve as a framing item alone, without a canonical quotative verb.

The hypothesis that *like* grammaticalised independently of collocations with *verba dicendi* is underlined by the frequency with which *like* occurs in quotative situations in slots directly before quotations.<sup>28</sup> Let me give you a few examples:

(4.22) US English - *Billboard hypocrisy*

- A I- I- they- they- they have all these neat phrases,  
B yeah,  
A you know- you know- you know **like** “critically acclaimed”,  
that's what they- when the thing hasn't won any award.  
B oh I see,

In this extract, A makes fun of the hyperbolic descriptions used by the movie industry in order to lure people into cinemas. She represents one of the *neat phrases*, namely *critically acclaimed*, as an example of all the deceptive slogans currently found in cinema advertisements, and frames it with *like*. *Critically acclaimed* is one example and as such an instantiation of the *neat phrases*. Hence, *like*'s simulative semantics plays a large part in its use as an introductory item for reported events. Also, *like* marks the focus of the utterance, the actual rendering of the stereotypical billboard phrase. Furthermore, the fact that the co-occurrence and repetition of several fillers indicates production problems opens the possibility for *like* to be interpreted as a filler. As can be seen in this example, all functions of *like* play into its grammaticalisation as a quotative introductory item, as all of them can hold at the same time. Consider also the next example:

(4.23) US English - *The lilac bush*

- A and I'm looking at him **like** “jeez,  
people would kill to have a lilac bush like this one”,  
B yeah.

The scenario in which example (4.23) is uttered is the following: A is renting a house with a beautiful lilac bush in front. One day her landlord declares his intention to cut the bush down. The ensuing exchange between A and her landlord hosts the excerpt given in (4.23). Speaker A compares the look on her face to a speech act “*jeez people would kill to have a lilac bush like this one*”. We can interpret this

---

<sup>28</sup> This hypothesis does not preclude that *like* and canonical  $V_{\text{quot}}$  cannot occasionally co-occur. It only suggests that this collocation is not the only and decisive one for *like*'s development into a quotative item.

quote as a “representation of what might have been said aloud (if the participant were in a ‘state of talk’)” (Ferrara and Bell 1995:283). A probably didn’t utter *jeez ...* but she felt like uttering it (Romaine and Lange 1991:237). If A had spoken out her thoughts, she would have uttered *jeez ...* but as it happened, she didn’t. Only her face betrayed what she would have said. The following example makes this point even more explicit:

(4.24) US English - *The dog*

A: and it would not have been fun trying to retrieve her but uh this dog,  
Dennis jumped in and got this look on his face **like** “what do I do now” he he,  
and he’s floating down the river then finally discovered that he could swim.

Here, the dog Dennis certainly cannot talk. A presents him in a way that if he could, he would utter the words rendered as a quote.

In the next example, a house is compared to the typical reaction people have when they see it:

(4.25) US English - *The party*

A: just her whole house **is hhh like** “wow” he he,

Obviously, *wow* is not attributed to a speaking house. People who perceive it exclaim *wow* or would do so if they were in a state of talk. We see that in cases like the above, the comparative function and the quotative function overlap: the house is compared (or likened) to people’s reaction at the sight of it. In the same vein, Romaine and Lange (1991) say that once *like* has taken on the role of a (comparative) conjunction, it is only a short step for it to introduce speech or thought. This is possible when the narrating speaker reports the speech or thought as if it was an exemplification or as if it was to be compared to what came before. They give the following example of a checkout woman who is about to put a roll of coins that a customer has just bought back into the cash register drawer. In this case the event of the customer’s standing there is compared/equated to a proposition belonging to the *de dicto* domain “*what is she doing*”

(4.26) US English - *The checkout woman* (quoted from Romaine and Lange 1991: 262)

Why am I putting these in my drawer? These are yours. You're just standing there like, "What is she doing?"

The cashier is animating what the customer could have thought or would say if she had put her thoughts into words. It is also a state of affairs to which the customer's standing is compared. Romaine and Lange point out that there is a metonymic association between comparing a state of affairs and the (constructed) speech/thought event going on in it (cf. also Ochs 1996). *Like*, if used repetitively to compare events with quotations, is on the road towards grammaticalisation. The sheer frequency with which *like* occurs in such contexts in my corpora is highly suggestive for their importance in its reinterpretation. A cursory look through the BNC, collected before 1993 in BrE (notice the first attestation of quotative *like* in this variety is in the COLT (Andersen 1998) collected in 1993) shows that *like* in such de dicto situations occurs in this variety as well. Thus, we can assume that this road to grammaticalisation is not a variety-internal phenomenon idiosyncratic to the US.

The last few paragraphs have given evidence that if *like* occurs within the realm of quotation in the context of *say like "quote"* as well as in cases such as the above, it is independently associated with this context via two routes. This greatly heightens the probability of its grammaticalisation.

I will now discuss another route to *like*'s grammaticalisation which has been almost entirely ignored in the literature to date. In my corpus, one of the most frequent contexts of occurrence of quotative *like* is the sequence *3<sup>rd</sup> pers. neuter pronoun + be+ like*, which occurs almost invariably in the reduced form *it's like*. Tagliamonte and Hudson (1999: 170) mention *it's like* in a footnote "Interestingly, although representing a very small proportion of the data, N=23 (3.4%), the most highly favoured context for *be like* in Britain is with existential *it...*". Let me give an instantiation of *it's like* from my corpus:

(4.27) US English - *Recycling*

B: it- I know I've got a two year old and a four year old and that's, my daughter thinks that any time she sees newspapers bundled up they are being= = recycled.

- A: uh huh,  
 B: **sohhhh it's like** "well, (.)  
 no honey",  
 so that - like so that was one of the reason we went to show her what people did=  
 =with their- their stuff,

Existential *it* plus dummy *be* and *like* make up the quotation frame in this example. *It's like* introduces speech attributed to self. The question is: why did the speaker not use *I'm like*? In the following paragraphs, I will discuss the role of *it's like* in the grammaticalisation of *like* as a quotative. I will explain the patterning of *it's like* in my data and discuss it in the light of the few comments on it in the literature, which are, interestingly, usually found in footnotes.

Consider Romaine and Lange's (1991:254) very telling discussion of an example of *it's like* in the media where they disagree with the punctuation proposed by the author. R&L write of the 1<sup>st</sup> and 3<sup>rd</sup> instance of **LIKE** (which they represent in bold capitals) that they are "indeterminate in usage". Let me give the excerpt and their comment in full. I want to focus on the third instance, preceding *I live here*.

56. **IT'S LIKE** (sic.) you want to come home and it's no left turn, no right turn, go this way, come back that way, and then the cops look at you **LIKE** "where do you think you are going?", **IT'S LIKE**, I live here [Charles Fishman, The Washington Post, 14 Dec, 1985, A3]

"They both follow the pronoun *it* rather than pronouns which indicate a specific speaker, and it may be that these are regarded by the writer or editor as cases where *like* is used for comparison rather than for the introduction of speech. However, we would argue that these still might be thought of as part of the quotation frame, particularly in the third instance where the statement "I live here" seems to be what the speakers might have said or thought. Likewise, in the second instance, the frames *they said* of *they thought* could be substituted for *like* with no substantive change in meaning...."

Romaine and Lange point out that the frame *it's like* seems to be particularly ambiguous in that it can often be paraphrased using quotation frames such as *I said* or by discourse markers *like*, *look*, *I mean* and so on. ... The speaker seems to be using *like* here to describe a more general situation as well as to introduce a quote.

According to Hopper (2000, 2001), the most vital element in the formation of new, creative patterns of grammar is frequency of patterning. The sequence *it+be+like* provides the contexts with the highest frequency where *like* occurs in the

direct vicinity of a quote – 56% of all *like*-tokens in the American data occur with neuter *it* (for British English see earlier). I argue that this fact is key to *like*'s grammaticalisation as a quotative, at least in US English.<sup>29</sup> Note Fleischman and Yaguello's (to appear) comment (again in a footnote) "The clause initial *it's / it was like* occurs with sufficient frequency in our American data as to raise the question of *it's* being a construction in its own right." Furthermore, the sequence *it's like* is the principal site where we have ambiguity between two structures, one of which is the comparative, core meaning of *like* (cf. Romaine and Lange's discussion above), the other of which is quotation. In such an environment, due to the process of reanalysis, *it's like* can acquire the function of a quotative frame. Let me show this with an example:

(4.28) BrE - *Devoutness and worldly success*

01 T: cause like I didn't go to the church and stuff and,  
 02 and more or less blamed to god,  
 03 like- it's **like** "God was responsible for her passing her A levels and everything"=  
 04 = you know,  
 05 and I'm **thinking** "that's a bit stupid because it's-it's what's in your head that counts,  
 06 and exercising your brain".

Let me discuss line 03 first. Speaker T, telling her interlocutor of her feelings towards her religious cousin, contrasts herself, a non-church goer, with her cousin's weekly service attendance. She then discusses the possible link between her cousin's success at school and her devoutness. *It's like* introduces one possible interpretation for her cousin's success at school: that her cousin was helped by divine intervention, which she later dismisses as *that's a bit stupid...* *It's like* thus precedes a description of a situation from T's subjective point of view and can be glossed as *It (=the situation) was like, at least I perceived it that way*. Line 05 is the point in the transcript, where T refutes her earlier interpretation, presenting a rational argument *it's what's in your head that counts*. She contrasts this perspective with her previous proposal and marks its subjective character as well as the fact that it is her opinion by framing it as inner thought with *and I'm thinking*. Again, T presents a situation

---

<sup>29</sup> Consider in this vein Winter (2002:13,20) who shows that innovative *be like* in Australian English is largely confined to third person singular use. She claims that this fact suggests a very different introduction pattern for the innovation from that documented for North America, Canada, and British English. Winter only reports 8% *like+existential* it.

through her subjective interpretation. *I'm thinking* has the same functions as *it's like*: to signal an interpretative description of a situation as seen through the eyes of a protagonist and to present it in the form of a quote. In cases of formal ambiguity of *it's like* as in the example above, the construction of arguments as contrastive stances in a parallel structure -first the religious position, then the rational one- gives weight to the possibility of interpreting *it's like* as a quotative.

Consider also the following example which hosts a repair from *I think* to *it's like*.

(4.29) BrE - *College talk*

Y: ahh you do the MAB one,  
X: uh uh,  
and I think- it's **just like** eh "that gives us an advantage now",  
because one MAB is based on your own,  
you're doing geoscience.  
and not having done geology GCSE,

We witness a case where there is a transition from *I think* to *it's like*. Speaker X, starts by representing the situation as inner monologue, framed by *I think*. She then repairs and reverts to using the less subjective form *it's like*. Both are modalising utterances: *I think* portrays the description as perceived from the point of view of the speaker, *it's like* hedges on the basis of subjectivity as well. We notice that the construction *it's like* is on the brink between a parenthetical quotative (Schneider 2002) and a real quotative.

In the next sequence, *it's like* clearly takes on the function of a situational quotative as described in Chapter 3.

(4.30) US English - *Politics and ethics*

B: but it's still you know,  
A: I think,  
→ B: I guess I keep **that one part of me that's like** "well hhhhh,  
I don't know I want to trust them",  
→ and I- **it's like** "I can't".  
and maybe it's because you know we've had our own propaganda over the years,

In this example, speaker B is describing his feelings “*well hhhhhh I don’t know I want to trust them*” framed by *that one part of me that’s like*. The item *like* stands between and compares the speaker’s feeling to a quote. Again, we have the situation in which a quote would have been uttered if the speaker it is attributed to had been in a state of talk. B is representing inner talk.

In the line with the second arrow speaker B states *I can’t*. This is first of all a description of the situation at the time that B gives to her interlocutor. B cannot trust *them* because of all the propaganda that had been disseminated previously. On the other hand, *I can’t* can be interpreted as a continuation of the previous quote this time framed by *it’s like*. The whole original line of thought, now broken up in two quotes, would thus run “*well hhh I don’t know I want to trust them (but/and) I can’t*”. The sequence *it’s like* frames linguistic material that is at the same time a description of a situation and a quote of the situational category. Note in this respect Schourup (1985:59-61), who feels that *it’s like* has the same evincive meaning as *like* alone. He interprets existential *it* as referring to what the speaker has in mind to express. Hence, as has already been established, *like* has a similarity-marking function. In collocation with existential *it*, forming the sequence *it’s like*, it is used to describe how speakers perceive situations. When descriptions take the form of previously occurring, reported speech, the sequence *it’s like* can take on the function of a quote introductory marker. Note that the comparative function of *like* is an important precondition for this transfer. The functional extension lies in the metonymical association of a situation with a quote.

Ferrara and Bell (1995) state that one of the reasons for *like*’s triumphal entry into the quotative system is its flexibility. As complex choices have to be made as to how to represent spoken discourse or inner speech, as well as gestures, sound, and voice effects, a one-size-fits-all quotative simplifies planning. By choosing the form *it’s like*, speakers can cut calculation time even more because it can occur with all grammatical persons, tenses and aspects. Functioning like a wild-card (cf. Chapter 3) and being able to frame all re-enactments in all pragmatic and formal contexts is undeniably one of the main reasons for the enormous momentum and speed with which *it’s like* has been introduced into the quotative system.



Singler (2001) says that *it's like* might have been an important step in the grammaticalisation of *like* but does not elaborate how: “arguably *it's like* evolved from *it's as if X said*, where X ordinarily has an indefinite referent (...) but can have a specific referent (...). The construction may have played a crucial role in *like*'s grammaticalisation as a quotative” (Singler 2001:261). I will now sketch a possible alternative scenario for *like*'s genesis as a quotative. The frequency distribution of *it's like* “quote” where *it's like* has clearly quotative (amongst other) functions in my US English corpus (1988-1992) shows that 56% of all *like*-quotes are framed by existential *it*.<sup>30</sup> This is an astonishing number given how little it has been discussed in the literature. Blyth, Recktenwald and Wang (1990:221) found that in their corpus, *like* is rarely used in the third person. However, they do not give any indication whether they have looked at the possibility of *like* occurring with existential *it*. Ferrara and Bell's (1995:278) findings indicate that *it's like* amounts for 7% of instances in their two earlier corpora (1990 and 1992) and only for 2% in their latest corpus (1994). Thus, compared to my data (1988-1992), the number of 3<sup>rd</sup> person inanimate pronoun with *like* found by Ferrara and Bell is substantially lower in the first corpus (1990-2) and even lower (1994) in the second. I suggest that the discrepancy might lie in the fact that earlier findings might have overlooked cases of existential *it*, simply because they were looking at the quotative frame through the lens ‘personal author + quotative’, hence failing to notice existential *it*. Such an interpretation of not all but most of the early literature is underlined by how seldom existential *it+like* is pointed out and by the fact that, if it is mentioned, it usually occurs in footnotes (cf. Tagliamonte and Hudson 1999). In sum, the possibility of *like*'s grammaticalisation via the sequence *it's like* “quote” in situations where descriptions and quotes overlap is one that has been neglected in previous research.

As has been confirmed in publications on *like* over and over again (Dailey-O'Cain 2000, Ferrara and Bell 1995, Johnson 2001, Tagliamonte and Hudson 1999), *like* has generalized into all pronominal environments. It is usually assumed to come from first person sg.; no mention is made of *it*. Data supporting the claims about

---

<sup>30</sup> Consider also Chapter 5. Those tokens are the ones which have been decided on as definitely quotative by the author. In case of doubt, I have consulted with M. Meyerhoff and at least one other native speaker. Only if all instances agreed on the interpretation of an instance as having a quotative function, was this token kept in the analysis.

generalization could also be seen as evidence of paradigmatic spread (Hopper and Traugott 1993) of quotative *like*, arising in its quotative function in a certain environment (existential *it*) and expanding into others. Hence, it seems that *it's like* is a strong contestant for the environment in which *like* could have grammaticalised to a quotative. While claims in this chapter can only be of a suggestive nature, it seems that the diachronic association of *like* with existential *it* provides a fascinating ground for future research.

A comparative glance at the BNC (data 1991-1993) reveals quite a number of instances of *it's like* with somehow ambiguous status which cannot yet be interpreted as situational quotes but which clearly have some sort of quotative notion implied. At the time the Derby and Newcastle data was collected (1994/5), 3<sup>rd</sup> person sg. neuter accounted for 34% of all quotes (after 40% for 1<sup>st</sup> person). Thus, while we do not know how important the role of existential *it* is in *like*'s grammaticalisation across both varieties of English, we can safely say that *it's like* is one of the constructions via which *like* has acquired its quotative introductory function. The question whether this construction represents the incipient grammatical form in general or one which is particular to the American context also remains a question for further research.

It looks as if in BrE, where *like* can occur after the item it has scope over, it is predominantly the in-between quotes position which made it grammaticalise in this variety as well as the association with non-canonical *go*, whereas in the US, the *it's like* construction led to its use as a quotative. But it is by no means the case that we have to postulate one grammaticalisation model for British English (the sequence “quote” *like* “quote” leading to *like*'s grammaticalisation as a quotative) and a completely different model for US English (*it's like* “quote” being the precursor structure to its quotative function). Instead of trying to pin down the grammaticalisation to either of the two scenarios, and blurring the commonalities, I suggest that the most intuitively satisfying hypothesis is that the triggering factor lies in *like*'s association with quoted material in the largest sense. The above discussion has shown that we have to accept the fact that we have to work with fuzzy categories, which are hard to pin down and ambiguous on multiple accounts. Thus, I claim that

all the contexts mentioned above have contributed to the association of *like* with quotation and all environments in which a reanalysis is possible need to be taken into account.

#### 4.3.6. Discussion

The above paragraphs have shown that all uses of *like* still have a semantic trait of comparison / approximation. Whether it is *like*'s hedging function on a pragmatic or semantic level, or its filler or focus function, all can be conceptualised as direct extensions from one common core of similarity. I have explicated in great depth how these functional outgrowths have contributed to the development of *like*'s quotative function. Again, the metaphorical and metonymical extensions which have led to this new function are motivated by the notion of similarity. This is in accordance with Hopper's (1991) principle of the persistence of meaning: *Like*'s semantic core meaning is still present in all the derived uses which are linked to each other in various ways. *Like*'s older uses still persist in the language and continue to play a role in its further development. The development is additive rather than suppletive. Behind the overlappings and ambiguities that result from *like*'s multifunctionality lies an interrelated net of semantic-pragmatic links around the core semantic property.

Schourup (1982a:32) gives data from languages such as Sierra Miwok, Lahu, and Raluana in which *like*-lexemes function simultaneously as comparative items and as, what he calls, 'evincive'. This seems to be supportive cross-linguistic evidence for a grammaticalisation path linking comparative items and hedges, as traced for Standard English by Romaine and Lange (1991), and for *olsem* in Bislama by Meyerhoff and Niedzielski (1998, 1995).

Hence, my results are sustained by much cross-linguistic evidence (cf. Güldemann 2001, Meyerhoff and Niedzielski 1998, 1995, Schourup 1982a: 32). As is evidenced in Table 4.1 below, multiple discourse functions within the same related field are quite widespread among unrelated languages.

**Table 4.1:**<sup>31</sup> Cross-linguistic distribution of *like*-lexemes

	approximative	comparison	rep. speech	rep. thought	focus	hedge	filler
English <i>like</i>	x	x	x	x	x	x	x
Bislama <i>olsem</i>	x	x	x	x	x	x	?
Japanese <i>nanka</i>	x	x	x	x	x	x	x
French <i>genre</i>	x	x	x	x	x	x	?
Chinese <i>xiang</i>	x	x	x	x	x	x	x
Buang	?	(na) be	(na) be	(na) be	?	be	?
Finnish <i>niinku(in)</i>	x	x	x	x	x	x	x
Thai <i>bæ:p</i>	x	x	x	x	x	x	x

This chart is to be read as follows: question marks denote cases where no data was available. Brackets mark optional constituents. The chart gives cross-linguistic evidence for a close semantic link between the notion of approximation, hedging, focusing, and introducing speech and thought.<sup>32</sup> These results are highly suggestive of a universal cognitive / perceptually salient pragmatic-semantic field lying behind such coherent findings (cf. Fleischman and Yaguello's (to appear) proposal that a theory of pragmatic universals is at order). If source-items in two or more different languages, especially languages that are unrelated, such as English and Thai, follow parallel paths of development without any evident contact, this supports the assumption that there is a general link between the notion of similarity and its derived functions. Hence, Traugott (1995) asks if one can make cross-linguistic generalizations about the development of discourse particles both in terms of their semantic sources and their semantic-pragmatic paths. Consider also Mosegaard-Hansen (1998:85) who points out that, synchronically and diachronically, discourse markers can be traced back to a number of related uses. I claim that such evidence supports the claim that a model, such as the radial structure model, which relies on a semantic field as an explanatory parameter, is best able to cope with the linguistic reality as discussed in this chapter.

It has often been claimed in the literature that discourse markers are highly language specific items (Brinton 1996, Mosegaard-Hansen 1995), hard to translate, and only understandable within their specific linguistic system. But Table 4.1 shows

<sup>31</sup> For the information in this chart I am indebted to the following people: Mie Hiramoto, Kazumi Yoshihara, Aaron Tsang, Sumittra Suraratdecha, Preena Kangkun, Gillian Sankoff, Hannele Buffy Nicolson, and Miriam Meyerhoff.

that similar lexical items in unrelated languages are generalized to serve the same discourse functions. Also, studies like the ones by Meyerhoff and Niedzielski (1998, 1995), Fleischman and Yaguello (to appear), and this present one underline this claim and show that there are cross-linguistic parallels between the source items and the outcome of such semantic-pragmatic developments. This means that it may be possible to relax the claims about specificity and untranslatability of discourse particles.

The proposed radial structure model shows how superficially messy facts can be linked in an orderly way. The diverse functions that *like* has assumed synchronically are motivated by this model - they cannot be predicted but they are explained. The analysis presented in the preceding paragraphs has shown that the synchronic functions of linguistic items cannot be fully explained by a chain model but are best understood as a net of relations. Given the overwhelming evidence for interwoven functions and the numerous cases of case-wise overlap and ambiguity presented in this section, it seems naive to try to force functional extensions into separate unilinear chains. The situation with *like* is one of ambiguity not only between the core item and the functional outgrowths, but also between all functional extensions. As I have argued throughout this chapter, this ambiguity is constitutive of how different functions are understood and conceptualised. If we want to render justice to what is really going on, we will have to try to free ourselves from simplicistic goals, no matter how “neat” they might seem. Unfortunately, with respect to *like*, things are complicated.

As linguists, we not only have to accept messy facts but we also need to represent them the way they are and give comprehensive accounts for the underlying factors that might have caused the status quo. What we need in order to account for *like*'s patterning is a model that is subtle enough to be able to cope with multidimensional and multifunctional entities. I claim that a radial field model is a strong contender for such a job (cf. the depiction earlier on). It comprises numerous ‘paths’ which I prefer to call functional extensions in order to avoid possible

---

<sup>32</sup> For an extensive cross-linguistic study on the link between items meaning *like*, quotative verbs, and various other domains see Gldemann (2001).

infusions of directionality. The different functional extensions can be represented as movements in semantic/pragmatic space. Consider Figure 4.4:

**Figure 4.4: The functional extensions underlying *like*'s semantic core field**

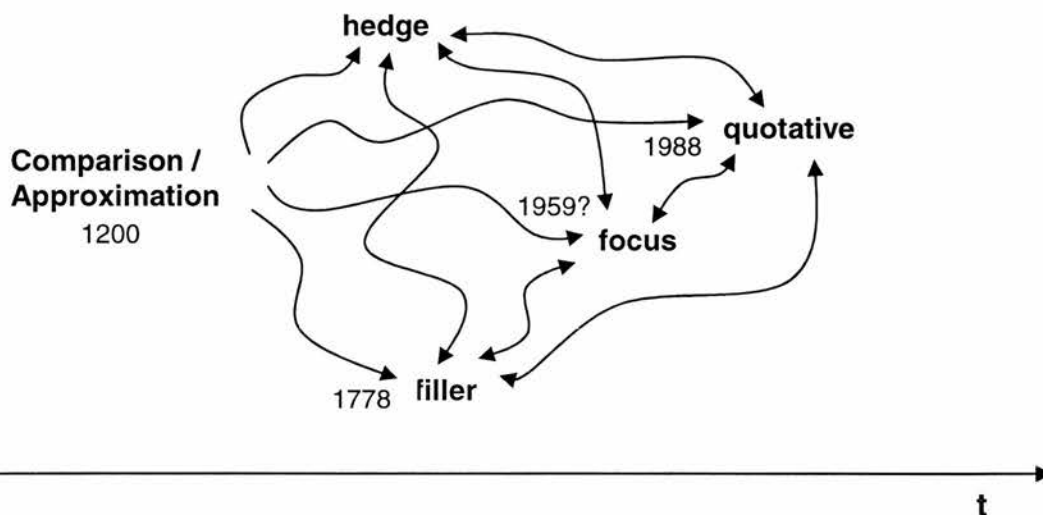


Figure 4.4 depicts the semantic field of *like* where functions are radiating out from the original core meaning of approximation / comparison. The model, which is obviously only an impoverished representation of the multi-dimensional semantic field, is presented so that it has a temporal dimension to it, with older functions grouped more to the left and newer ones depicted as appearing on the right hand side. In the ensuing interwoven field, dates indicate the first mention of the functional extension. It is reasonable to assume that the hedging function arose quite early. While we do not have any concrete evidence for its first attestation in the OED (hence, the schema contains no date for the entry *hedge*), I argue that it is reasonable to suppose that an approximative item can be used with hedging function from its very beginning (see in this vein Fleischman and Yaguello: to appear). Except for the quotative function, which Butters first noted with dummy *be* in 1982, the dates are taken from the OED. Concerning this function, the OED gives an instance of *like* as a quotative complementiser in the year 1970 under the heading “also in less analyzable constructions” which is as follows: “1970 *Time* 31 Aug. 19 Afterward, a girl came up to me and said, ‘You kinda look interested in this; did you know there are civil rights for women?’ And I *thought like* ‘wow, this is for me’”(italics mine). If we accept *like*'s occurrence as a quotative complementiser as a precursor to its full quotative function, we can conceptualise this instance as situated within the realm of

quotation. But it is still removed from the full quotative function and situated somewhere in-between in semantic space along the intersecting dimensions hedge, focus, and approximation and quotation. As those categories are fuzzy and have intersecting areas, ambiguous cases - as attested numerously in real occurring data - can be positioned somewhere in-between. Singular instances of *like* are located in the scalarly aligned n-dimensions of the field, closer or less close to the prototypical functional categories.

#### 4.3.7. Conclusion

The above discussion has shown that *like*'s grammaticalisation does not proceed unidimensionally. Its development does not progress step-by-step, but within a synsemantic<sup>33</sup> field of mutually overlapping and reciprocally reinforcing functions. Also, the boundaries of *like*'s synchronic functions are hard to determine and are heavily dependent on the context (prosody, lexicon, syntax etc.) as well as on speaker intention and hearer reception, leading to case-wise ambiguity and overlapping functions the interpretation of which must necessarily be subjective. This is underlined by Bredin's (1998:75) statement about the "polysignificance of comparison markers".

Building on those findings, I have shown that a unidirectional grammaticalisation model cannot render the synchronic multifunctionality of *like*. Rather, I have demonstrated that a semantic core structure is able to capture the linguistic reality by giving enough plasticity and allowing for ambiguity on multiple dimensions. This is underlined by Ferrara and Bell who claim that "the present pragmatic plasticity of *be+like* - that is, its ability to introduce constructed dialogue, thought, and quotable gestures of self and others - arises from systematic variation and expansion from a core paradigmatic case" (1995:281-2).

Similarly, Romaine and Lange (1991:245, 262) say that "the meaning of 'approximative' and 'similarity' as well as the focus function have contributed to

---

<sup>33</sup> Synsemantic in the sense of having (a) semantic trait(s) in common.

both the discourse uses of *like*, and that these are natural outgrowths of existing uses and functions of *like*... [Hence,] a simple linear model of grammaticalisation is inadequate to account for these developments. What may emerge from grammaticalisation is a network of related meanings of an item” (cf. Claudi and Heine 1986:313). They have tried to render justice to this fact by using a branching model. In this chapter, I go one step further. I have argued that the radial structure model first introduced by Lakoff (1987), allows for one core meaning with metaphorical and metonymic extensions more or less closely linked amongst themselves while retaining their original core function. I have discussed this model in the light of the distribution of *like* in my corpus and have shown that it is able to motivate and conceptualise all instances of *like*. In the next section, I will show that a radial structure model can also explain the status quo of *go*'s functions.



#### 4.4. The *go*-Scenario

This section investigates the grammaticalisation of *go*, another highly multifunctional item that can be used to introduce quotation. Several of its functions, such as its use as a future marker, have been examined frequently in the literature (Bybee and Pagliuca 1987, Bybee, Pagliuca and Perkins 1991, Bybee, Perkins and Pagliuca 1994, Danchev and Kytö 1994, 2002, Pérez 1990, Poplack and Tagliamonte 1999, Royster and Steadman 1923/1968, Wekker 1976). But no study has ever given a coherent account of how *go*'s quotative function can be linked up with its other synchronic functions nor has there been any data-driven study which attempts to motivate the ambiguity and multifunctionality which individual cases of *go* may display in real occurring speech. This section is an attempt to give an account of how the quotative function of *go* interacts and is informed by the rest of *go*'s functions. Investigating *go* from a synchronic perspective, I will propose a solution as to how these functions can be conceptualized as motivated and interrelated. As with *like*, I will argue that *go*'s synchronic meanings are functional extensions from one common core and that they are linked via diverse channels.

The OED's most general definition for *go* is the following:

- “an intransitive verb of motion, serving as the most general expression  
(I) of a movement viewed without regard to its point of departure  
(II) for a movement away from the speaker, or from the point at which he mentally places himself  
(III) to or towards a place which is neither in fact nor in thought that occupied by the speaker”.

It further adds that *go* is to be understood as in contrast to *come* and that it later acquired its semantic application to walking in order to distinguish this form of movement from other modes of progression.

A contemporary dictionary entry for *go* yields a number of seemingly unrelated functions. But looking at a corpus of talk-in-interaction, the only functions that occur with great frequency in contemporary spoken English can be grouped under the following four headings:<sup>33</sup>

---

<sup>33</sup> This list excludes phrasal and idiomatic uses, as in (i) she goes to (i.e. attends) Edinburgh University; (ii) I've gone and burned the toast; (iii) Blair will go to the country (i.e. hold an election) next month. Nigel Love (p.c.) rightly pointed out that my proposed model can only claim to be a full account of all

- movement as in (4.31) *She went by train.*<sup>34</sup>  
this category also subsumes more metaphorical cases of movement as in  
(4.32) *The prize went to a sophomore.*
- intention: (4.33) *I didn't go to do it.*
- introducing quotes containing linguistic material and/or sound and voice effects  
and/or gestures: (4.34) *The gun went "bang".*
- future: (4.35) *What is going to happen to us?*<sup>35</sup>

In the spirit of Lichtenberk's (1991) query as to whether meanings have structure or whether they are to be thought of as unanalyzed wholes, this chapter will explore all of *go*'s functions (found above the level of significance in 2 corpora of talk-in-interaction) and will try to represent them in one coherently structured model. It will become evident that Lakoff's (1987) concept of a radial structure is again a very useful tool to conceptualize *go*'s linguistic reality. Through the mapping of diachronic developments in synchronic semantic space, the ensuing model serves both as a description of the synchronic status quo of the development of the lexeme *go*, as well as a motivation for the development of further functions.

#### 4.4.1. The radial structure

The semantic-pragmatic structure of *go* is quite complex, a finding which directly taps into Güldemann's (2001) claim that "quotative verbs are notoriously polyfunctional outside the quotative frame". Bybee, Pagliuca, and Perkins (1991:31), however, argue that movement-derived futures do not provide the richness of semantic nuances that modality-derived futures do. While this study is not in a position to compare *go*'s functional wealth with other future markers, it is worth noting that *go*, as a movement-derived future, does display an array of other functions, which is also contra Bybee, Perkins, and Pagliuca's (1994) findings that movement verb futures typically have only one other use or no other use.

In this chapter, I will argue that all of *go*'s functions (movement, future, intention, and quotation) are 'alive and well', co-exist and even overlap. This again

---

of *go*'s uses once I have been able to incorporate these functions. As it stands, this section is to be read as an account of all of *go*'s tokens that occurred with any frequency in a multi-million word data-base of two varieties of spoken English.

<sup>34</sup> This and the following examples have been adapted from Merriam-Webster's 2001 online dictionary.

<sup>35</sup> From Collins (1998).

questions the appropriateness of a chain as a synchronically explanatory parameter. Within my model, I assume the notions **GOAL** and **SOURCE** as the basic semantic components of the synsemantic field around *go*. This is both due to their early appearance (first attestation in OED as **GOAL** 971, first attestation in OED as **SOURCE** 1000) and because they are the persistent semantic primitives which are ‘shared’ within the developmental schema of extended meanings. Note that *go* can also have the sense of ‘unspecified movement’, as in *They went shopping*. Its first record in this function (attestation in OED 825) slightly predates the **SOURCE** and **GOAL** meaning of *go*. Even though a slightly earlier date of occurrence (by 150 years) does not necessarily point to an earlier use, especially in cases of patchy data-records from the earliest periods of English, it nevertheless questions the status of allative and telic *go* as a diachronic centre of the schema. This question certainly deserves further study but I will not pursue it any further here as my focus is the conceptualization of *go*’s synchronic uses in one coherent model.<sup>36</sup>

Earlier accounts of grammaticalisation maintained that while meaning gets lost or at least eroded, the newer functions usually retain traces of their original meaning (Hopper 1991). Heine, Claudi and Hünnemeyer (1991:156) have called this phenomenon “isolating abstraction”: the separation and possible further development of a particular property or feature which is not necessarily the core or nucleus characteristic of the source concept. I claim that it does not really matter whether the features that give rise to a functional extension are at the semantic core of the original source concept. What matters is that the feature(s) of the source concept is/are at the core of the new radial structure. I will provide evidence for this claim throughout the remainder of this chapter.

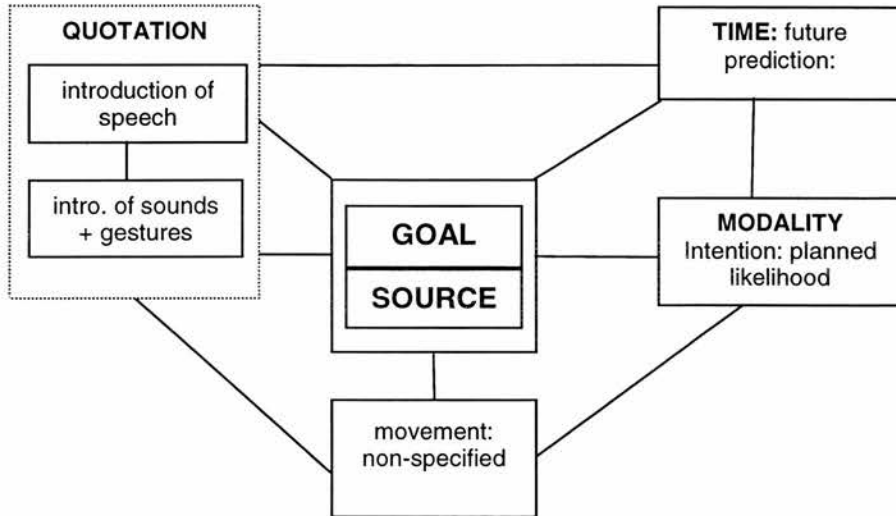
It will become evident that the core notions **SOURCE** and **GOAL** give rise to various meanings which still contain the core semantic properties to a greater or lesser degree. This claim follows the grammaticalisation literature in that it considers the original meaning(s) of the source structure as determining the extended uses the item(s) will have later in their history (cf. Bybee and Pagliuca 1987, Givon 1973).

---

<sup>36</sup> Note also that as Therese Lindstrom (p.c.) has pointed out, there also could have been a split to unspecified and specified movement in the history of *go*’s development with the latter meaning becoming the centre of the grammaticalisation schema as depicted here.

This study shows that original semantic-pragmatic traits are the centrepiece of the synchronic structure. They are not prone to being lost at all. Consider now the radial structure model:

**Figure 4.5: The radial structure model of *go***



The above radial structure outlines *go*'s non-suppletive grammaticalisation into multi-layered meanings via functional extensions. Those offshoots are based on processes that have been identified as the driving forces of functional extensions: metaphor, metonymy, and conversational implicatures. In the following paragraphs, I will discuss the meaning extensions of *go* starting from its core. The discussion will show that synchronically, a single grammatical chain joining uses that are more and more remote from one original cannot cope with the situation as we find it in real occurring speech. In my data, *go*'s co-occurring functions are highly ambiguous and overlap with one another, building up a multidimensional field. Note also that this radial structure, contrary to the model proposed for *like*, splits up the function quotation into two sub-functions, introduction of speech and introduction of sounds and gestures. The reason for the split is that the different underlying parameters have given rise to the functional extensions, as will be explained later in this section.

Güldemann (2001:38) shows that in some languages a verb which is capable of marking mimesis may mean 'make, do, act', such as Greek *kano* reported by Tannen

(1986:317) or German *machen* (Golato 2000, 2003). He proposes that a meaning like ‘do’ may actually be a descriptive confession that a monosemous account is impossible because of the extreme semantic-functional versatility of the item. He maintains that a verb in a quotation frame does not necessarily have to mean ‘say’ and proposes, given the cross-linguistic evidence, several source items for the class of quotatives that he provisionally calls performance or mimesis verbs. These verbs have no de dicto meaning outside of the quotative frame and are usually generic performance verbs. More specifically, Güldemann (2001) refers to the example of the Dongala verb *án* ‘go, become’, which, amongst others, has telic and non-telic locative functions and can be used as a quote introductory item (cf. also Ambruster 1960). If more cross-linguistic evidence for the general existence of the cognitively salient pragmatic-semantic link between these functions can be found, one could argue for a general semantic field underlying *go*-verbs. To date, it remains to be seen if the synsemantic field of *go* with the above functions is a largely idiosyncratic phenomenon.

#### 4.4.2. The Temporal Function

Dahl’s (1985) cross-linguistic study provides a number of common developmental types of futures.<sup>37</sup> It has given rise to the hypothesis that all futures develop from an identical, reduced set of lexical items and assume a roughly parallel course of development (cf. Bybee and Pagliuca 1987 for typical channels of future-marking morphemes). Movement verbs are among the most frequent sources of futures and are most frequently periphrastic with transparent etymology (Bybee, Perkins, and Pagliuca 1994, Güldemann 2001, Ultan 1978). This could be seen as evidence that they are relatively recent developments<sup>38</sup>. In the case of English *go*, there

---

<sup>37</sup> For the discussion of whether *will*, *going to* and other future markers have the same semantics and for an overview over the semantic shadings involved see Fleischman (1982), Nicolle (1997), Quirk et al (1985), Royster and Steadman (1923/1968), Ver (1993) and many others. I will align myself with Poplack and Tagliamonte (2000) who claim that it is impossible to establish an objective means to identify or even measure highly subjective and contextualized speaker intention and listener inference of semi-equivalent variants (see also Traugott 1972, Visser 1970).

<sup>38</sup> Cf. Fleischman (1982) on the cyclic development of the future markers in Romance Languages where the Latin synthetic form has been replaced by the periphrastic construction infinitive + *habere*. According to her, this development is being repeated with the emergence of the *going to* future in French and Spanish.

is reason to believe that it is the entire construction *be going to GOAL* and not only the movement verb, which is the source of the future function (Hopper and Traugott 2003, Perez 1990, Poplack and Tagliamonte 2000). Cross-linguistically, Ultan (1978) observes that future markers are more likely to be periphrastic than other tense and aspect markers and that their etymology is often still observable whereas in other affixes it is often obscure. This is the case with the future marking function of *go* and suggests a recent development for this function, a claim which is underlined by findings from Danchev and Kytö (2000) who have shown that the first instance of *go* with future connotations in the Helsinki corpus dates from 1438 only.

Bybee (1988) has convincingly argued that semantically unspecified movement is not sufficient in itself to give rise to future meaning. The notion of a **GOAL** is an important element in the grammaticalisation of a movement verb to a future marker. *Go* in its telic meaning implies that the agent's point of view is situated at the moment of the speech act and any newly occurring event is seen as a step forwards in space as well as in time. The origo of the deictic reference is moving away from the spatio-temporal here and now towards a locally displaced event in a prospective future. This is concurrent with the metaphor "future is forward and past is backward" (Fillmore 1997) and the underlying general metaphor "time is space". The following example shows this overlapping property between future and the goal of a movement as an inchoative, telic entity.

(4.36) adapted from SWB (US English)

A: I am going to see a real good movie tonight.

Let us assume that speaker A describes himself as physically moving at a later time to achieve event <e> *see a real good movie* (the context makes clear that A is planning to leave the house for this movie and is not referring to a rental film). The spatial precondition for <e> is state X, being in the cinema, a place distant from where A is in the deictic now at the time of uttering (4.36). Such an event <e> at state X is placed in the future and implies a temporal move as well as a spatial one. The construction could be considered as an elliptical form of *I am going (to the cinema in order) to see a real good movie tonight*. The event, *to see a good movie tonight* can, via metonymic transfer, be interpreted as a temporal achievement. Hence, if the spatial **GOAL** is

encoded as an infinitival phrase, as above, is can be read as a temporal destination. Today, future *go* most frequently occurs with such infinitival complements but Perez (1990) reports its incipient grammaticalisation to have occurred with NP complements such as the example given in Poplack and Tagliamonte (2000), *I am going to Nova Scotia*.

Also, a purposive situation needs to hold in order for the semantic transfer between a movement verb and a prediction future to take place (cf. Bybee 1988, Bybee and Pagliuca 1987, Fleischmann 1982, Givon 1973, Ultan 1978, 1972 for the link between imperfective and purposive). When speaker A produces an utterance like (4.36), he is not simply announcing movement towards a spatially distant place. Going somewhere to do something also involves the speaker's intention to bring about this event, why else would one be moving towards it (see this section)?

Hopper and Traugott (2003) point out that originally, the imperfective aspect was an important part for goal-directed movement to be interpreted temporally: simultaneity of prediction of later event <e> and action in order to achieve it (cf. also Poplack and Tagliamonte 2000). If the agent A is already on his path to the movie theatre and the movement is in progress at the time he utters (4.36) the target of the agent of *to be going to* can be interpreted as a purposive future.<sup>39</sup> Note that the overlapping semantics of moving forward in space and travel ahead in time can only be co-present in a scenario where both hold simultaneously. This scenario is when a continuous goal-directed action is intentionally executed. An utterance like *I'm going to get my jacket* typically uttered when the person is (about to be) on his/her way to the current location of the jacket is a representative example of such a co-presence of travel in space and in time. Similarly, a speaker's movement is motivated by their desire to achieve event <e>, *be at the movies* and this movement is at the same time in the course of happening. We have simultaneity of prediction of a later event <e> and action in order to achieve it. Due to a generalization of the form, the imperfective-

---

<sup>39</sup> Note that Traugott and Heine (1991) point out that if we compare the cross-linguistic grammaticalisation of *go*-verbs to future markers, we notice that such a trajectory does not necessarily rely on an overt purposive construction (such as *to*-constructions in English) or imperfective marking. However, even if such components are not explicitly given in the co-text, we might still assume that they can be inferred from the context.

requirement, of course, is not a constraint for the use of *go* in its temporal function any more now.

In every day discourse, as Traugott and Heine (1991:14) point out, the future marker *go* mostly blends with the auxiliary as well as with the proposition *to* into *gonna*. Heine and Reh (1984:25) define the phenomenon of fusion as the disappearance of the boundary separating two (or more) morphemes “these morphemes thus being reduced to one phonological unit”. If we accept Givon’s (1990:826) cognitive form-function parallel, which he refers to as “diagrammatic iconicity” and which maintains that the more semantically or pragmatically integrated events are, the more the linguistic material expressing them will be grammatically integrated, we can assume that in the case of *go* in its future reading, the event of movement and the intended goal are morphologically fused into one marker, *gonna*.

The question of whether the ablative element (=SOURCE) is necessary for the grammaticalisation of future is less clear and has not been addressed much in the literature. According to some researchers it is important (Bybee, Perkins and Pagliuca 1994, Hopper and Traugott 1993, Fillmore 1997) but need not be explicitly marked. The entity which *goes* necessarily has a deictic origo which is (at least implicitly) defined by temporal, spatial and personal parameters. As stepping forward in time (and space) implies movement *from a certain reference point in the here and now onwards*, there seems to be a semantic overlap between the concept of a SOURCE and future meaning. Thus the notion SOURCE is at least implicitly present when a telic movement verb such as *go* grammaticalises into a future marker. I therefore follow the literature and postulate a link between the future function and the semantic trait SOURCE in the radial structure for *go* without claiming that this trait is a necessary condition for *go*’s development into a future marker.

The analogical extension of temporal *go* to contexts that have nothing to do with movement has also forced a more general interpretation onto *go* as a future marker. As the semantic feature of spatial movement has been lost, temporal *go* has reduced selection restrictions in present-day English (cf. Bybee 1988, Danchev and Kytö 1994, Perez 1990). Historically, spatial goal-directed *go* is the basis for the future marking



function of *go*. Synchronically, though, *go* has more than one potential meaning, that is, we have a case of polysemy/semantic layering (Hopper 1991, Torres Cacoullos 2001).

“... we lose the sense of physical motion (together with all its likely background inferences). We gain, however, a new meaning of future prediction or intention – together with its likely background inferences. We thus cannot be said to merely have “lost” meaning, we have, rather, exchanged the embedding of this image schema...” (Sweetser 1988:392)

Consequently, speaker A in (4.36) need not be on the way to the cinema at the time of his/her speech act just as much as the speaker in Poplack and Tagliamonte’s example does not need to be on the way to Nova Scotia. In fact, due to the time adverbial *tonight*, the most probable interpretation is that speaker A is not. In this vein, consider the following example:

(4.37) personal example (BrE)

She is going to become a massage therapist.

If (4.37) is said in a situation where the agent is already at the place of professional or academic training, this assertion does not involve any physical movement component. While temporal movement without spatial movement is a frequent scenario, it is quite hard to imagine spatial movement without temporal movement (Einstein 1938). Hence, (4.37) does entail metaphorical movement towards a career **GOAL**: becoming a massage therapist. By the time (4.37) was uttered (in 2002), *go* had already sufficiently grammaticalised to be used as an expression of future reference which does not involve spatial displacement. In this case, it overlaps with the intention function use as will be discussed in the next section.

Full grammaticalisation to future has taken place when *go* can be used in cases where the agent and/or the event are incompatible with the concept of motion, i.e. cases like *I am going to regret this*. Due to its generalization as a future marker, *go* can now have future marking function in cases where the agent’s temporal goal is clearly not physical movement. Consider the following utterance *o.k. so I am going to stay put then* (<http://www.writerspace.com/chat/chat082601.html>, 26.06.03)

Here, the event *to stay put* is incompatible with physical movement. *Go* as a temporal marker has extended its use to the future marking of static events.

Furthermore, all agents in the previous examples were human subjects which are endowed with wilful behaviour. The realization of the **GOAL**, event <e> may thus be determined and intentionally executed by the agent deciding on and moving to perform the action in the event clause. *Go* with a more generalized future meaning can also encompass subjects which are not capable of self-propelled physical movement or willpower as in *The party is going to be held Saturday*. Being fully grammaticalised as a future marker, *go* has no selectional restrictions concerning the animacy of its subject any more.

To wrap up this section, historically, *go*'s spatial function is a semantic precursor to its temporal function. Synchronically, both functions can hold simultaneously, resulting in ambiguity of meaning as in example (4.36), where we have temporal, as well as, spatial displacement. I have thus provided evidence for the motivation of the link between the core **GOAL** and *go*'s future time reference use in our semantic field model. It has to be pointed out that this link also holds when no movement component is present as in example (4.37). It is, again, not the movement which is the motivating factor for the new function but the semantic trait **GOAL**. While telicity can be physically concrete, as well as volitional and abstract, diachronically, it was the telic physical movement component that allowed for the functional extension to arise. The notion of **GOAL** underlies *go*'s newer future time reference function. Thus, the future marking property of *go* is metaphorically and metonymically motivated by the telic trait of movement inherent in *go*'s semantics.

#### **4.4.3. The Modal Function**

As most activities have both a physical and a mental component, the transfer from the literal meaning of 'the subject is on a path moving towards a goal' to 'the subject intends to do the event indicated as the goal' is quite straightforward (cf. the discussion under 4.4.2.). The movement and the path only have to be taken figuratively for the internal, purposive sense to arise (cf. Nicolle 1997, Poplack and Tagliamonte 2000, Royster and Steadman 1923). Conversational implicatures of

intentionality easily arise with respect to movement (a self- or other- sustained activity). These implications can then become conventionally associated with *go*.

Our world knowledge tells us that most activities entail physical displacement of some sort, especially when the context makes clear that they are conducted at a geographically distant location from the deictic reference point. Hence, according to Grice's Maxim of Quantity, the implication is that stating that someone *goes* to do something should involve new information, not only displacement. It is open to attract implications based on contextual inferences. In the same vein, Horn's (1984) principle of least effort states that speakers say no more than they must and that hearers require that speakers say as much as they can. Because of this tension, of which both parties are aware, the hearers are required to extract all the meaning possible from the message, which leads to an "enriched interpretation" (Enfield 2003, Wilson and Sperber 2002). As concrete, spatial and abstract, volitional goals occupy the same syntactic slot, modal and motion meaning can overlap given the appropriate context. Consider (4.38), where both the telic movement and the intention reading pertain.

(4.38) US English - *Missionaries*

- B I mean they have a training center where they teach them,  
it's called the missionary training center in Utah  
and they have to be taught the langu[age customs and all those kinds of things.
- A [um-hum,
- B and then they um you know **go** and actually live in that country for,  
the the uh young men do it for two years and the young women for eighteen months,

The line marked by the arrow expresses both intention and movement. Speaker B instead of uttering *they um you know actually live in that country...* chooses to say *... go and actually live in .....* Her choice of words could be taken to mean that she wants to stress the fact that the event <e> (the missionaries going abroad) involves motion. But as the conversation has been about missionaries serving in countries other than the US, this has already been asserted and would constitute old, non-reportable, information. Consequently, in line with Grice's Maxim of Quantity, the *go*-construction can be interpreted as expressing that those missionaries have the intention to do <e>, *live in that country*. If such implicatures regularly arise from a certain linguistic surface structure, they can become interpreted as inherent in the form.

Traugott (1989) has called this process, where inferences become so associated with the verbal expression that they become part of the meaning, “pragmatic strengthening”.

Note also the occurrence of the adverbial emphasiser *actually* in this chunk of speech (cf. the OED entry for *actually* “(5) ..... added to vouch for statements which seem surprising, incredible, or exaggerated: ‘He has actually sent the letter after all.’”, see also Quirk et al. 1985:583 and Clift 2001:28 who says “the very presence of *actually* marks the TCU AS informative”). What makes the chunk of speech above reportable or informative is the fact that the missionaries have the intention to do <e> in spite of all the obstacles such as the different language, the unknown customs etc. as cited in the chunk of speech above. Example (4.38) could be paraphrased with purposive *to* (Hopper and Traugott 2003) as *go in order to (actually) live in that country*. Hence, *go* expresses physical motion as well as the intention of the actors.

In this vein, consider the difference between

(4.39) She hit him with a bottle<sup>40</sup>

(4.40) She went and hit him with a bottle

Where (4.39) is a matter-of-fact description which does not necessarily express assault. A continuation such as ... *but actually she didn't mean to* is possible, hence intentionality is a conversational implicature and cancellable. Example (4.40), on the contrary, expresses intention to *hit him with a bottle* and, if uttered in a context of a court hearing, would not leave much leeway for any claims to the accused person's innocence<sup>41</sup>. Note that (4.40) expresses more intention than (4.39), as *go* is used here with a meaning which is situated closer to the intention function. This is even more evident in the next two sentences:

(4.41) She killed him

(4.42) She (actually) went and killed him

Whereas the agent in (4.41) could be held for manslaughter, the agent in (4.42) is implied to have committed an intentional killing. This is because *go*, with its telic meaning, expresses the intention to achieve some kind of **GOAL**.

---

<sup>40</sup> Example 38 was taken from the State of Michigan Court of appeals, Wayne Circuit Court No. 221289.

<sup>41</sup> These interpretations obviously only apply given a continuous, non emphatic intonation curve.

Through the aforementioned conventionalization of conversational implicatures, *go* has long extended its use to cases where no displacement is, at least tacitly, implied and the *go*-construction is now also found in contexts where it is contextually clear that the agent is not moving in order to fulfill the intention/event expressed in the complement. *Go* then operates on an abstract level without any spatial connotations. Example 4.43 gives such a situation where no movement is implied in the context, which leaves room for *go*'s interpretation as modal as well as temporal. Speaker A reports an incident when, on a lazy afternoon, he gets an invitation to go out. He tells the girl who phoned him that he and his friends prefer to stay in and watch golf.

(4.43) US English - *The lazy Sunday*

A     it's like we just woke up we had a hangover and everything you know,  
        and so we just woke up and uh we're watching TV,  
 →     and I go "no today we're just **going to** dedicate ourselves to watching golf".

Speaker A and his buddies are watching TV (*and so we just woke up and uh we're watching TV*) when he gets the phone call. Consequently, it seems likely that event <e>, *watch(ing) golf* at a later point in time does not imply any spatially telic movement. As A and his buddies' future activity is continuous with their present one, we can assume future action to be equilocal to the one executed at time zero. Thus the possible interpretation of *go* as spatially telic is cancelled out. It can only have figurative uses which have to be inferred from the context: As the event *watch(ing) golf* is to take place at a later temporal stage than the deictic reference point marked by *and I go*<sup>42</sup> "[quote]", the verbal expression *going to* can be interpreted as signalling future tense (prediction). But *going to* in this case can also be interpreted as expressing intention on the part of speaker A and his buddies. This is because intention and future functions are target-oriented, the epistemic and temporal expression of a **GOAL** (cf. Leech (1971:54) "future of present intention", Fleischman 1982, also Hopper and Traugott 1993). A and his buddies' staying at home is not some mere chance coincidence in future time but is presented as executed intentionally, a situation of planned likelihood.

---

<sup>42</sup> For an interpretation of *go* as a quotative consider section 4.4.4.

On a similar note, Dahl (1985) pointed out that the central function of future is less a temporal category and more a category resembling agent-oriented and epistemic modality with important temporal implications.

Leech's example *\*I wonder if she's going to know you* underlines the intentional function of *go* in the future sense: it strikes one as odd "because one cannot will oneself into knowing somebody" (1972:55). Hence, Ultan (1978) claimed that while cross-linguistically, future markers are often used with atemporal functions, they seem to be especially often associated with mood or modality, from which they are frequently diachronically derived. Often, their primary function is the expression of modality and the future time reference is only a secondary, or at least temporally later function (cf. also Güldemann 2001, Bybee 1995). Bybee, Perkins and Pagliuca (1994) and Sweetser (1988) maintain that in a movement construction the co-occurrence of spatial as well as temporal meaning is given from the very beginning and that the spatial is explicitly expressed first and the temporal follows from it (cf. Section 4.4.2). Hence, there is reason to believe that, diachronically, intention even seems to be the crucial bridge to the development of the prediction function.

While examples (4.37) and (4.43) have shown the functional ambiguity of *go* between modal and temporal function, I will not make any claims here concerning the temporal order of the functional developments. This chapter has shown that *go*, a verb of movement, can have strong modalised functions. I have illustrated the close association between directional movement and intention. The connecting tie has been identified as the notion of a **GOAL**, concrete in the first case, figurative in the second. The link between intention and movement is thus motivated. Furthermore, a close association between intention and future meaning has been pointed out. When it is asserted that an agent intends to achieve a goal (event <e>), we can safely infer (unless otherwise advised and as long as the goal is realistic) that the agent will carry out <e>. In this vein, Langacker (1985:23) states "the speaker/conceptualizer... traces mentally along the path in order to situate the process in relation to a reference point". The common denominator between the future and intention sense is salient: an abstract **GOAL**.

Within the radial structure, we have to assume a tightly-knit triangular cognitive structure between goal-directed motion, intention, and future reference as depicted in Figure 4.5. This point is underlined by Danchev and Kytö's (2002) research into the future marking function of *go*. Looking at the Helsinki corpus, the first English example with future connotations which they found dates from 1438 and combines features of movement, intentionality and near futurity. Thus, the close association of those notions in a synchronic radial structure is sustained diachronically.

A conventional unidirectional model would have problems depicting such a close association between the dimensions of modality, temporality and spatially directed displacement. It would not be able to represent the functional three-way ambiguity of some examples. But note that the radial structure model proposed here can cope with the linguistic reality of closely knit, at times even overlapping, multiple functional categories by representing them as interconnected functional extensions from one semantic core. The ensuing semantic field model depicts them in more or less close associations with potential ambiguity stemming from and relying on one common semantic origo, the notion of a **GOAL**.

#### 4.4.4. The quotative function

The advent of two newly<sup>43</sup> grammaticalised quotatives *go* and *be like* into the formerly stable pool of quotative introductory items has triggered a wave of interest in their patterning with respect to intra and extralinguistic factors (Buchstaller 2001a,b, 2002a,b, 2003a,b, Cukor-Avila 2002, 2003, Butters 1980, Ferrara and Bell 1995, Macaulay 2001, Tagliamonte and Hudson 1999, Singler 2001, Winter 2002, and many others). Example (3.1), here replicated as (4.44) illustrates *go* in this use:

(4.44) US English - *The long haired guy*

A:     the other day I went into a bar,  
          and this guy asked me to dance.  
          all he saw was my hair,

---

<sup>43</sup> Newly grammaticalised in the sense of newer than the other members of the quotative cohort, which has been stable for centuries (Ferrara and Bell 1995).

and he **goes** “do you want to dance”?  
 I turn around and **go** “what”?  
 he **goes** “do you want to dance”?  
 I **go** “no no”.  
 he **goes** “oh oh I’m sorry”,  
 I **go** “yeah, you better be”.  
 I **go** “[you better be”.  
 B [that's hilarious,

Example 4.44 gives a whole row of *go*-lexemes in quotative function. Here, *go* precedes reported speech and has taken on a function roughly corresponding to *say*. (The details of the functional differences between *say* and *go* are discussed in Chapter 3 or Buchstaller 2001a) In this example, *go* introduces linguistic material, such as “*do you want to dance*”, “*what*” etc.

#### 4.4.4.1. *Go as an introducer of voice effects, sounds and gestures*

In what is considered its more prototypical use (Tagliamonte and Hudson 1999) *go* introduces quotes containing non-linguistic material. This function is exemplified in (4.44) and (4.45), below:

##### (4.45) US English - *Kids’ games*

B: my kid didn’t care,  
 A: I know,  
 B: He picks up a stick and **goes** “bang”,

##### (4.46) BrE - *Home sweet home*

B: and I have got home and after dinner,  
 and **you are just kind of going** “urghhhhh”,

Here, *go* introduces the sound effects *bang* and *urghhh*. But it is also frequently found with quotes which contain voice effect and gesture as in (4.47).

##### (4.47) US English - *Watching baseball*

A: so I tu-I watched entire games of baseball  
 and **they’re going** “o::h.hh my ↑Go:d”



Those re-enactments of (selective aspects of) previous events based on voice, sound or gesture have been called “mimesis”. Clark and Gerrig (1990) show that the enquoting person “does not say what the content of the quote is (i.e. what was said) instead he does something that enables the hearer to SEE for himself what it is, that is to say, he *shows* this content” (1990:802, cf. also Goffman 1981). *Go* in (4.45) introduces a mimetic enactment of a previous non-linguistic event, *bang*. By replaying her son’s playful conduct via a sound effect, speaker A is demonstrating the WHAT and the HOW of the original speaker’s action. She re-plays a selectively depicted aspect of her son’s behaviour at the time of the original occurrence.

Güldemann (2001) claims that mimetic quotes are marked and by that feature are better enquoted in a structure that is marked in some way as well (cf. also Yule and Mathis 1992). This claim is substantiated by his findings on African languages as well as by a study of the enquoting constructions in American English. As discussed in Chapter 3, the newly grammaticalised quotatives are principally used as markers of mimesis. *Like* and *go*, the newcomers in the quotative complex and still more marked constructions, highlight the HOW, the demonstrative-enacted side of the material. Romaine and Lange (1991:238) claim that *go* lacks the explicitness of *say* and has special connections with the auditory-vocal channel. It serves as a cue for that channel and introduces sounds or onomatopoeic expressions rather than words. “*Go* translates loosely as ‘makes the sound of’” (240). In contrast, the most frequent dialogue introducer *say* is generally used much less with mimetic performances. It is pragmatically unmarked and thus less compatible with marked quotes (Buchstaller 2002b).

According to Butters (1980: 307), *go*’s enquoting function for onomatopoeia and non-linguistic sound imitations precedes its use as a quotative item for speech. “The imitative *go*, present in the language for centuries, would seem in a more general way the most likely candidate for the source of the semantic extension”. Butters sees imitative *go* as the developmentally primary, narrower function. He says that “it is actually just a small movement from this specialized use to the broader one” (ibid 1980: 307). Güldemann (2001), while agreeing on the diachronic scenario traced by Butters, reverses its interpretation. He claims that while *go* originated as a mimesis marker, its use as a quotative of direct discourse is the more specialized function.

Counter the standardly assumed “speech-verb channel” (Lord 1976, 1993, Saxena 1995), he argues for a broadening of *go*'s functions from an introductory item for sounds, gestures etc. to a general mimesis marker.<sup>44</sup> He points out that mimesis is actually a more general domain than direct reported discourse and that a mimesis marking construction is generally a potential source for introducing linguistic elements. This claim is backed by findings by Buck (1915). Tracing a diachronic scenario for the Indo-European languages, Buck shows that verbs introducing noise can become general discourse introductory items.

Despite the details of their interpretative difference, both Güldemann and Butters stress the close relationship between the enactment of sounds, voice effects, and gestures and the rendering of linguistic material in direct discourse. In my model, I maintain the category quotation as a superordinate concept for quotative verbs that frame more or less mimetically enacted renderings. But due to differences in their slightly idiosyncratic developmental histories and due to the fact that we can construct functional extensions involving different links, I divide the realm of quotation into 2 subcategories: introductory items for sounds, gestures etc. (mimesis) and introductory items for linguistic material, being well aware that this distinction is arbitrary due to the fluidity between (conventionalized) speech and (not -yet- conventionalized) sound.<sup>45</sup>

Güldemann (2001) shows that, in general, mimesis is embedded in the surrounding linguistic material as follows: (1) It does not require an embedding construction such as a predicate (cf. Cameron 1998), (2) if it is co-occurrent with a discourse introductory construction, it does not enter in a syntactic relation with it (cf. Partee 1973:418 “the quoted sentence is not syntactically or semantically part of the sentence which contains it.”, Munro 1982), (3) it is usually set apart from the co-text by a pause or other suprasegmental features (Klewitz and Couper-Kuhlen 1999,

---

<sup>44</sup> A comparative case is German (*so machen* (Golato 2000), which occurs with sound imitation and gesture but rarely with direct discourse (yet). Time will show whether *so machen* grammaticalises in the same way and whether it generalizes into a fully-fledged  $V_{quot}$ .

<sup>45</sup> As the reader will have noticed, mimesis can be understood to be a more or less encompassing category. It has been used in the literature mainly as the superordinate term for more expressive re-enactments like sound, gesture and voice effects. Because of its usefulness as an umbrella term, I will employ it in this sense here. Güldeman (2000:10) is right, though, when he reminds us of the fact that any kind of re-enactment is to be considered as mimesis, no matter how expressive it is. Direct reported discourse lies within the functional domain of mimesis. Thus, mimesis actually is the superordinate term for linguistic and all other demonstrations.

Günthner 1998), and (4) it constitutes an independent and focal intonation unit by itself. Along the same lines, Clark and Gerrig (1990: 772) point out that quotations are embedded in the discourse but not in the sentence that hosts them. The internal and the external structure of embedded quotations have to be kept apart.

While I do not think that their statement can be generalized to all quotations (cf. Cameron 1990:771/2), with respect to the interpretation of *go*, I follow Clark and Gerrig's reasoning that "internally the quotations again depict sentences but externally they are a manner adverb and predicate nominals or adjectives" (1990:772). Hence, we can interpret the quote following *go* externally as a manner adverb. Let us consider Butters (1980) example

(4.48) ....and this little pig **went** wee-wee-wee all the way home.

*Go*, as a verb of movement, predicates the displacement of the agent *pig*. This movement is qualified by the "adverb" *wee wee wee* which – if we had to put it into words while losing a lot of expressive force – is roughly equivalent to *crying*. It further qualifies the action of the agent *this little pig*. *Wee-wee-wee* co-occurs with linguistic material but is not embedded syntactically in the sentence. It qualifies the way in which the pig went home, a marker of the HOW. The above scenario can obviously be replicated with other sorts of mimesis. Gesture, for example, can be produced simultaneously with a speech signal as the two signing techniques rely on different media (the same is true with facial demonstrations). Scholars researching in paralinguistic signs and bodily movements (Kendon 1980, McNeill 1985, 1992, Streek 1988) show that verbal and bodily behaviour are to be considered closely associated if not dependent on one another. They form complementary channels to convey information about the referents' mental and physical state, where the linguistic material is the more symbolic rendering and the gestural-sound channel is the indexical rendering. Also, rhythmic enactment, singing, etc. can be superimposed on verbal behaviour and complement its impact.

Once speakers repeatedly qualify movement with manner adverbials such as *wee wee wee*, *go* co-occurs more frequently with mimesis and, via the process of metonymy, it can become associated with its environment. The co-occurrence of mimetic performances with *go* then becomes conventionalized because its context of

use encroaches upon its interpretation (Sperber and Wilson 1986). It is exactly small-scale loci such as the occurrence of *go* preceding mimetic enactment which lead to the infusion of contextual implication into the meaning of an item and bring about an alternative analysis (cf. Stern 1931:353). Again, the conventionalization of conversational inferences, which is especially frequent in the earlier stages of grammaticalisation (Hopper and Traugott 1993) can go so far as to become part of the inherent meaning of the form or even replace its original meaning. As Grice puts it, it is not unusual “for what starts life... as a conversational implication to become conventional” (1975:58). In our case, though frequent association of *go* with mimetic enactment of whatever sort, *go*’s mimesis introductory function starts becoming a conventional use within its field of functional extensions.

Note that it is not the telic sense of *go* that is implied in its grammaticalisation as a mimesis marker. In (4.49) below, no direction is implied:

(4.49) US English - *Latin American Dancing*

- M: I don't know what those Brazilian women are doing,  
 J: what are they do[ing],  
 M: [but their hips are like,  
 J: are they going-are they *going* like really fa::st?  
 → I mean are **they going** [one two three, (rhythmical)  
 M: [I mean,  
 J: one two three-,  
 M: first you had this wi::de sw:::ay of their hips this way,

Here, an atelic motion (dance) is portrayed via enacted material, *one two three*, which is rendered as if uttered or thought by someone doing dance steps, in a rhythmic, beat-following manner. The dancers are represented as moving in a *one-two-three-ish* fashion, as if going with an imaginary rhythm.

*Go* as a movement verb is modified by an adverb of manner which qualifies the HOW: *one two three*. Thus, given its placement in the clause before the mimesis, *go* is in the right kind of syntactic position to be reanalyzed as a verb of quotation. The material following it is a selective demonstration of the agent’s behaviour at an earlier time, thus a mimetic re-enactment in Clark and Gerrig’s (1990) and Wierzbicka’s (1974) sense. *Go* frames a mimetic quote and can thus be interpreted as functioning as a quotative verb. Note again that the scenario in (4.49) involves no telicity. The

dancers have no concrete goal, neither in a spatial telic sense – they are not dancing towards something – nor in the form of an interlocutor to which they direct the words/action represented by *one two three*. The important notion is that of a **SOURCE**. The mimetic enactment is stemming from actors, the dancers in question. Whether it is intended for someone else (i.e. whether there is a potential goal) is irrelevant. What matters is the performing actors who produce what is later rendered as mimesis, the point in space where the enactment comes from. It is the **SOURCE**-encoding function of *go* that can give rise to its use as a mimesis marker. In this case we have an overlap between *go* as a mimesis marker and between its ablative atelic movement meaning. The mimesis stems from, “move(s) away, depart(s) from, leave(s)” (OED II, 21) the author who produces it. Therefore, *go* in this sense should be grouped under heading II “uses in which movement from a place is the primary notion”.

We saw this, too, in example (4.48). The little piggy has some sort of goal (*all the way*) *home*. But this fact is tangential to the interpretation of *go*. I claim that even though a goal is present, it is not a crucial component for *go*'s functional enlargement and that the analysis of *go* as a mimesis introduction maker does not rely on a telic element in the context. If we substitute a non-telic adverbial such as the PP *along the street* for the goal-oriented adverbial *all the way home*, we get the same interpretation for *go* as shown in (4.50).

(4.50) ....and this little pig **went** wee-wee-wee along the street.

In example (4.50) the context does not contain any spatial goal. Yet, *go* does take the function of introducing reported action even if it is not directed at anyone and even if it is not heard. Note that even in *go*'s earliest attestation as a mimetic introducer (OED 1791) *And his noble heart goes “pit a pat”*, there is no addressee, explicit or implicit, for the sound/movement *pit a pat* which is re-enacted in the quote. *Go* can be interpreted as a verb of motion further qualified by an adverb of manner, at the same time, it is a quotative verb which frames mimetic re-enactment. This example is significant because it shows that *go*, in its early stages of grammaticalization, did not need a contextually given telic element. Quite the opposite seems to have been the case: it is the element of a **SOURCE** which is the decisive factor in *go*'s grammaticalisation to an introducer of sounds and gestures. I will show in the next

section how the notion of a goal comes into play with respect to *go*'s functional generalization which then also includes reported speech. By now, the link from ablative movement to mimesis introduction in *go*'s radial structure has been motivated via the notion of a **SOURCE**. Here, again, the two functions are not discrete. They overlap in some contexts and can be ambiguous.

#### 4.4.4.2. *Go as an introducer of reported speech*

Reported speech in this chapter refers to direct reported speech only. In my US English data, *go* cannot be used as an introductory item for indirect reported speech (cf. Golato 2000, Hudson 1985, Schourup 1982b). Although very infrequently, there are some tokens of *go* framing indirect speech in my British English corpus (cf. Chapter 5, as well as Buchstaller 2002a). This finding does not violate my claims, as indirect speech, albeit involving a different deictic orientation, can nevertheless involve telicity as in *She told me that she wasn't sure*, where *me* functions as the telic addressee, a **GOAL**.

Once *go* is already grammaticalised as a marker of enquoting non-directional sounds or gestures (according to the OED, its first attested use as a mimetic enactment marker is 1791), all it takes is a goal-directed performance on the part of the speaker to an explicit or implicit addressee for it to enter the terrain of a quotative in the interactional, directional sense. This can be exemplified with the following stretch of talk:

#### (4.51) US English - *The party*

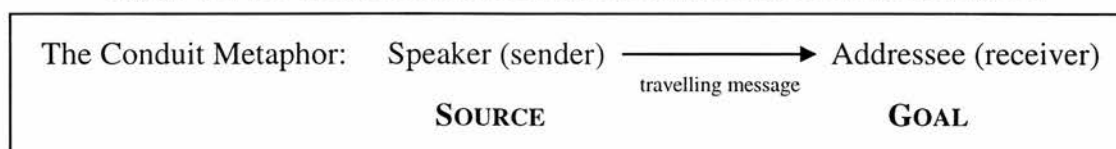
A:     so Cathy calls me up,  
          and she wants Jonathan's ph- address.  
          and I **go** "for what"?  
          she **goes** "well we have to invite him to the New Year's Eve party"?  
          and I **go** "why"?  
          she **goes** "well XX opened his big mouth",  
          and I said "oh so you're going to host them are you"?  
          she **goes** "what do you mean host them".  
          I said "if you invite them up there for a party,  
          they're gonna assume that they are staying with you",

Example (4.51) contains several question and answer scenarios (Q: *why*, A: *well XX opened his big mouth*, Q: *what do you mean host them*, A: *if you invite them up there*

*for a party...* ). Arguing with the next turn proof procedure used in conversation analysis (Sacks, Schegloff and Jefferson 1974), according to which a second pair part sequentially needs a first pair part, the fact that an interlocutor provides answers to questions speaks for the fact that these questions were interactively realised, goal-directed instances of speech. *Go* enquotes linguistic material which stems from an author and is intended for an interlocutor, the receptor. The same situation also holds for example (4.44). The quotes in both stretches of speech contain other-directed, telic speech which involves the notion of a **GOAL**. And in both cases the addressees confirm their role as an interactional goal: they acknowledge the receipt of the question by giving an answer. Hence, *go* has taken on the function of a fully fledged introductory item for quotes in a context that is defined interactionally as telic.

*Go*'s functional transfer from a verb of motion to an item associated with reported speech can be explained by what Fillmore (1997) referred to as the Place/Time parallelism. Sweetser (1987) argues that the lexical fields of physical motion, action, location and of mental states and speech acts are distinct semantic systems but that they are metaphorically connected and indeed overlapping. As our understanding of abstract processes (such as speech acts and mental states) is modelled on the concrete physical world, the source vocabulary for intellectual and speech domains taps into the expressions for the corresponding physical activities. Via this analogy, a communicative act can be metaphorically seen as an instance of motion: a travelling message from one interlocutor, the sender, to the other, the addressee. A communicative model relying on the conduit metaphor (cf. Lakoff and Johnson 1980, Lakoff 1987, Foley 1997) as depicted in Figure 4.6 makes the overlap between temporal and quotative function explicable.

**Figure 4.6: The metaphorical base for *go* as an introducer of reported speech**



According to Figure 4.6, we can conceptualize communication as a travelling message from the speaker (the sender, thus the **SOURCE**) to the addressee (the receiver, hence the **GOAL**). The temporal vantage point is to be taken to be the point at which the

encoding of the utterance takes place. Decoding is understood to occur in the prospective future as signalled by the forward trajectory of the telic movement verb.

Consequently, the semantic primitives of ablativity ‘starting point’ as well as telicity ‘endpoint’ are the decisive factors in the grammaticalisation of a movement verb into an introductory item for reported other-directed speech. Both, the notions **SOURCE** and **GOAL** play into the grammaticalisation of *go* as a quotative in the interactive sense. Hence, we have to postulate functional extensions from the core which rely on two different semantic traits: telicity as well as ablativity.

*Go* in its functions as a mimesis introducer and as an introducer of reported speech share one semantic feature (**SOURCE**). We can postulate a diachronic scenario where *go* first occurred as a mimesis introducer (*And his noble heart goes “pit a pat”* OED 1791) which relies on the notion of **SOURCE** only. Later, it took on the quotative function for interactive reported speech, which also relies on the notion of a **GOAL** as a second feature. This means that once *go* was grammaticalised in its function as a marker of enquoting non-directional sounds or gestures, all it needed was a **GOAL**-directed performance on the part of the speaker for *go* to take on the function of a quotative in the interactive, directional sense. Synchronically, the two functions of *go* still overlap in their common underlying feature **SOURCE**. Other-directed *go* has a second functional link to the centre, **GOAL**. Note that this explanation further corroborates Güldemann’s (2001) claim: because adding a semantic trait to an item increases its specificity, *go* as an introducer of reported discourse has the more specialised quotative function.

At this point the division of quotation into sounds/voice effects/gestures on the one hand and linguistic performances on the other hand has to be reconsidered. While it is true that sounds/gestures are prototypically not intended for an interlocutor, there are certainly situations where they are. Consider a typical example.

(4.52) personal example - USE

This guy **went** “xxx”

(giving the finger)

Here, the enquoted gesture is clearly other-oriented. It stems from an author, a **SOURCE** and is intended to convey an attitude to an opponent, the **GOAL**. Interactive



mimetic enactment is framed by *go*. The same phenomenon can also apply to sounds and voice effects. Consider, for example the whistle that certain men (and women) in some cultures utter upon seeing an attractive member of the opposite or same sex. This whistle can be non-directional, not meant for the ears of anyone else, a mere acknowledgement of one's inner states, an outloud. But the prototypical macho-whistle is so annoying exactly because it is directed at the person in question (and possibly even bystanders) and does not only signal inward approval. When quoted with *go*, the whistle as mimetic enactment could be a telic or non-telic performance depending on the situation it occurred in (whether or not there was a **GOAL** who was intended to hear the sound). Regardless of the situation, the whistle stems from a **SOURCE** and it is this notion which is important for the grammaticalisation of *go* as an introducer of non-directional quotes, irrespective of their content.

At the same time, there are reported speech events which are not telic. An example is (4.53).

(4.53) US English - *Money comes and goes*

B: when you go out into the working world,

A: uh huh,

B: it- it changes quickly,

A: I am sure,

→ B: um you look at your paycheck and **you ghhhhooh** "oh my gosh where did it all go"?

A: hhhhh,

Notice that in this case, the original utterance – which is replayed in the quote in the marked line – is not other-directed. The **SOURCE**, from which the utterance *oh my gosh where did it all go* stems is speaker (B), who is reporting the incidence. At the time B looked at her paycheque and uttered *oh my gosh where did it all go*, there was no telic **GOAL**, hence no addressee to whom B could have addressed her utterance. Rather, B is reporting her reactions to the dwindled amount of money to A in the form of a hypothetical quote (see Chapter 3), which puts into words (at least in the speaker's and hearer's now) what was going on in B's mind when she saw the cheque. As Goffman (1981:97) rightly points out, with no one present, speech acts like this are quite likely to have been omitted altogether (cf. also Tannen 1986, Yule and Mathis 1992). The utterance reported in (4.53) is not telic, as there is no real or contextually inferred

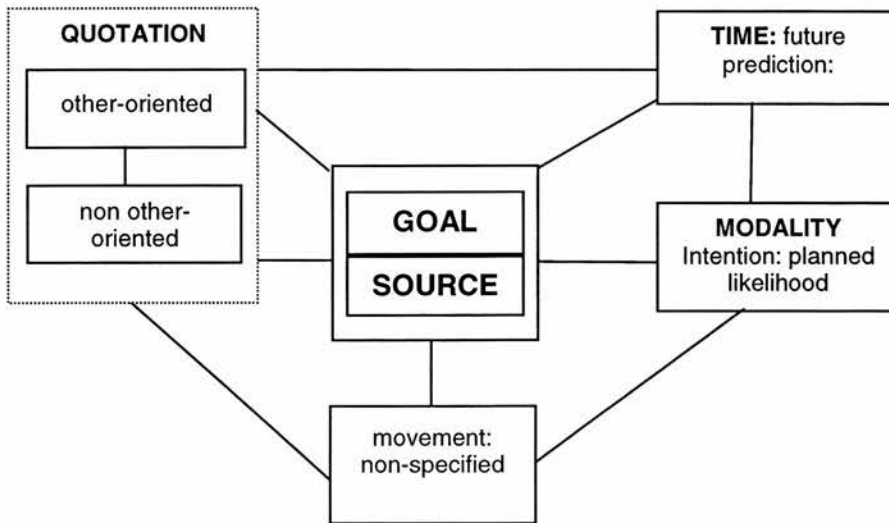
interactive, communicative **GOAL**. Hence, direct quotes containing linguistic material are not necessarily other-directed. This means that quotative *go* does not inevitably rely on the semantic feature telicity when framing direct reported speech of a linguistic nature. Even though (4.53) is a direct quote with linguistic material, the motivating functional extension in this case is that of **SOURCE** only.

Examples like (4.52) and (4.53) show that the decisive criterion for a subdivision of the quotative realm cannot be whether the reported material is linguistic or non-linguistic. We have seen that the semantic features **SOURCE** and **GOAL** can be present in communicative situations irrespective of the nature of the quote. What matters is the interactive, telic nature of the quote and the presence or absence of an interlocutor to whom speech/ sound/ gesture is addressed. Diachronically, it seems that *go*'s quotative function was first restricted to non-telic quotes, motivated by the functional extension **SOURCE** only. Later, it enlarged its repertoire (cf. Butters 1980) and acquired the more specified function of a quotative in the interactive sense (cf. Güldemann 2001). This was possible through the addition of a further functional extension already present in *go*'s radial structure, the notion of a **GOAL**. The above claim is motivated by the fact that the examples cited in the OED (ranging from 1791-1892) still all imply non-telic mimetic re-enactment. Butters (1980) is the first to mention full quotative *go* in its interactive telic sense.

Synchronically, the two categories of quotes have one underlying semantic primitive in common and one that distinguishes them; i.e. shared notion of a **SOURCE**. Both interactively realised quotes and verbally non-committed thought stem from (an) author(s). But only the interactively realised, other-oriented quotes need the telic notion of an interlocutor. Consequently, the decisive criterion is whether or not there is an addressee, a **GOAL**.

These different motivational parameters justify the division of the category quotations into two sub-categories: the diachronically earlier non-directional ([+source] but [-goal]) and the later directional ([+source] and [+goal]). This calls for a re-labelling of the headings of the 4.2.4.1 as *Go as an introducer of other-oriented quotes* and 4.2.4.2 as *Go as an introducer of non-other oriented quotes*. The same applies to the functions in the realm of quotation of the radial structure. Consider Figure 4.7 below:

Figure 4.7: The radial structure model of *go*



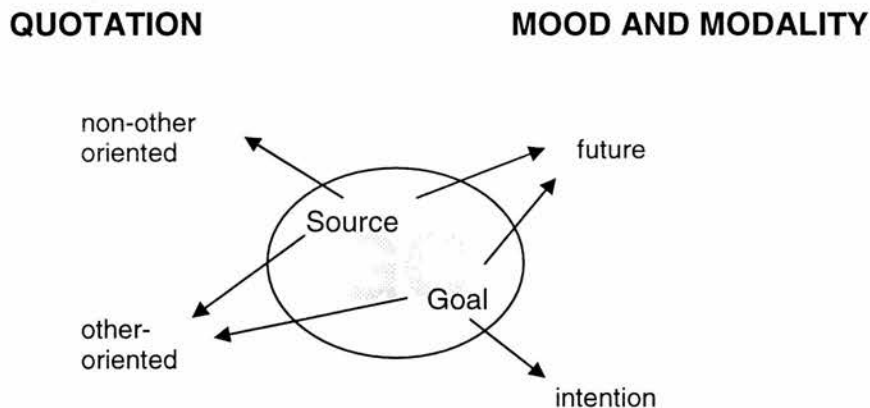
Notice that the radial structure model depicts the close association of the two quotative functions by conceptualizing them as sub-functions pertaining to one common category, quotation. They have one semantic trait in common, the notion of a **SOURCE**. Their difference lies in whether or not they rely on the existence of a **GOAL**.

Notice also that the possibility of ambiguity and overlap between the functions movement and quotation, time and modality and between all the above categories with the central semantic primitives of **GOAL** and **SOURCE** can be captured by this model. As with *like*, the lines between the functions should be considered as clines linking the focal members of prototypical functional categories. Singular tokens of *go* are thus situated somewhere along the line between the endpoints, oscillating more towards the one or the other pole. Far from forcing linguistic structure into a linear model, this highly flexible radial structure provides an explanatory model that can cope with the linguistic reality. As poly-ambiguity and non-determinism are an integral part of the structure, cases of where one example has several functions simultaneously do not constitute a problem for this model. Rather than accounting for ambiguities one by one in terms of respectively earlier and later functions, the full linguistic messiness of fuzzy categories is provided for via a poly-dimensional approach.

#### 4.4.5. The substructures of the grammaticalisation schema

The above discussion has given evidence of the complexity of *go*'s radial structure. Its synsemantic field hosts multiple links of various underlying motivations (metaphor, context induced interpretation etc.). Its grammaticalisation can be understood synchronically as an interwoven net of functional extensions from one common core. Within the network there are functions which are linked to one another to greater or lesser degrees. All display more or less close semantic –pragmatic relationships to a core meaning. Another finding underlines the complexity of the model: there are two sub-structures involved. The links within the network are based on two different functional extensions based on semantic core traits. The structures originate and overlap in the core and spread out independently, forming a superimposed web of connections. The constitutive overlaid structures of the model are depicted in Figure 4.8.

Figure 4.8: The substructures of *go*'s radial structure model



This model shows on which semantic traits the functional extensions in *go*'s core-structure model rely. One motivating semantic trait that has crystallized through the discussion is the notion of a **GOAL**. As we have seen, the telic use of *go* can be shown to be at the base of the future and the intention function. It also gives rise to the enquoting function for other-oriented speech/sounds and gesture.<sup>46</sup> The notion of a

<sup>46</sup> The semantic trait **GOAL** that these notions have in common justifies the overlap that Comrie (1985:44) has shown between intention and future meaning as “alternative worlds” with lower epistemic stance. We can easily link up quotation-introduction as a further function that implies an alternative world: the opening of new mental space (Fauconnier and Sweetser 1996) representing the speech/thought of another individual nested into the account by a narrator.

**SOURCE** is involved in the development of the future function and the function as a quotative introducer for non-other directed speech/sound/gesture and thought.

#### 4.4.6. Conclusion

I have shown in this section that all of *go*'s functions can be interpreted as direct functional extensions of the common semantic core. They are linked to this core via the semantic traits allativity and telicity, which all functions share to a greater or lesser extent. As has been shown throughout the discussion, **SOURCE** and **GOAL** are the semantic primitives which made *go*'s reanalysis possible. They are the heart of two intertwined substructures and they link the outgrowths to each other as well as to the centre. Notice that both the future and the other-oriented quotative function involve both underlying notions. The superimposed structures centering around **SOURCE** and **GOAL** overlap at the two functional nodes which rely simultaneously on both explanatory parameters.

Figure 4.8 shows that not only is a radial structure not unidirectional, it can also rely on multiple functional extensions. As cognitive structuring is a flexible and perpetual process, it can lead to structures which rely simultaneously on various explanatory parameters and diverse pathways from a common core. Motivational similarities leading to reanalysis (via metaphor, metonymy etc.) are everywhere and can turn into co-occurring, superimposed motivating forces in a grammaticalisation schema.

#### 4.5. General Conclusion

I have shown that in American and British English, the synchronic status quo of *go* and *like* is best captured by a radial structure model, which is capable of rendering the versatility of the lexeme within its synsemantic field.<sup>47</sup> A unidirectional chain is

---

<sup>47</sup> On a similar note, cf. Güldemann's (2001: 28) statement that the hypothesis that a quotative introductory element is derived from something other than a speech verb can often explain in a more

not able to cope with their synchronically complex polyfunctionality and with their ambiguity of function in the singular case (between the future and intention function as in (4.37), between future and movement as in (4.36) and between movement and quotative as in (4.49), for *like* the filler and focus function as in (4.6), the hedge and quotative function as in (4.15)). These instances of multiple ambiguity speak for a more multifaceted scenario. In view of the individual and highly local development of lexical items (cf. also Buchstaller 2001a, Lakoff 1987, Lakoff and Johnson 1980), a more complex premise than a unilateral cline seems in order.

The unidirectionality claim has long been one of the basic assumptions within the field of grammaticalisation. The historically underlying claim is Kuryłowicz's (1965:52) fundamental statement that grammaticalisation involves "an increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status". The movement is considered to be unidirectional from less to more grammatical and not vice versa. But even this basic underlying claim is not without exception. (Jeffers and Zwicky 1980, Kahr 1976, but refuted by Lehmann 1982/1995:16-20). Cases like the degrammaticalisation of *up* from a preposition to a verb (Givon 1975:96) show that the process can be reversed and that cognition and, accordingly, language do not follow rigid one-way lanes of development.<sup>48</sup> The unidirectional claim can be held in stronger and weaker forms and the literature has quite divergent views of how encompassing unidirectionality is to be interpreted (Traugott 1995, 2001). In its strongest forms, it claims that all grammatical items derive diachronically from lexical ones and that no reverse process is possible.<sup>49</sup> This claim postulates an early stage in a language where there are only lexical elements and no grammatical morphology. As there is no evidence for such a stage in Indo-European nor is it empirically attested cross-linguistically (Tabor and Traugott 1998), this claim raises serious questions and is nowadays upheld by only a few

---

natural way many of its characteristics in and outside the reported discourse domain. Furthermore, it involves much less bleaching of lexical meaning than has been assumed in the grammaticalisation of quotatives according to the unidirectionality-cline model.

<sup>48</sup> There has been ample discussion of whether unidirectionality is actually a hypothesis, thus testable and falsifiable, or whether it is inherent in the definition of grammaticalisation per se, thus a postulate (cf. Lass 2000, Lindstrom 2001, Traugott 1995, and many others).

<sup>49</sup> For an overview of counterexamples to unidirectionality see Janda (2001), Joseph (2001). For a criticism see Traugott (2001).

scholars (Haspelmath 1999, Lehmann 1982/1995 see also Heine et al 1991, Heine 1994). A much weaker version of unidirectionality states that it should be seen as an assertion about the orderliness and traceability of semantic change (cf. Bybee, Perkins, and Pagliuca 1994, Hopper and Traugott 2003).

Working from two case studies, this chapter has given evidence that synchronic orderliness does not depend on unidirectionality in grammaticalisation (cf. also Buchstaller 2001a, Lakoff 1987, Meyerhoff and Niedzielski 1995, 1998, Meyerhoff 2002, Mosegaard-Hansen 1997, 1998). Rather, systematic patterning can have multiple, non-linear but highly ordered forms. As has been displayed throughout the previous paragraphs, orderliness may just as easily result from an underlying radial structure and rely on functional extensions from one common core. More specifically, Güldemann (2001:4) raises serious reservations against the overgeneralization of the “speech-verb channel”.

I have given evidence that in conversational English the quotative introducers *like* (Buchstaller 2001 a,b) and *go* (Buchstaller in preparation) are polyfunctional and build up a whole network of relations. In cases such as these it is clearly contrary to the synchronic evidence to postulate a unidirectional channel with *say* or any other single starting point and to conceptualize derived functions as linearly aligned in a chain-like fashion. Rather, this chapter shows that change results from context-induced reinterpretations triggered by the in-sequence positioning of the item, and due to more general metaphors. A radial structure can capture the fact that the outcome of the grammaticalisation process is dependent on the small-scale environment of a particular item (cf. Hopper and Traugott 1993). As form, meaning and context play into the functional extension, a neat, all-encompassing linear generalization is impossible to make. This chapter proposes a radial field model which captures the semantic space in which functional movements co-occur and which most readily expresses the full ambiguity and overlap we find in real occurring data. The proposed radial structure shows how the superficially messy facts can be linked in an orderly way. This model is able to motivate all functions that *like* and *go* have assumed synchronically and has mildly predictive value in the sense that some functional extensions (within the same semantic field, relying on already established dimensions) seem less surprising than others. Even Bybee, Perkins and Pagliuca (1994:282), researching the cross-linguistic

propensities of future markers, contend that “close analysis of changes in progress may reveal a complex network of mechanisms applying together in sequence”.

In conclusion, the only general principle which holds in the light of *go* and *like*'s synchronic status quo and which can be applied to a common core model as well as to chain structures is the innovativeness principle mentioned by Andersen (1989, cf. also Hopper 1991), which basically builds on the “principle of exploitation of old means for novel functions” (Werner and Kaplan 1963:403). Consider also Traugott and Heine (1991:86/7)

“Grammaticalisation can be interpreted as the result of a process which has problem-solving as its main goal, its primary function being conceptualization by expressing one thing in terms of another. This function is not confined to grammaticalisation, it is the main characteristic of metaphor in general.”



From the fact that language is a social institution, it follows that linguistics is a social science...we must determine which social structure corresponds to a given linguistic structure and how general changes in the social structure are translated into changes in linguistic structure (Meillet 1926:17-18, quoted in Labov 2001: 22-23)

## Chapter 5

### The new quotatives in US English and British English

Ever since Butter's birth announcements of the quotatives *go* and *like* in US English in 1980 and 1982, the new quotatives were attested in numerous varieties of English, such as Canadian English, Nigerian English, Indian English, Scottish English etc. (Singler and Woods 2002, Macaulay 2001, Tagliamonte and Hudson 1999) and it is commonly assumed that they spread out from a US epicentre via the mass media (Singler and Woods 2002, but see Buchstaller 2003, Meyerhoff and Niedzielski 2003). US hegemony has been claimed to leave its imprints in the language in the form of the newest addition to the quotative system in other varieties of English just as it has on a cultural/societal and economical level. Hence, in a cross-variety study, Tagliamonte and Hudson write that "the diffusion of *be like* beyond the United States presents a possible test case for the examination of putative mega-trends currently underway as English increasingly becomes a global language" (1999:149).

*Like* and *go*'s widespread attestation has lead researchers to compare how they pattern with respect to social factors in different places. Such an investigation into extralinguistic constraints presupposes the variants' general linguistic comparability across varieties. To date, researchers have taken as given their status as the same variants in several varieties. Cf. Tagliamonte and Hudson (199: 147) "the linguistic trajectory of the innovative form *be like* is remarkably parallel, not only across the British and Canadian corpus, but also comparable with previous reports of this form in the United States". But sociolinguistics has shown in numerous studies that regionality has an impact on language structure. Prior to that, works in dialectology have taught us that we cannot assume a priori likeness across regional varieties (e.g.

Dutkova-Cope 2001, Kretschmar 2003, Macaulay 2002). It seems that assuming that superficial resemblance (in the shape of the same surface form) indicates overall equivalent linguistic constraints in various varieties runs the risk of comparing apples with oranges. But so far, sociolinguistic research on the new quotatives has fallen short of investigating their overall comparability on linguistic grounds. I would argue that failing to clear the linguistic ground before embarking on extralinguistic investigations is an oversight which ultimately threatens the validity of cross-variety comparisons (cf. Rickford and McNair-Knox 1992). I suggest that the first question we have to answer when doing sociolinguistic research of any kind is: are we looking at the same thing? Hence, in this study, I have to ask: do I have the same variant in both varieties?

In the same vein, in their 1999 study on the use of *like* in Canadian and British English, Tagliamonte and Hudson ask for the “assessment of the presence and use of *be like* in other countries, particularly the differential trajectories of linguistic change and diffusion in which it may be engaged, as well as the relationship to other members of the English quotative cohort in each context”(149).

This chapter has two concerns. First, I will address the issues raised in Tagliamonte and Hudson (1999). Looking at the patterning of the whole quotative cohort in British English (henceforth BrE) and American English (henceforth AE), I will show that *like* and *go* are separate but equal in the two varieties under investigation. This claim is based on the fact that they have *fundamentally* the same properties with regard to their co-occurrence with levels of hypotheticality (section 5.2.1), the mode of their surround (section 5.2.2), and their co-occurrence with mimesis and exclamation (5.2.3). But it will become evident that this is not the whole picture. While it is true that we are witnessing a restructuring of the quotative system in both varieties, *like* as well as *go* follow slightly different trajectories in both.

The second aim of this chapter is as follows: I will show how we can fruitfully incorporate insights from grammaticalisation into variationist sociolinguistics in order to ascertain general comparability across varieties, registers and other lects. Following Hopper and Traugott's (1993:30) plea that “a fuller integration of sociolinguistic (...) research with research on grammaticalisation still remains to be

worked out”, this chapter will show that it makes sense not only to turn to grammaticalisation in search for generally valid principles of language change but also that a grammaticalisation approach will help define what a sociolinguistic variant is.

Within variationist sociolinguistics, the extension of sociolinguistic analysis “above and beyond phonology” (Sankoff 1980a) sparked a cacophony of voices arguing for and against the fundamental comparability of phonology and other areas of linguistic inquiry.<sup>1</sup> The main problem at the time was the (in)-equity of phonology and syntax. According to Labov (1972a:271), “social and stylistic variation presume the option of saying ‘the same thing’ in several different ways: that is, the variants are identical in reference or truth value, but opposed in their social and/or stylistic significance”. Hence, semantic equivalence is the primary justification for including surface forms within one same category. But Lavandera (1978) challenged the assumption that there ever is semantic equivalence between syntactic variables. Shortly afterwards, Dines defined a discourse variable as having variants with the “same function in discourse” (1980). This raises the question: how is equivalence to be measured?

Cheshire (1987) and Romaine (1984) called for an interdisciplinary approach to be taken in order to grapple with the problem because “a wider perspective is needed than a purely sociolinguistic one ... [and because] ... sociolinguistics can both enrich and be enriched by other disciplines” (Cheshire 1987:274). Since then, numerous studies in sociolinguistics have drawn upon the framework of grammaticalisation and have incorporated explanatory mechanisms, such as clines (Meehan 1991, Poplack and Tagliamonte 2000, Romaine and Lange 1991), layering (Macaulay 2001), and persistence (Torress-Cacoullos 2001), but none has actually tried to closely intertwine grammaticalisation methodology and variationist sociolinguistics.

Looking at variation in discourse, I will argue that grammaticalisation helps us establish that the potential candidates for a sociolinguistic investigation (in this case *like* and *go*) can actually be considered the same variant. It will be shown that we can

---

<sup>1</sup> Cheshire (1987), Dines (1980), Labov (1978), Lavandera (1978), Romaine (1984), and others.

go about ratifying their status as the same variant by ascertaining that they are grammaticalising in the same way. Hence, sameness is defined here in functional as well as in distributional (cf. the notion of the surround introduced in Chapter 3) terms. Only when sufficiently comparable status has justified treating surface items as one same variant (viz. the same variant of the same variable), can we proceed to see whether there are differences in the use of those variants with respect to extralinguistic or intralinguistic factors such as variety. This will be further explained below.

This and the following chapter of my thesis can be read as a case-study which demonstrates how a cross-variety variationist investigation could proceed. First, sameness is established and then, as a second step, I conduct a variationist investigation into the patterning of the variable with respect to extralinguistic parameters (Chapter 6).

## 5.1. Emergent Grammar

Chapter 4 demonstrated *like* and *go*'s recent grammaticalisation as quotatives. New functions, more explicitly new more grammatical functions, developed out of already existing material (Kuryłowicz (1965 [1975]:52). The approach within grammaticalisation that proves most compatible with variationist sociolinguistics is that of emergent grammar. The term emergent grammar was coined by Hopper, who claims that he was influenced by historian James Clifford's statement that "culture is temporal, emergent, disputed" (1986:19). Hopper maintains that the same holds for language "a language is not a circumscribed object but a loose confederation of available and overlapping social experiences" (Hopper 1998:171).

Emergent grammar considers grammar not as an abstract system that exists independently of language, where language is only the implementation of a logically prior system; rather, grammar is emergent in the sense of being continuously negotiated by the speech community or communities. Grammar, then, is not the source of regular patterns of distribution but the outcome. After some time,

unplanned spoken discourse results in recurrent strategies for the expression of certain phenomena.

“Structure, or regularity, comes out of discourse and is shaped by discourse in an on-going process. Grammar is, in this view, simply the name for certain categories of observed repetitions in discourse. It is hence not to be understood as a prerequisite for discourse, a prior possession attributable in identical form to both speaker and hearer. Its forms are not fixed templates but emerge out of face-to-face interaction in ways that reflect the individual speakers’ past experience of these forms, and their assessment of the present context, including especially their interlocutors, whose experiences and assessments may be quite different.”(Hopper 1998:156)

Hence, grammatical structure is always provisional, as language incessantly changes and evolves. Recurrent strategies are constantly being negotiated in speech communities in which they develop. Grammar is thus social in that it reveals the practices of speech production of a given community at a given time on their way towards (potentially new) structure. This structuring is a “sedimentation of routine ways” (Couper-Kuhlen 1998:1) of engaging in talk.<sup>2</sup> In the same way, structure must always be anchored in the concrete situation of the utterance; grammar “is context-dependent” (Hopper 1998:157). And this is how we can link the approach of emergent grammar with variationist sociolinguistics: the concrete situation or context of an utterance is defined by the sociolinguistic variables of gender, age, socio-educational level, ethnicity, regional and social affiliation etc. We can expect language patterns to sediment differently in different social groups and, indeed, sociolinguistics has shown in numerous studies that different communities of practice have their own linguistic norms and customs (Bucholtz 1999, Eckert and McConnell-Ginet 1999).

This social aspect of grammar is of primary importance for our present study; the consideration of how *like* and *go* grammaticalise in different regional varieties. In this chapter, I will compare the patterning of *like* and *go* as quotatives with respect to their function and distribution within the cohort of quotatives in BrE and USE.

---

<sup>2</sup> Hopper says (1998:166) “language is, in other words, to be viewed as a kind of pastiche, pasted together in an improved way out of ready made elements” (cf. also Goffman’s (1981:151) concept of the “lamination of language”, Bakhtin’s (1981:281) concept of “heteroglossia”, Derrida’s “graft metaphor” (Culler 1982:134-135), Becker’s “prior texts” (1988:69), Volišinov’s (1973) “voices”. Buchstaller (2001a) refers to “mosaic”. Cf. also the related concepts of “conversational routine” (Coulmas 1981) and “routinization” (Haiman 1991, 1994).

Chapter 3 has shown that the entrance of the new variant *like* causes an important restructuring of the USE quotative system as a whole (cf. Buchstaller 2001b, Tagliamonte and Hudson 1999). The following chapter will now investigate whether we can generalise our earlier findings on US English and whether the functional niches that *like* and *go* have appropriated for themselves are identical across varieties. I will show that while the reorganisation amongst the quotative variants has generally resulted in the same functional distribution, *like* and *go* follow slightly different trajectories in the two varieties. It will become evident that there are some crucial differences in the grammaticalisation of *go* and *like* as quotatives on both sides of the Atlantic. This chapter focuses on four aspects of their systematic patterning: *go* with telicity (section 5.3.1), *go* with direct and indirect quotes (section 5.3.2.), the collocation of *like* with grammatical person (section 5.3.3.) and with other verbs of quotation (section 5.3.4).

## **5.2. Separate but equal**

In this chapter, it will be seen that both of the new quotatives *go* and *like* behave equally within the linguistic system of the two varieties. I will first discuss *like* and *go*'s behaviour with respect to the hypotheticality level of the quote they are introducing (5.2.1). Then I will talk about their distribution with respect to the epistemic level of their surround (5.2.2). Finally, I will address their co-occurrence with mimesis and exclamations (5.2.3). It will be shown that *go* as well as *like* behaves fundamentally the same on all three accounts in the two varieties.

The fact that *like* and *go* have similar status within the cohort of quotative devices and their comparable distribution point to a grammaticalisation along the same lines in both varieties. This ratifies the cross-variety comparison of *go* and *like*.

### **5.2.1. The Co-occurrence with Levels of Hypotheticality**

In chapter 3, we saw evidence that in US English the different variants of the (say) variable pattern differently with respect to epistemic stances. In this section, I

will explore *like* and *go*'s patterning with hypotheticality levels across varieties and I will investigate whether the quotative system has been adjusted functionally in the same way in the two varieties to accommodate the intrusion of the two new members.

I will again consider the three categories, realis, hypothetical, and situational, which can be conceptualized as stretches of the continuous, bi-polar cline of epistemicity. Realis quotes are situated at one pole, situational quotes at the other, with hypothetical ones filling the intermediate sphere and spanning various degrees of probability. We find that *go* and *like* in both Br and USE have a similar tendency to co-occur with realis (for *go*) and as a wildcard (for *like*).

**Table 5.1: Frequency of distribution of *go* (in %) with degrees of hypotheticality by variety**

	realis	hypothetical	situational	SUM
go BrE	54	44	2	238
go USE	63	33	4	73

ANOVA post hoc (Scheffé) :  $go_{BrE-USE}$ :  $p = .288$

Table 5.1 shows that, generally, *go* is used in BrE as well as in USE for introducing realis quotes (54% and 63%). It can also be used for hypothetical quotes (with 44% and 33%) but it is not used very much for situational quotes (only 2% and 4%). This equivalent functional load is underlined by the results of a multivariate ANOVA test, which was performed on the two quotative systems. Post hoc tests, which reveal which pairings of variants are significant, showed that the difference in the use of *go* with hypotheticality levels is not significant across varieties ( $p = .228$ , n.s.).

When we look at the frequency-based patterning of *like* by variety, the same pattern emerges in both USE and BrE.

**Table 5.2: Frequency of distribution of *like* (in %) with degrees of hypotheticality by variety**

	realis	hypothetical	situational	SUM
like BrE	22	59	19	80
like USE	27	39	34	106

ANOVA post hoc (Scheffé):  $like_{BrE-USE}$ :  $p = .791$

Across the board, *like* can be used with all hypotheticality levels but it is most prevalent with hypothetical quotes (see Tagliamonte and Hudson 1999 for equivalent results). There is a slight difference in patterning in that *like* is used more frequently with situational quotes in USE (34% vs. 19%) and more frequently with hypothetical quotes in BrE (59% vs. 39%). But in both, BrE as well as USE, we witness the wildcard effect, viz. no commitment to any level of hypotheticality and no low values in any category. Fundamentally, the use of *like* is the same in the two varieties. As with *go*, an ANOVA post hoc test yields no significance for its distribution according to variety ( $p=.791$ , n.s.) and underlines the claim about *like*'s equivalence across varieties. Hence, we have statistical evidence that *go* as well as *like* is functionally equal in both locales when it comes to their co-occurrence with hypotheticality levels.

By way of comparison, consider the co-occurrence of *say* and *think* with hypotheticality levels. Table 5.3 gives an overview of how all verbs co-occur with all levels of hypotheticality in the two varieties:

**Table 5.3: Frequency of occurrence of  $V_{quot}$  with level of hypotheticality per variety**

	realis		hypothetical		situational	
	USE	BrE	USE	BrE	USE	BrE
like	27	22	<b>39</b>	<b>59</b>	34	19
go	<b>63</b>	<b>54</b>	33	44	4	2
say	<b>70</b>	<b>71</b>	24	26	6	3
think	1	2	<b>83</b>	<b>89</b>	16	10

$F_{BrE - US \text{ hypotheticality}} (1, 1363) = 0.2, p = .888, n.s$

As expected, *say*'s primary locus of occurrence is the framing of real occurring quotes and this, again, in both varieties (70% and 71% respectively). And *think*, in USE as well as in BrE, is used for inner monologues, which feature the category hypothetical quotes (83% and 89%). But note that it also functions as a way of introducing situational quotes (10 % and 16%).

All quotatives, including the new ones, *like* and *go*, pattern in their own way with respect to hypotheticality levels. Importantly, this functional allocation across different verbs of quotation looks the same in BrE and in USE: ANOVA tests show the difference to be non-significant. Overall, we witness a similar task sharing amongst the members of the (say) variable. Hence, after the entry of the new



quotatives, the same re-functionalisation appears to have taken place in both systems. Every quotative variant has found its own functional niche and this is the same across varieties.

The above discussion has shown that we can indeed make cross-varietally valid claims about the patterning of quotative variants. Generally, we can show that different quotative options are correlated with different hypotheticality levels and that they are by that feature associated with quotes of a certain epistemic value. As in Chapter 3, I will now transfer this notion to the concept of the epistemicity of the surround, whereby expressions with a certain hypotheticality level attract/co-occur with other expressions of the same epistemic stance.

### 5.2.2. Priming

Again, I tagged every quotative frame with respect to the expression of mental verbal or ambiguous expressions in the surround (which was determined as the preceding and following 5 TCUs). Table 5.4 gives the cross-variety results for all quotative variants.

**Table 5.4: Co-occurrence of quotatives with mental/speech expressions (in %)**

	USE				BrE			
	mental	speech	ambig	not	mental	speech	ambig	not
like	28	21	22	29	34	36	19	12
go	16	<b>36</b>	21	27	19	<b>44</b>	31	6
say	38	<b>53</b>	3	6	20	<b>68</b>	7	5
think	<b>68</b>	21	2	9	<b>47</b>	35	6	13
	$\chi^2(9): 136.867, p < .001$				$\chi^2(9): 227.801, P < .001$			

T-tests:  $like_{BrE-USE}: p = .196$ ,  $go_{BrE-USE}: p = .068$ , n.s.

$say_{BrE-USE}: p < .001$ ,  $think_{BrE-USE}: p < .01$ ,  $unframed_{BrE-USE}: p < .001$

We see that the noncommittal stance of *like* in its general environment holds in BrE as well as in US English. *Like* can occur in a surround characterised as speech, thought, or it may be ambiguous verbs or in a context unspecified for repeatedly occurring speech at all. While it is true that in BrE *like* is used more frequently with

mental and speech environments (and less in ambiguous and 'not'-contexts), the fact that there is no single favouring surround in its distribution holds here as well.

Also, the direct association with *go*'s high realis levels, as shown in Table 5.3 (63% and 55%), is expressed on a distributional level in a preference for speech surrounds. We notice in both varieties a clear peak in the correlation between the occurrence of *go* and speech environments (36% and 44% respectively). *Go* prevalently occurs in an 'envelope' of speech. *Say*, is associated generally with verbal activity (53% and 68%). Notice that *say* is much less surrounded by ambiguous verbs (with much lower percentages than *go* or *like*). *Think*, the introspective quotative, is mainly surrounded by verbs of thought, much more so in USE (68%) than in BrE (47%) but clearly in both varieties. Also, it does not occur much with ambiguous contexts in both varieties.

T-test results show that the patterning of *like* and *go* in the two varieties is non-significantly different (with  $p = .196$ , and  $p = .068$ ). Hence, the two new quotatives occur in the same surrounds in both varieties; their distribution with mental or speech environments is the same in both varieties. Note that that the functional difference for *say*'s and *think*'s patterning per variety came out significant. But looking at the table I think it becomes obvious that the main locus of occurrence for *say* in both varieties is speech environments (and even more so in BrE, hence the significant difference). *Think* mainly patterns with mental environments (with an even stronger tendency to do so in USE, hence significant).

The main point to note here is that, as with the hypotheticality level of the quote, the new quotatives pattern in an equivalent way in BrE and in USE.

### 5.2.3 Verbs of Quotation, Mimetic Effect and Exclamation

In the following sections I will investigate the patterning of all quotative variants with respect to mimesis, such as voice and sound effect and gesture, as well as with respect to exclamations. Table 5.5 gives an overview of the co-occurrence of mimesis with the most important quotatives in US American English and British English. The numbers in this table are percentage values out of all tokens of the particular quotative (i.e. 69% of all of *like*'s tokens frame mimetic enactment in BrE).

**Table 5.5: Co-occurrence of quotatives with Mimesis**

	mimesis	
	BrE	USE
like	<b>69</b>	<b>75</b>
go	<b>74</b>	<b>73</b>
say	43	60
think	47	<b>75</b>
unframed	<b>72</b>	59

$\chi^2$  (df 18):205.55 , p<.001       $\chi^2$  (df 18): 132.7, p<.001

T-tests: *like*<sub>BrE-USE</sub>: p=.356, *go*<sub>BrE-USE</sub>: p=.616, n.s.

*say*<sub>BrE-USE</sub>: p<.001, *think*<sub>BrE-USE</sub>: p<.001, *unframed*<sub>BrE-USE</sub>: p<.001

The patterning of quotatives with mimesis is highly significant in both varieties as is shown by the  $\chi^2$  statistics. A t-test was run to see if the quotative options pattern differently with respect to mimesis in the two varieties. The frequency distribution shows that both varieties converge in the fact that *go* and *like* are amongst the main mimesis introductory items (69% and 75% for *like* and 74% and 73% for *go*). But the results show that the functional difference for *like*'s patterning and *go*'s patterning with respect to variety is not significant. *Like* and *go* in USE as well as in BrE are used to introduce voice effects. Tagliamonte and Hudson's (1999) results underline this claim and show that *like* and, to an even greater extent, *go* are both used with what Tagliamonte and Hudson call "inarticulate sounds" in British as well as in Canadian English.

But Table 5.5 also shows that the members of the quotative pool which take on mimesis-introductory function are not necessarily the same across varieties. In BrE, mimesis co-occurs most frequently with *like*, *go*, and unframed quotes (69%, 74%, and 72% respectively). But in USE *like*, *go*, and *think*-framed quotes most frequently introduce mimesis (75%, 73%, and 75%). The difference in mimesis-framing function is significant across varieties for *say*, *think*, and unframed quotes.

Hence, the frequency distribution gives evidence that different quotatives occupy different functional niches (cf. the same effect with hypotheticality levels in section 5.2.1). Presumably, after the intrusion of two new quotative options, functional re-allocation takes place in both varieties and functional tasks such as the introduction of mimesis are re-distributed over the whole quotative cohort. There is statistical evidence that *like* and *go* have taken on the same functional load – viz. they are mainly used to introduce mimesis – in both varieties. Other members of the

cohort behave differently across the two varieties and have taken on different functions in different varieties. Overall, we can see that the association between mimetic enactment and quotative introducer is conventional, dependent on the variety.

Let me briefly show that *like* and *go*'s functional equivalence in their correlation with mimesis is underlined by their patterning with exclamations such as *oh*, *wow*, *oops* in USE and BrE. Exclamations, which have been termed “quotable gestures” (Kendon 1992), “meaningful noises with no conventional spellings” (Ferrara and Bell 1995:282), or “response cries” (Goffman 1981) are usually drawn from 2 sources – mild taboo words and non-word vocalizations – and are used to dramatise the situation and to create listener involvement (Tannen 1989:25-26). Ferrara and Bell draw attention to the earlier occidental tradition of drama which often featured soliloquies –revelations of inner states in order to cue the audience into the mental landscape of the characters. While lengthy introspective sequences in the form of soliloquies have gone out of fashion (Goffman 1981), Carbaugh (1988) claims that they nowadays find their continuation in the (American) tendency for self-revelation, which finds its epitome in shows like Donahue, Oprah or Jerry Springer. These media-manifestations of the inner self abound with dramatic re-enactments of personal experience, most of them in the form of narrative (Labov and Waletzky 1967). This creates a need for a frame that can introduce what was going on in the mind of the actor and present it as if the participant was “in a state of talk” (Goffman 1981). Ferrara and Bell (1995) found that *like* had filled the slot of a conventionalized theatrical device which marks revelations of inner state to the audience. Similarly, it seems clear that quotative *go* emerged from an earlier stage as a mimesis marker (OED 1791 *my noble heart goes “pit a pat”*). Consequently, we would expect *go* also to feature highly in the category ‘framing exclamations’.

For the purpose of this study, I define exclamations as non-referential, non-lexical expressions (cf. Goffman's 1981 response cries, “exclamatory expressions which are not fully fledged words”). Like Cameron (1998), I include laughter and weeping in this category because they share the element of emotional distress or intensity associated with the use of interjections. Unsurprisingly, the same quotatives

that have been shown to introduce mimesis most frequently co-occur with exclamation.

**Table 5.6: Co-occurrence of quotatives with Exclamation**

	exclamation	
	BrE	USE
like	<b>29</b>	<b>33</b>
go	<b>40</b>	<b>40</b>
say	15	17
think	<b>29</b>	<b>35</b>
unframed	<b>26</b>	16

$\chi^2(4)$ : 88.211,  $p < .01$ ,       $\chi^2(4)$ : 38.102,  $p < .01$

T-tests:  $like_{USE-BrE}$ :  $p = .473$ ,  $go_{USE-BrE}$ :  $p = .912$ ,  $say_{USE-BrE}$ :  $p = .338$ ,  $think_{USE-BrE}$ :  $p = .231$ , but  $unframed_{USE-BrE}$ :  $p < .01$

Table 5.6 shows that, in both varieties, *like* and *go* are the main items for introducing exclamation (at 29% versus 33% and 40% versus 40%). This distributional evidence underlines the findings for mimesis. *Go* even has the highest percentage of occurrence in both (40%). In USE, *think*, which has been already identified as strongly associated with mimesis (cf. Table 5.5), also occurs with high frequency with exclamations. In BrE, it is again unframed quotes (cf. Table 5.5), and also *think*, that are strongly associated with exclamations. Hence, *think* in BrE, while not necessarily associated with sound and voice effects, is nevertheless one of the quotatives that have a high frequency of co-occurrence with exclamations.

Again, we witness task-sharing amongst the quotative options. Different verbs of quotation take on specific functions, here the introduction of exclamations. But we notice that the allocation of quotatives into functional niches is a variety internal matter. Such a re-distribution can and indeed does yield a different pattern in different varieties. The case of *think* has already been pointed out. We also note that the unframed variant is much more associated with the introduction of exclamations in BrE than in USE. T-tests show that all lexical verbs pattern non-significantly across the two varieties but that the functional load of unframed quotes again differs significantly across varieties. Also, the functional difference between *go* and *like* across varieties is not significant. This means that *go* and *be like* have assumed the same functions both in USE and BrE.

Comparing these results with the variants' patterning with respect to mimesis (Table 5.5) shows that the new quotatives, *like* and *go*, have assumed the same functions in both varieties. Other quotative variants, which had to be re-functionalised after the enlargement of the quotative pool by two options, sometimes occupy different functional niches across varieties. With respect to mimesis, *think*, unframed, and *say* pattern significantly differently and, with respect to exclamations, unframed quotatives show a significant effect. Functional reallocation does not automatically proceed along the same lines and hence does not necessarily lead to the same distributional patterning in both varieties. This makes the finding that *like* and *go* show identical patterning in both varieties even more important.

We are now in a position to assert such sameness for *like* and *go*. Initially, I stated that when looking at a potential variant, it is important to answer the questions of 'what is it doing? (what is its functional load)' and 'where is it occurring? (in which contexts)'. Only if the results show that those are comparable in the most fundamental sense can we speak of the variant as being the same thing across varieties.

We have now seen evidence that *go* and *like* have the same functional patterning in two different varieties. *Like* as well as *go* has the same distributions with regard to the epistemicity of their quotes and the contexts they occur in, both in BrE and USE. In both cases, the quotative system has been re-structured so that different quotatives take on specialised functions, i.e. occupy functional niches. It has been established that mimesis is associated with *like* and *go* in both varieties and also with *think* in USE, but with unframed quotes in BrE. Research into the distribution of quotatives with exclamations has yielded similar results.

In her criticism of Labov's criterion of semantic equivalence, Lavandera (1978, cf. also Dines 1980 and Romaine 1984) came to the conclusion that functional equivalence is a much better criterion for establishing sameness than semantic equivalence. In this chapter, I have demonstrated that both function and distribution with respect to the key of the surround help us decide whether we can assume that something is to be taken as the same thing. The mere fact that the surface looks the same for *like* and *go* in different local varieties does not mean that they are actually

grammaticalizing as the same thing. After having proven with statistical results that we are actually comparing same with same, that *go* as well as *like* behaves fundamentally the same in the two varieties, we are at the point where we can conduct a variationist study and investigate the variants' patterning with respect to linguistic, and external, social variables. I will now investigate whether there are regional differences with respect to the new quotatives and if so what these differences are. The remainder of this chapter is a study of their intralinguistic constraints. Chapter 6 presents an investigation of their patterning with respect to extralinguistic factors.

### 5.3. The Crucial Differences

This section shows that when we look at the linguistic details of *like* and *go*'s patterning, we notice that while they are generally comparable, there are nevertheless indications that they grammaticalise independently in different varieties. In the following sections, I will give four key differences. The fact that we have two slightly different surface outputs means that we have to consider the developments in the 2 varieties to be autonomous to at least a certain degree. I will conclude that

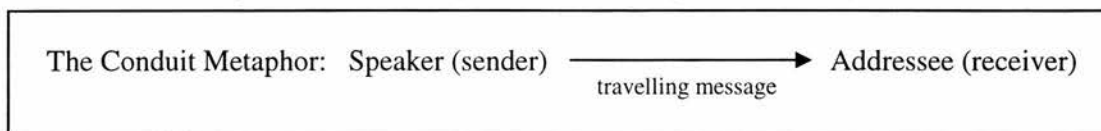
- a) due to their overwhelming functional and distributional equality and in spite of minor differences in their patterning, *like* and *go* are to be considered the same variant across varieties, and
- b) we cannot speak of the grammaticalisation of *like* and *go* as a process in English per se because we witness divergent grammaticalisation patterns in different regional varieties.

#### 5.3.1. Telicity with *go*

Chapter 4 has shown that *go*'s metaphorical extension from a verb of movement to a quotative is made possible by the fact that it encodes directional/telic movement (cf. Buchstaller in preparation). Figure 5.1 (below, as reproduced from Chapter 4, cf. also Fillmore 1997, Foley 1997, Lakoff 1987, Lakoff and Johnson

1980) depicts the communicative act as a travelling message from the speaker (the sender) to the addressee (the receiver).

**Figure 5.1: The metaphorical base for *go*'s use as a quotative**



Hence, we can conceptualise communication as the encoding and sending of a message by a speaker to a potential addressee. An addressee is a communicative goal, an interlocutor to whom speech is directed and who can be encoded into the quotative frame as a participant to whom speech is directed.

The prototypical sequence for a quotative frame is the sequence SPEAKER V<sub>QUOT</sub> “QUOTE” without such an overt encoding of an addressee (cf. example (5.1) for BrE with *go*).

(5.1) BrE - *Coming home from work late*

X: whenever she's tired she'll get ratty [won't she,  
 Z: hh hih[hi  
 X: [and irritable,  
 → and she'll **go** “get me a cup of tea,  
 I've been at work all day”,

This sequence holds for both varieties in English, BrE and USE, as well as for any quotative such as *think*, *say*, *scream*,... But there is also the non-prototypical quotative frame, which has the form SPEAKER V<sub>QUOT</sub> ADDRESSEE “QUOTE”. Here, the addressee is explicitly encoded as a participant in the quotative frame. This sequence is exemplified for BrE in (5.2) below with a quotative framed by *say*:

(5.2) BrE - *The new car*

A: and he was just so,  
 Mandy was around the other night when we phoned Tom during yeah,  
 B: uh huh,  
 → A: and he was **saying to Mandy** “I took her show around in the car,  
 I know she doesn't like it,  
 but tough”.

In (5.2), in the line marked by the arrow, we have the sequence SPEAKER (*he*) V<sub>QUOT</sub> (*was saying*) ADDRESSEE (*to Mandy*) “QUOTE”. The surface addressee is encoded in



the PP *to Mandy*. While this sequence occurs with most quotatives in both varieties, the crucial difference is that **only** in BrE do we have *go* in this structure with a surface addressee. The sequence SPEAKER V<sub>QUOT</sub> (*go*) ADDRESSEE is exemplified with a 1<sup>st</sup> person addressee in (5.3), where the surface addressee is encoded in the PP *to me*. Example (5.4) contains a 3<sup>rd</sup> person surface addressee, *to the police*.

(5.3) BrE - *Jogging*

- A: she goes “did you you ever went jogging”,  
I said “yeah”,  
B: of course you’d heard,  
→ A: and she’s **going to me** “w-well will you speak to her today”,  
and I went “well yeah”,

(5.4) BrE - *Youth today*

- X: and I mean my brother was terrified in case the police had said something you know,  
Y: uh huh,  
X: had said something to me mum and me dad,  
→ Y: I mean nowadays they would just **go to the police** “ehhhh”.  
you know,  
Y: yeah they think nothing of it,

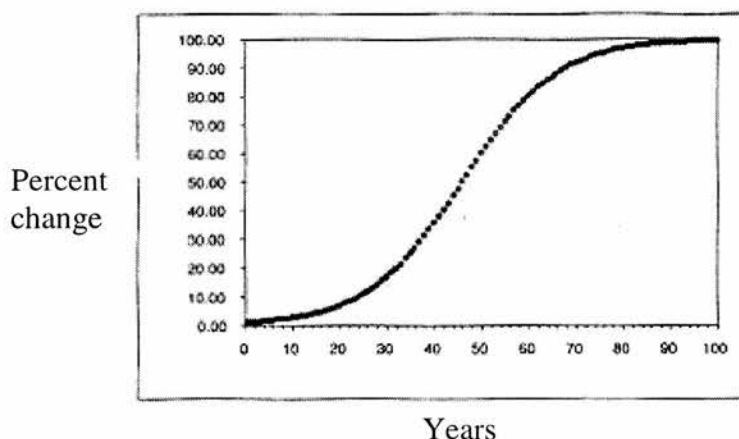
For (5.4) the previous context (which I have not included due to its length) makes clear that the addressees *the police* were already present at the moment of the quote. This excludes an interpretation whereby *go* has the function of expressing telic movement of the speaker towards the police. Hence, the *go*-quotative frame hosts an expressed surface addressee *to the police*.

What the two BrE examples show is that *go* as a quotative does occur with an expressed addressee in the frame, a sequence which is not attested in the USE data (0 tokens out of a total of 234 *go*-framed quotes). Hence, the structure SPEAKER *go* “QUOTE” is possible both in BrE and in USE, whereas the structure SPEAKER *go* ADDRESSEE “QUOTE” is possible in BrE but **not** in USE. Altogether, the BrE data contain 4 tokens out of 291 quotes framed with *go* (1.37%). Note that the tokens are not from the same speaker but have been produced by 4 different speakers, 3 from Derby and 1 from Newcastle. In three cases the addressee was expressed as a PP with *to* and once a bare NP. The direct encoding of such an addressee in the quotative frame as a PP or NP means that telicity is overtly expressed as an argument. A comparative survey of all of *go*’s tokens suggests that the two varieties differ with respect to whether telicity can be overtly expressed or not. Hence, the

data suggest that telicity can grammaticalise differently during the emergence of *go*'s quotative function. In BrE, *go* is grammaticalising as a quotative which can take an overtly expressed/surface addressee. In USE, *go* appears to have co-occurrence restrictions. It does not take an overt surface addressee. We witness different patterns of grammaticalisation for *go* in different varieties.<sup>3</sup>

A short note on the small numbers involved in this and the following sections: during the initial stages of grammaticalisation, we necessarily expect low numbers of tokens of the linguistic phenomenon in question. Changes start off slowly and only acquire impetus as they are picked up by the speech community. Labov (2001:450) gives the estimated curve of a change in progress, replicated here as Figure 5.2:

Figure 5.2: 100-year progress of a change with incrementation (from Labov 2001:450)



The low numbers of tokens given in this section may not show any statistical significance, but what they may represent is the initial appropriation of new contexts during the early stages of grammaticalisation. In this respect, this section is more of a qualitative nature. The importance of the numerical difference between the varieties (albeit small) lies in the fact that one integer is consistently zero. Whereas the new quotatives have appropriated themselves a certain linguistic context in one variety, they have not done so in the other.

<sup>3</sup> In order to check whether *go* as a quotative verb is generally grammaticalising differently with respect to the encoding of telicity in the two varieties, I conducted an investigation into the occurrence of *go* in structures where it is ambiguous between a spatial telic verb and a quotative. Ambiguity exists in constructions such as SPEAKER GO "QUOTE" PP/ ADV<sub>SPACE</sub>, as in

X: Simon went down the hill on the way to college,  
**went** "bang" **into** the back of a van.

### 5.3.2. Indirect quotes with *go*

American authors have pointed out that *go* occurs only with direct quotes. When Butters first noted the occurrence of *go* as a quotative verb in USE in his much-cited 1980 article, he wrote “so far, it does not seem to have spread to indirect discourse ... though (...) [that] might come in due time”. (cf. also Blyth et al. 1990, Cukor-Avila 2002, Romaine and Lange 1991, Schiffrin 1981, Tagliamonte and Hudson 1999). Schourup writes that *go* “... unambiguously cues listeners to the onset of a direct quotation [so that it] (...) neatly solves a small problem in spoken English” (1982b:149-150). The problem-solving property he refers to is the fact that, whereas *say* and other verbs of quotation are often ambiguous with respect to the enquoting of direct or indirect reported speech, *go* apparently always enquotes direct reported speech.

To recapitulate, for the purpose of this thesis, I define direct quotes (cf. Chapter 2 for a more detailed definition) as representation of past speech or thought which have a deictic orientation to the experiencer of the quote in spatial, temporal and personal orientation. I have not taken as a distinguishing characteristic for a direct quote the absence or presence of voice or sound effects (Klewitz and Couper-Kuhlen 1999) or of whether or not there is a complementiser *that*.

Having defined direct quotes by those criteria, we can investigate *go*'s patterning in the two varieties. Looking at the USE corpus, we find Butters and Schourup et al. ratified. There are no indirect quotes with *go* (0 out of 234 tokens). But in the BrE corpus, we find instances of quotes like example (5.5):

(5.5) BrE - *Spooky night*

X:       and the first thing that went through me mind was,  
          I was telling Mandy was ehmmmm,  
          when Angela stayed that night,  
→       and **she goes** .hhh she heard a noise downstairs,  
          someone speaking.  
          and I thought “flipping hell,  
          who is that”?

---

I found no token in US English (with 234 overall) and 3 tokens in BrE (1.03 %). This confirms earlier results in which the telicity of *go* is overtly grammaticalising only in BrE but not in USE.

In (5.5), the TCU marked by the arrow contains the quote (.hhh) *she heard a noise downstairs someone speaking*. The 3s reference *she* marks the agent as a person *other than* the experiencer – otherwise it would be the 1s pronoun, *I*. Angela is clearly the person referred to by both tokens of *she*. Hence, the deictic orientation of this utterance is to the speaker relating it, not to the experiencer. The reported stretch is an indirect quote. Another possibility is to interpret this construction as a structural break. According to this interpretation, X, who is telling a story about Angela, starts with *she goes* (movement or quotative *go*). Immediately after *go*, X breaks off and starts a new construction with *she heard a noise downstairs*. This new construction has no structural connection with *she goes*, thus cannot be interpreted as subordinated under quotative *go*.

There is a slight breathing pause after *she goes*, which could be interpreted as a sign of reorientation followed by a different structure. But due to the fact that we have none of the typical signs of an anacoluthon, such as stutter, lengthenings, extended pausing, fillers, repairs, etc., I have not analysed this as an example of a structural break. I take this stretch of speech to be an indirect quote framed by quotative *go*, an interpretation which is underlined by a continuous intonation pattern.<sup>4</sup>

*Go*, an item which can have quotative function, is occurring directly before an indirect quote (.hhh) *she heard a noise downstairs* and can be interpreted as framing the reported utterance. *Go* thus functions as an introductory item for indirect reported speech.

Overall, my BrE corpus contains 4 tokens of indirect *go*-framed quotes out of 291 *go*-quotes (1,37% versus none at all in USE). Even though there are problems in using negative evidence (Lass 2000, Lindstrom 2001), I conclude that while *go* as a quotative for direct speech occurs both in USE and in BrE, it **can** be used to introduce indirect quotes in BrE but **not** in USE. While *go* can take on the function of framing indirect speech in British English, *like* certainly does (still) have this disambiguating function and never occurs with indirect speech in both varieties.

---

<sup>4</sup> I would like to thank the participants in the Language in Context group at the University of Edinburgh for their helpful comments and suggestions concerning this analysis.

Having shown that *go* patterns independently in the two varieties with respect to indirect enquoting and telicity, I will now give evidence that *like* as well, has slightly idiosyncratic behaviour on both sides of the Atlantic. Far from directly mirroring the developmentally earlier American surface item, British speakers use *like* in a distinctive fashion.

### 5.3.3. The Collocation of *like* with other verbs of quotation

In this section, I will focus my investigation on collocations of *like* with quotative introductory items leading to sequences of the form SUBJECT V<sub>QUOT</sub> LIKE “QUOTE”. The co-occurrence of *like* with quotative verbs has been commented on in the literature as “mixed forms” (Macaulay 2001)<sup>5</sup> or “transitional forms” (Singer 2001). In my data, *like* is attested as co-occurring with an array of quotatives, *say like*, *think like*, *tell like*,.... Example 5.6 shows a case of the collocation of *like* with *go* in BrE:

(5.6) BrE - *The barbecue*

- Y:     cause eh-hm I was talking to him you kn- you know Mandy’s barbecue at XX,  
          well ahmm he was tell- he was **going like** “I don’t know anybody”,  
          and all this-
- X:                    -did he go?

The collocation of two items with quotative potential, *go like*, here frames the reported speech sequence *I don’t know anybody*. There is no noticeable pause, stutter or any other sign of disfluency which would speak for a repair of one quotative (viz. *go*) for another (viz. *like*). This leads to the interpretation that *go like* together take the function of introducing a quote. An investigation of the distribution of such collocations in the two corpora yields an interesting pattern:

---

<sup>5</sup> While Macaulay (2001) refers specifically to quotative frames with the form *go like that* “quote”, I will not concern myself with the deictic here.

**Table 5.7: Collocations of *like* with quotatives across varieties**

BrE	N	USE	N
say like	14	say like	2
think like	7	think like	3
go like	4	tell like	1
feel like	3	feel like	76
Total:	28		82
% of all quotes framed by ( $V_{quot}$ ) <sup>6</sup> <i>like</i>	23.1%		40.4%

Table 5.7 shows that, in BrE, *like* co-occurs with all sorts of quotatives but mostly with *say*. Fourteen tokens (50%) out of all 28 collocations are combinations of *say* and *like*. In USE *like* enters much more frequently in collocations with quotative verbs (35.9% vs. 18.8% in BrE), but it is specialized with *feel*. Seventy-six of all 82 tokens are collocations of this sort, i.e. 92.7% out of all combined frames. The following 2 excerpts exemplify the prototypical collocations of *like* in the two varieties, first the combination with *feel like* in USE (5.7), then *say like* in BrE (5.8).

(5.7) USE - *Planting*

- B: I've got a few plants here,  
but I'm not really knowledgeable.  
I I I feel real good if I water them and they continue to grow,  
you know I **feel like** "oh I've accomplished something",  
A: uh huh

(5.8) BrE - *The deja-vu*

- X: he thought I was mad,  
he **said like** "what are you going on about"?  
and I was like "oh I know that was scary .hhhhhh  
I didn't like that",

Those examples show how speakers use *like* with quotatives associated with the verbal as well as with the mental domain. It is regularly combined both with verba dicendi that are associated with real occurring and with hypothetical speech. In fact, different varieties favour one or the other collocation (*feel like* versus *say like*). This is not surprising given the wildcard status of *like*. As I have shown in earlier sections, it is not committed to any epistemic stance. It can introduce quotes of all hypotheticality levels (Chapter 3.4), and it can co-occur in contexts of expressions of

<sup>6</sup> The parenthesis indicates that  $V_{quot}$  is an optional item.

verbal and of mental activity (Chapter 3.5). These findings underline that *like* has few co-occurrence restrictions and collocates relatively freely with quotative verbs. It just sediments out, i.e. grammaticalises, differently in different varieties.

### 5.3.4. Verb-person correlation for *like*

In Chapter 4, I investigated *like*'s occurrence with existential *it* and the role this collocation might have played in the grammaticalisation of *like* as a quotative (note also the reservations about the incompatibility of my results and earlier studies). Example (5.9) gives another illustration of this sequence:

(5.9) USE - *mood changes*

- A: and the people their moods are extreme you know,  
 B: right,  
 A: **it's like** "God if I don't get to see the sun I'm going to be in this bad mood all this=  
 =entire time" I mean,  
 and this is how I see people reacting.

Recall Tagliamonte and Hudson's (1999:23) footnote, which reads as follows: "interestingly, although representing a very small proportion of the data, N=23 (3.4%) (*viz.*), the most highly favoured context for *be like* in Britain is with existential *it* ...". If "existential *it*" is the most highly favoured context, why does it only account for 23 out of 120 *like*-quotations?

Indeed, in my corpora, *it*-quotations constitute the most favoured context. Certainly, as Tagliamonte and Hudson point out, the question of the occurrence of *like* with *it* is worth investigating from a cross-variety perspective. Table 5.8 gives an overview of *like*'s patterning with grammatical person in both varieties.

**Table 5.8: Verb-person Correlation for *like* in both varieties (in %)**

	1st	2nd	3rd	it	not
LIKE BrE	<b>40</b>	4	19	34	1
LIKE USE	29	3	13	<b>56</b>	0

$$F(1, 26.474) = 15.973, p < .001$$

We see significant differences ( $p < .001$ ) across varieties in the use of *like* with respect to *like*'s occurrence with 1<sup>st</sup> person, 3<sup>rd</sup> person and existential *it*. Notably, US

English speakers use *like* 56% of the time for impersonal *it*-quotations in comparison with only 34% for British English speakers. The most frequent environment for *like* in BrE is 1<sup>st</sup> person (40%) whereas in USE it is most frequent with impersonal *it*. The higher *it*-frequency in US English patterns with lower frequencies of 3<sup>rd</sup> person personal and especially of 1<sup>st</sup> person marked quotes (13% USE versus 19% BrE and 29% USE versus 40% BrE).

Whether this different distribution is due to different narrative styles, as hinted at in Tagliamonte and Hudson (1999), or whether the fact that in US English, *like* as a quotative is at a much more advanced stage and thus intrudes in territories that BrE *like* has yet to explore is a matter for further investigation. Without any time-depth study that is based on data that includes impersonal *it*, it is impossible to draw any conclusive results.

But in any case we see that synchronically, *like* has assumed very different distributional preferences with respect to person marking. While it is mostly used to frame first person quotation in BrE, it is first and foremost a quotative that occurs with impersonal *it* in USE.

## 5.4. Conclusions

The conclusions to this chapter are twofold. I will first talk about the descriptive findings and will then address the methodological consequences of this study.

### 5.4.1. Descriptive

I have shown that both *go* and *like* as quotatives have fundamentally the same properties in BrE and USE. *Like* as well as *go* has the same function with respect to the hypotheticality level of the quote (Chapter 5.2.1) and the enquoting of mimesis and exclamations (Chapter 5.2.3). Furthermore, in both varieties, *like*, as well as *go*, has the same distribution with respect to the mental and speech environments they occur in (section 5.2.2). We not only witness a surface resemblance but in function and distribution, both *like* and *go* are the same thing across varieties. We are now in a



position to claim that because of their form and their distribution, it is appropriate to consider them comparable variants of the same (say) variable across the two varieties. Hence, the innovative forms *like* and *go* have spread in the quotative system of two regionally separate varieties of English in astonishingly similar ways, alike enough for us to categorise *like*<sub>BrE</sub> and *like*<sub>USE</sub> and *go*<sub>BrE</sub> and *go*<sub>USE</sub> as the same variants.

But a closer look at *like* and *go*'s intralinguistic behaviour has shown that there are different *patterns* of grammaticalisation in different local varieties. Macaulay (2001) refers to "innovations or confusion in the transmission" and he points out that the grammaticalisation of the quotatives in Scottish English is not a case of the colonial lag phenomenon mentioned by Trudgill (1999). This claim is underlined by the comparative BrE-USE data presented in this study. With respect to the encoding of telicity, this chapter has given evidence that *go* can co-occur with surface expressions of telicity in BrE but not in USE. Also, *go* can frame indirect quotes in BrE, whereas in USE it appears to have co-occurrence restrictions.

Looking at *like*, I have shown that with respect to collocation with other quotatives, it co-occurs most frequently with *say* in BrE, whereas in USE, where we have a generally very high occurrence in collocations, it most often occurs with *feel*. This has led me to conclude that *like* specialises with different quotative verbs in different varieties. Also, *like*'s patterning with person marking shows distinctive patterns on different sides of the Atlantic.

Overall, *like* and *go*'s idiosyncratic developments have provided evidence that the surface output of the grammaticalisation process differs with respect to locality, a fact that points to locally independent grammaticalisation. Hence Tagliamonte and Hudson's (199:169) statement that "British and Canadian youth both are following the same functional trajectory as that found in the United States. While the more traditional variants (e.g. *say*, *tell*) have somewhat different patterns of use in British and Canadian English, the patterning for *like* is the same" is to be taken with a grain of salt. While the data makes evident that *like* and *go* pattern functionally similarly enough in British English and US English to justify calling them the same variant across varieties, a closer look at the details of their grammaticalisation has shown that they both display idiosyncratic, variety-specific behaviour. The present study has

shown that while newcomers *like* and *go* sparked a reallocation of the functional load amongst **all** members of the quotative pool in both varieties, the individual quotatives have come to occupy slightly different niches in each variety. This leads me to conclude that we cannot speak of **the** grammaticalisation of *go* and *like* as a process in English per se. Language variety is a variable that must be taken into account.

#### 5.4.2. Methodological

Sociolinguistics and grammaticalisation theory have not traditionally been closely intertwined. This chapter shows that there are real advantages to tying grammaticalisation, most notably the notion of emergent grammar, into a variationist sociolinguistic study. A combination of approaches is possible as their field of interest is the same: both look at change or variation in linguistic practices.

While variationist sociolinguistics seeks to investigate the co-variation of social and linguistic variables, and quantitatively examines language variation and change, grammaticalisation argues that language structure is (re-)emergent and that these emergent regularities are the creatively achieved outcome of verbal activity.

Methodologically, grammaticalisation tries to empirically investigate in which linguistic contexts the innovations arise, strives to find explanations of how those new structures come about, to give an account of the mechanisms involved, and to understand how the innovations penetrate the linguistic system. This has implications for sociolinguistics: according to Labov, “social and stylistic variation presume the option of saying ‘the same thing’ in several different ways” (1972a:271).<sup>7</sup> Semantic equivalence has been challenged as the proper measure of equivalence (Lavandera 1978) and subsequent definitions of the sociolinguistic variable have relied on variants having the “same function in discourse” (Dines 1980). This chapter has shown how variationist sociolinguists can take on board grammaticalisation as a means of assessing whether our variants generally – overall - function within the linguistic system in the same way. Applied this way, grammaticalisation can help us ratify and rationalise our decision as to whether or not a variant is actually “the

---

<sup>7</sup> This is paralleled by Hopper and Traugott’s statement that grammaticalisation looks at the phenomenon of speakers “find[ing] new ways to say old things” (1993:65).

same". Once we are in a position to assert functional equivalence, we have a justifiable reason for claiming that a potential variant is behaving fundamentally as the same thing. And only once this general sameness is asserted can we go on to investigate how this variant patterns with respect to extralinguistic variables as is demonstrated in the next chapter.

One final note: what can variationist sociolinguistics give to grammaticalisation? Within the framework of grammaticalisation, which proceeds primarily diachronically and cross-linguistically, languages are still very much considered homogeneous blocks. Even though we do find the occasional allusion to sociolinguistic registers (Hopper 1991, Norde 2001, Traugott and Heine 1991), variation is very much seen as an intermediary, messy state of ambiguity which precedes the clean, clear-cut next stage in a grammaticalisation pathway which is assumed to hold for a whole language. But stable variation in the sense of different patterns with respect to different social variables has not been explored by grammaticalisation studies so far (cf. Lass 2000). As Janda (2001) pointed out, within grammaticalisation, linguistic change has to be reformulated in terms of speakers and their linguistic practices. This study shows that we cannot speak of the grammaticalisation of *like* and *go* as quotatives in English per se and thereby proves that extralinguistic factors, especially the variety in question, have an influence on surface output. It also shows that there is systematicity at each stage in the process. And that this systematic variation *can* ultimately give rise to what grammaticalisation theorists have commonly idealized as different languages.

**“For each individual, his language is thus a total recreation achieved under the influence of the milieu that surround him.”**

Meillet (1929/1936:74-75)

Not all variability and heterogeneity in language structure involves change; but all change involves variability and heterogeneity. (Weinreich, Labov and Herzog 1968:188)

## Chapter 6

### Extralinguistic factors

Ever since the 1960s with the ground-breaking work of Labov (1966a, 1966b), sociolinguistic research has been concerned with finding the orderly relationship between social factors and linguistic variables. Quantitative variationist sociolinguistic research aims to express the variability of a given variable as a correlation between the variants and a set of social factors. This probabilistic co-occurrence is considered not only as indicative of group membership but also stresses the agentive and constructive nature of variable use.

Numerous studies have shown that grammatical variables tend to stratify groups more sharply than phonological variables (see Cedergren and Sankoff 1974, Chambers 1995, Wolfram 1969). This chapter investigates how a discourse or morpho-syntactic variable, such as the (say) variable patterns according to social parameters. Indeed, earlier research on (say) variants has given convincing evidence that the quotative frame is a site that can carry extralinguistic information. Yule and Mathis (1992) show that quotative choice can organize speaker's topic, Mathis and Yule (1994) give evidence of how the zero quotative expresses both interpersonal and spatio-temporal immediacy. Romaine and Lange (1991), building on the work by Nordberg (1984), show that quotative choice can heighten the dramatic impact of the story.

With *like* being a highly salient innovative option within the quotative system and progressing at full speed across the sociolinguistic spectrum in many varieties of English, sociolinguists around the globe have seen their chance to investigate the birth of a linguistic variant. The plethora of sociolinguistic studies which report on *like*'s extralinguistic patterning have resulted in a variety of (sometimes conflicting) findings which I will summarise below. Most research to date has, however, only

looked at *like* (plus sometimes the most common other quotatives) without accounting for all other variants that make up the variable (notable exceptions are Tagliamonte and D'Arcy to appear, Cameron 1998, Cukor-Avila 2002, Johnson 2001 and Tagliamonte and Hudson 1999). But, as Cameron rightly argues, we need to “close the set that defines the variable” (Labov 1996:78) in order to fully account for the variation amongst the individual variants within the variable. Hence, this chapter discusses the extralinguistic constraints of *all* variants which interrelate within the (say) variable. While I will specifically focus on *go* and *like*, their patterning will be discussed within the systemic interplay of all quotative variants.

Indeed, Cameron's plea for a full consideration of all competing variants is underlined by two findings reported here: (i) the importance of a stable denominator in statistical evaluations within one variable and (ii) the fact that the introduction of quotative *like* resulted in a functional restructuring of the entire cohort of quotative variants. This large scale reorganization gives evidence of the systematicity of the (say) variable as a whole. And just as Chapter 3 has shown linguistic reallocation / refunctionalisation of *all* variants, it will be interesting to see whether the quotative system has become restructured on the extralinguistic plane as well. But we can only fully understand *like's* spread across the social spectrum when we are looking at its interaction with all other quotative options. It follows that we will have to consider the social reality of the (say) variable as a whole.

Furthermore, previous chapters have shown that even though *like* and *go* fundamentally follow the same lines of development in both varieties, their grammaticalisation route is nevertheless locally independent. I will now examine whether this finding can be extended to their social constraints by conducting a variationist sociolinguistic investigation of the two varieties. Given the fact that the allocation of social meaning is arbitrary and highly culture specific, we would expect an idiosyncratic national patterning of sociolinguistic variants according to locality with respect to social constraints just as much as with respect to the linguistic constraints. But we will also have to admit for influence from US English – the variety where *like* was attested 12 years earlier – to British English via the mass media (which facilitate the transmission of cultural goods as well as the social meaning attached to them, see Stuart-Smith 2001). Nevertheless, as social meaning is

high context information (von Hippel 1994), we would imagine a less-than complete transmission of social constraints (Buchstaller 2003, Meyerhoff and Niedzielski 2002). Rather, it is to be expected that the transfer of a surface form goes hand in hand with reallocation and reappropriation in the new locale (cf. Chapter 5). The following pages will test this hypothesis in discussing whether *like* has the same social distribution in the UK as in the US data, the variety from which British English speakers have allegedly borrowed the new quotative (Macaulay 2001, Tagliamonte and Hudson 1999). A comparison of the results of two separate variationist sociolinguistic investigations of the quotative system in both varieties will enable us to compare the social constraints which play a role in the (say) variable in the two varieties.

This chapter is structured as follows: After a discussion of the relevant sociolinguistic literature and of previous investigations of the new quotatives, I will give an overview over *like* and *go*'s social constraints in the two data sets under investigation. In the first sub-section, the correlation of social categories and quotatives in the American English data will be described. Here, *like* has already spread considerably across the social spectrum and had had the time to expand intralinguistically by the time of data collection (1988-92). Then, the patterning of the quotative system in British English, where *like* was at an incipient stage at the time of data collection (1994/5), will be investigated. I will point to crucial differences in the patterning in the two varieties and will give possible explanations why.

It will become evident that there are not many significant social effects with regard to *like* and *go* in British and US English in my data. Hence, while the previous chapters have given evidence of *like* and *go*'s important intralinguistic constraints, their extralinguistic profile turns out to be much less remarkable than their linguistic constraints. This finding complements a growing body of research which, looking at both intra- and extralinguistic constraints, has found linguistic factors to be much more important and robust than social factors (cf. Bell 1984, Fought 1999, Labov 1993, Meyerhoff 2000, Preston 1991, Rickford and McNair-Knox 1992, but see also Smith and Steele 2003). I will argue for a synthetic approach to variationist sociolinguistics. We need to account for social as well as linguistic factors that

constrain the variability of our (say) variable, as much as any other variable, in order to be able to understand as fully as possible the reality of the variable in question.

After giving an overview of the few social constraints that have an effect on the new quotatives, I will look into the sociolinguistic distribution of all members of the pool of quotative strategies in more depth, especially with respect to the interaction between the non-canonical variants, *go*, *like* and unframed quotes. I will discuss their patterning across apparent and real time.

## 6.1. Literature overview

In an important observed regularity of linguistic change, Milroy and Milroy (1992) show that in modern urban societies, members of different social classes tend to have very dissimilar social networks: middle class speakers exhibit less closely-knit networks with social ties outside their immediate neighbourhoods and families. Working class speakers, on the contrary, tend to be less geographically mobile with closer social ties. The Milroys show that this difference is directly reflected in their speech with the middle class speakers producing less localised and more standardised variables than the working class speakers.

Additionally, L. Milroy (1980) found a weakening of gender-effects on linguistic variables within the less locally oriented open network of the middle class speakers. Hence, in the present investigation, we would expect a bigger gender effect amongst the working class speakers for class-stratified quotative variables. Chambers (1995: 128, et passim) further adds that “loose-knit network structures also found in the MC go together with the blurring of gender roles ... [and] yet we consistently find the same linguistic differences between men and women”.

Labov (1990:214) comments that “women ... are said to be more expressive than men or use expressive symbols more than men or rely more on such symbols to assert their position” (cf. also Sattel 1983). This is underlined by findings that “female linguistic behaviour is viewed more tolerantly by local peer-groups, so that women have, in a sense, more linguistic freedom than men” (J. Milroy 1981:37). We would expect this higher degree of (socially permitted) linguistic creativity to

manifest itself in the overall higher use of non-traditional variants by women. And indeed, while we do find variationist studies that show men in advance (Labov's 1963 Martha's Vineyard study, Belfast as described by Milroy and Milroy in 1978, in Norwich, Trudgill 1972), women have been shown to be at the forefront in many studies on linguistic change in progress whether in New York City (Labov 1966b), Philadelphia (Labov 1984) in Panama City (Cedergren 1973), in Hong Kong (Bauer 1986) etc. In fact, differential statistical output according to gender is one of the most commonly cited diagnostics of language change in progress with females generally in the lead (cf. Labov 1972a, Labov 1990, Ladegaard 1998, Romaine 2003). Young women are consistently cited as the main innovators (cf. Eckert 1989, Eckert and McConnell-Ginet 1999, Labov 2001).

This apparent contradiction between women leading the way and women lagging behind has been pointed out by numerous researchers. Eckert (1989:248) calls for caution against sweeping gender-related statements and reminds us that "not only is it a mistake to claim that women are more or less innovative than men, but at this point in our research it is a mistake to claim any kind of constant constraint associated with gender". And Wolfram and Schilling-Estes (1998:187) spell out the conundrum: "women appear to be more conservative than men, in that they use more standard variants ... At the same time, women appear to be more progressive than men, because they adopt new variants more quickly".

Labov (2001) formulates this situation, whereby the same group of speakers is at the same time conservative as well as progressive, as the "Gender Paradox ... Women deviate less than men from linguistic norms when the deviation is overtly proscribed, but more than men when the deviations are not proscribed". However, as Eckert and McConnell-Ginet (2003:293) rightly point out, this only amounts to a paradox if we try to attach the same social meaning to non-standard language and use of innovative forms. Indeed, as sex and linguistic practices do not simply map up but because speakers can choose to construct particular aspects of their identity via (amongst other things) linguistic means, "gender has a variety of effects on the variable and (...) gender in variation cannot be reduced to notions of male and female speech as 'more or less conservative'" (Eckert 1989). While this chapter is not able to resolve the conflict that surrounds this topic, it will be interesting to see if



and to what extent male and female speakers in the two data-sets adopt and further the non-standard variants. Furthermore, following calls to investigate specifically the interaction between gender and other social categories (Eckert 1989, Eckert and McConnell-Ginet 2002, Romaine 2003), this chapter will investigate the social reality of the quotative system at the interface of gender, class and age.

In his 2001 book, Labov, extrapolating from the social patterning of numerous sound changes, has formulated a number of principles, which are intended to predict courses of language change. I will print principles 1, 3 and 4 in full.

Principle 1, or the Curvilinear Principle: Linguistic Change from below originates in a central social group, located in the interior of the socioeconomic hierarchy (2001: 188) ...

Principle 3: In linguistic change from above, women adopt prestige forms at a higher rate than men (2001: 274)....

Principle 4: In linguistic change from below, women use higher frequencies of innovative forms than men do (2001: 292)

In the course of this paper, the developments within the quotative system in British and US English will be examined with an eye on their conformity or non-conformity to the predictions. But we need to bear in mind that with the introduction of *like* into the hitherto stable pool of quotatives, we are looking at the very special case of a rapid change phenomenon. The variable consists of a whole pool of competitor variants, some of which are vernacular (namely *like*, *go* and unframed quotes).<sup>1</sup> And while the numerous lay comments in the media (e.g. List of All Time Banished Words) give evidence that the lexical items newcomer *like* and stigmatized *go* are above the level of consciousness, the status of unframed quotes, the variant without surface form, is not so easy to determine. Also, the question of prestige is problematic with a variant that holds prestige for the younger generation (viz. is highly endorsed by and associated with role models in the media) but seemingly stigmatised by their elders. Furthermore, given the fact that the innovative variant *like* potentially got transposed from one variety to another, it will be interesting to see whether the same principles of transmission and adoption can be shown to hold

---

<sup>1</sup> I employ vernacular in the sense of Holmes (1992:146) in that I use “the term vernacular as an alternative to non-standard... vernacular forms tend to be learned at home and used in informal contexts.”

in geographically separated locations with largely different social structures (Aukrust and Snow 1998).

If we want to investigate the overall idea of women as the innovators or as the more conservative speakers in more detail, Foulkes and Docherty's (1999:16) suggestion that "in the current climate, where non-standard varieties are becoming more and more influential, it follows that females are more likely to be the harbingers of oncoming variants, even if they are non-standard in origin" is a very useful point of departure. Mathisen (1999), Trudgill (1988), Przedlacka (2001), Watt and Milroy (1999) and Wells (1982) independently show that women's speech is more prone to being influenced by non-local forms. Foulkes and Docherty (1999) rephrase the dichotomy of variants being standard versus non-standard used in earlier work as local versus non-local (going back to Milroy et al. 1994 and Abdel-Jawad 1981). They go on to show that female speakers use more supra-local forms, viz. variants which do not characterize the regional variety of a local community. These findings are of great interest for the situation we are investigating. It will become clear during the discussion in this chapter that if we redefine our parameters and speak of supra-local versus local forms instead of prestige versus non-prestige forms, we can account for several patterns we find in the data.

Another important concept involved in this discussion is Britain's "contact, focusing and reallocation" hypothesis (1997), which states that dialect contact can lead to levelling and simplification. One common form of simplification is known as reallocation (Trudgill 1985:40, 1986:152) or rule governed contact (Taeldeman 1989) where "two or more ingredient variants are reallocated, in the new koineized dialect - whether to different phonological environments (Trudgill 1986:159), to different lexical environments (Taeldeman 1989), or to certain stylistic contexts (Trudgill 1986:110)" (Britain 1997:35). While Britain and Trudgill state that such processes can occur "when face to face interaction occurs at a high level between speakers of different but mutually intelligible language varieties" (Britain 1997:34), later research has shown that they do not necessarily need face-to-face interaction in order to take place (Kerswill and Williams 1999, Williams and Kerswill 1999). As I will illustrate below, the reallocation hypothesis can be used to account for the possible

scenario whereby *like* was borrowed from the US into British English.<sup>2</sup> Consider section 6.5. for more details.

Sociolinguistic and discourse-oriented publications which explored the recent changes within the quotative system have primarily concentrated on *like*'s gender affiliation and have resulted in highly conflicting results. For US English, where most investigations of *like* have been conducted, it has been reported to be a feature of women's speech by Ferrara and Bell (1995 for the 1990 corpus), Romaine and Lange (1991:228), Singler (2001) and many others. Blyth et al. (1990), Dailey-O'Cain (2000), Dougherty and Strassel (1998) and Lange (1986) report that in the US, independently of the actual linguistic reality, people tend to perceive *like* more as a feature of female speech. But Blyth et al. (1990), Dailey-O'Cain (2000) and Sanchez and Charity (1999) found *like* to be more prevalent in male speech. Some articles report no gender difference at all (cf. Ferrara and Bell 1995 for the 1992/4 corpus, Tagliamonte and Hudson 1999 for Canadian English, Eckert 2000<sup>3</sup> and Singler 2001). For a very useful overview of the literature see Cukor-Avila (2002).

In a real-time study spanning the years between 1990 and 1994, Ferrara and Bell (1995) show that the female bias has decreased to almost zero as fast as from 1990 to 1992 in Texas. This result is paralleled by Tagliamonte and Hudson's (1999) data from Canadian English in 1995, where they report that the difference in VARBRUL weights .54 (female) and .44 (male) is not significant.<sup>4</sup>

Note on the contrary the situation in British English: while *like* in my corpus (in 1994/5) is not stratified by gender at its incipient stage, later studies in York have

---

<sup>2</sup> I do not want to rule out the possibility of an independent development in US and British English. But *like*'s diachronic precedence in US English (attested as early as 1982) points to the fact that its occurrence in America has at least helped spark the development in British English. Consider Chapter 4 for an analysis of the semantic field underlying the grammaticalisation of *like*, which makes its development into a quotative likely. Consider also the cross-linguistic evidence given in Table 4.1 which justifies calling this process "natural", thus prone to occurring in several places at the same time, cf. Meyerhoff and Niedzielski 2003).

<sup>3</sup> Note that Eckert does not separate out quotative and discourse marker use.

<sup>4</sup> Tagliamonte and Hudson (1999) work with the VARBRUL system, widely used in sociolinguistic research, which gives weighted probabilities stemming from a multivariate analysis. The factorial weights give the probability of a variant occurring in a particular context. The closer these numbers are to 1, the more highly this context favours the occurrence of the dependent variable.

shown that the more diffused *like* gets, the bigger the frequency-based gender difference becomes (cf. Johnson 2001). Tagliamonte and Hudson (1999) speculate that the longer *like* persists as a full member of the quotative cohort, the more gender-stratified it becomes.

Prior to studies that endeavoured to investigate the quotative system as a whole, not much was known about the social or linguistic constraints on quotative *go*. Consequently, besides a short birth announcement by Butters in 1980, I have not come across any study which specifically investigates *go*. It is usually mentioned in passing when discussing quotative *like*. While broadly associated with male speech in North America, *go* is found to be sharply delineating male and female speech, preferred by the males, by Tagliamonte and Hudson (1995) for Canadian English (.35 and .73 VARBRUL weights) or at least to show significance as in Singler (2001, with VARBRUL weights of .55 male and .47 female). When it was first attested by Butters (1980), he associated it with male adolescents in North Carolina. But note that, contrary to the stereotype, Ferrara and Bell (1995) only find a 2% difference between younger speakers of both genders in 1992. Furthermore, Blyth et al. (1990) and Tagliamonte and Hudson's 1996 British English corpus show that *go* is non-significant with respect to gender.

In British English very little research has been done on the quotative system at all. All sociolinguistic accounts of the quotative system have been conducted after the introduction of newcomer *like* and specifically target its distribution (and later also its interaction with the other members of the quotative cohort). The only three variationist studies of the quotative system of British English which investigate *like* and *go* are Macaulay (2001) for Scottish English collected in 1997, Tagliamonte and Hudson (1999) with data collected in York in 1996 (from a cross-variety perspective) and a follow-up study by Johnson (2001) in 2001 in York. One major asset of the Tagliamonte and Hudson (1999) study is the fact that they draw attention to the fact that *like* is on the borderline between thought and speech. Consequently, *think* and, by extension, other verbs of introducing quoted thought are included in their analysis, which closes the set of the studied variable and makes their results more comparable to mine.

Interestingly, in my data, there is a big gap in *go*-production between only 5.8% in the US English and 12.8% in the British English data. The question is whether this fact stands in relation to an issue which has been raised in the literature (Bakht-Rofheart 2002, Blyth et al.1991), viz. whether the spread of *like* intrudes in *go*'s territory. Note in this respect Macaulay's (2001:6) proposition that "*go* became popular in the US among younger speakers about 25-30 years ago but soon was replaced by *be like*, although older speakers continued to use *go*." I will aim to test this claim. If *like* is indeed replacing *go*, this should show up in the apparent time data.

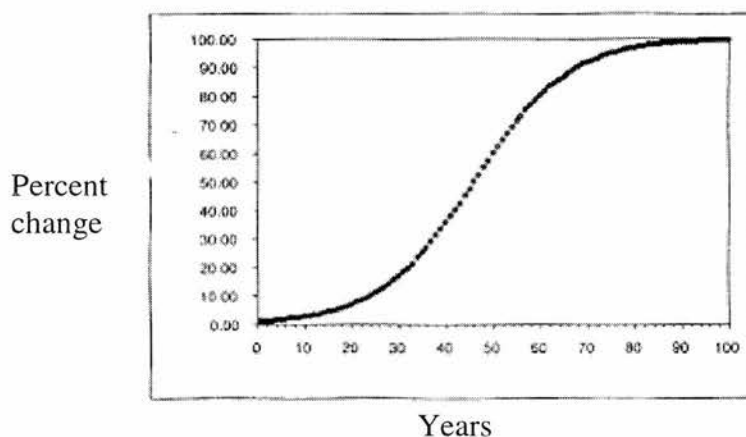
The following sections discuss the sociolinguistic patterning of the two corpora under investigation. While the amount of variation we find in the data sets explains only a small part of the overall variation observed, a study of the social constraints which govern the quotative system is nevertheless valuable for the following reasons: By comparing the patterning of the (say) variable as a whole in Britain and the US, we gain insights into the rule-governed variability of innovative features and their competing variants in two geographically distant varieties. Furthermore, the restructuring of all quotative variants during the rapid intrusion of the robust new variant *like* gives us an important test-case for various explanatory parameters for language change that have been raised in the literature. In the discussion to follow, I will consider my data with respect to Labov's principles, the reallocation hypothesis, Foulkes and Docherty's claims and the Milroys' findings. Looking at the data we will see that the social constraints of quotative *like* and *go* are not the same in the two varieties. Hence, even if *like* was transmitted from the US to Britain, the social information adhering to this item in the donor variety was not taken over but has been re-created in the borrowing variety. This finding constitutes an important contribution to the growing body of research, within sociolinguistics, on globalisation phenomena (Buchstaller 2003, Kerswill 2003, Meyerhoff and Niedzielski 2002, 2003, Tagliamonte and Hudson 1999, Trudgill 1983, 1985 and many others). Also, a comparative study of the apparent time patterning of non-canonical variants within the quotative pool reveals important insights into

phenomena such as reallocation, competition within one socio-pragmatic field, as well as the interaction of variants of a single variable.

## 6.2. A note on statistics

Before beginning variationist sociolinguistic research into the two new quotatives, a few preliminary words of caution concerning the statistical analysis are in order: Labov shows that changes in their incipient stages move very slowly and have low token numbers which do not get picked up by statistical analysis (because they do not achieve significance). With respect to the Philadelphia Sound Shifts, he finds that the “relatively low involvement of social factors is characteristic of the earlier stages of change, when the major independent variables concern group membership rather than social status” Labov (2001:301). Note again Figure 5.2 (drawn from Labov 2001:450) here replicated as Figure 6.1, which depicts the calculations of language change as a logistic curve underlying the increments from one year to another. The vertical axis tracks such changes from 0 to 100 over the horizontal time-axis.

Figure 6.1: 100-year progress of a change with incrementation (from Labov 2001:450)



The low values on the left hand side show the slow start of linguistic change with increments picking up force until the midway point and slowly decelerating symmetrically after that. This is concomitant with earlier observations that changes begin slowly, pick up speed when they are in mid-course and slow down in their later

stages (Bailey 1973, Kroch 1989, Weinreich, Labov and Herzog 1968). Judging from this figure and bearing in mind the fact that *like* is only a few years old in British English at the time the data was collected (cf. Chapter 2), we would necessarily expect low token-numbers in the British corpus. An incipient linguistic variant, *like* is expected to be situated on the extreme left of the continuum. In American English, on the other hand, we would expect *like* to have picked up speed. Incrementation tends to accelerate when the variant starts carrying social or stylistic weight; *like* may be socially, and stylistically, more stratified.

In general, as Cameron (1998:57) points out, discourse variables – such as quotation – are relatively rare in speech<sup>5</sup> and the collection of a sufficient quantity of tokens requires an enormous amount of textual data. In my data, the Derby and Newcastle speakers combined produced 2064 quotations and the US speakers 1371. Note again the break-down of the (say)variable:

**Table 6.1: Frequencies of occurrence of variants of the (say) variable**

	British English		US English	
	N	%	N	%
like	93	4.5	121	8.8
go	264	12.8	80	5.8
say	967	46.9	562	41.0
think	175	8.5	121	8.8
unframed	462	22.8	298	21.7
tell	23	1.1	42	3.1
other	80	3.9	147	10.7
<b>Total</b>	<b>2064</b>		<b>1371</b>	

As this investigation is concerned with all quotative options within the (say) variable, a further split of this multiplicity of variants per numerous social categories, viz. independent variables, reduces cell size significantly. This turns out to be especially problematic with respect to the variant *like* in British English and with respect to quotative *go* in the US data. The low token numbers mean only few subdivisions of the social constraints are possible. In order to keep cell size reasonable, I have

<sup>5</sup> The occurrence of quotation has been shown to be very strongly genre-related, with narratives providing the most common environment of occurrence. As not all speech consists of narratives, a bias towards naturally occurring talk-in-interaction has proven to yield highest frequency of quotation (cf. Chapter 2 for an overview about the data used in this investigation).

restricted this investigation to three categories: age, gender, and socio-economic status.<sup>6</sup>

As I noted in the data section in Chapter 2, even with a carefully designed corpus that aims at selecting a speaker pool which represents all social categories equally in terms of numbers of speakers, it is a well-established fact that different social categories of speakers produce uneven amounts of speech (cf. Macaulay 2001, Tannen 1986). I will now lay out in some detail the important consequences of this numerical bias on the statistical evaluation of the raw data.

Consider the production of quotatives in my two corpora split by gender, age and class as depicted in the following tables:

**Table 6.2: Production of quotatives per gender**

	N	%
male	1500	43
female	1930	56

**Table 6.3: Production of quotatives per class<sup>7</sup>**

	N	%
working	1890	55
middle	1540	45

**Table 6.4: Production of quotatives per age**

	N	%
young	2107	61
old	1324	39

As already pointed out by Ferrara and Bell (1995), Johnstone (1988, 1990, 1993), R. Lakoff (1973), Romaine and Lange (1991) and many others, the production of quotation in any corpus is skewed towards the female speakers. Table 6.2 shows that this applies also to my data, where women produce 56% of all quotes.<sup>8</sup> A further

<sup>6</sup> The concept of socio-economic status was predetermined by the respective research projects in which the data was collected. Consider Chapter 2 for details of the methodology involved in sampling and coding.

<sup>7</sup> Note that the parameters for the inclusion of individual speakers in socio-economic categories were slightly different in the US and in Britain. They have been grouped together here for demonstrative purposes only. Please refer to the data section (Chapter 2) for details.

<sup>8</sup> There is evidence that there are different speech styles according to gender (Macaulay 2001, Vincent and Dubois 1999). Therefore, there is a possibility that the differential number of quotations produced is a direct function of the speaker's gender. But the question of to what extent unequal overall quotative production is due to an underlying social factor (such as speech style) or whether it simply reflects unequal speaking times (with one or the other gender claiming more speaking time within the mixed sex-pairings) will not be investigated in this thesis.



skew in the data stems from a much less noteworthy numerical bias towards the working class. In spite of being given the same amount of speaking time in the corpus, working class speakers quote more (55%) than middle class speakers (45%). This effect is consistent with Rimmer's (1988) findings in Liverpool and Birmingham, where working class speakers with less specialized training used the most quotations (cf. Macaulay 2001, Vincent and Dubois 1996: 370). Again, we have a numerical skew leaning towards one social category.

Finally, Table 6.4 shows that the overall quotation production is considerably biased towards younger speakers, who are responsible for 61% of all quotations in the corpus, whereas older speakers produce fewer quotations (39%). Again, this is consistent with previous results (Dailey-O'Cain 2000, Ferrara and Bell 1995, Macaulay 2001, Romaine and Lange 1991).

To sum up, given potentially the same speaking time, members of different social categories will produce highly unequal numbers of quotes. As Tables 6.2-4 show, my data is significantly skewed, far from normalised. But as this investigation is not concerned with the raw token numbers of quotative production but rather aims at describing the composition of the quotative *system* as a function of social categories, the uneven numerical distribution of overall quotative use does not constitute an impediment to statistical investigation per se. Rather, it is a noteworthy reality and indeed a determining factor for the way statistical analysis is to be carried out. I will demonstrate this point in the discussion to follow.

The research question underlying this chapter is: what is the composition of the quotative system per social category. How often do, say, female speakers produce *like* or *go* out of all quotations they produce? This can be stated even more specifically: what is the proportional use of the new quotative *like* within its functional domain, the pool of quotatives? How does it interact with *go*, *say* and unframed quotes in social space? In the following paragraphs, I will show that an investigation starting with such a research question also avoids a fundamental error commonly found in calculating such data.

If we are interested in the distribution of quotations per social category, the frequency-based distribution can be calculated in two ways: firstly, we can allocate each quote-introductory strategy to one of the levels of the social variable involved. Table 6.5 represents such a vertical split of all quotatives in my corpora by social category, here gender. At first glance, such a vertical subcategorisation seems to represent token-distribution per social category.

**Table 6.5: Vertical distribution of quotatives per gender (m. = male, f. = female)**

	like		go		say		think		unframed		tell		other	
	%	N	%	N	%	N	%	N	%	N	%	N	%	N
m.	41	89	35	119	44	665	42	123	49	369	51	33	45	101
f.	59	125	65	225	56	863	58	173	51	391	49	32	55	125

Table 6.5 answers the question of which percentage of a given quotative is produced by male and which percentage is produced by female speakers. For example, out of all *like* tokens in the whole corpus, 41% are produced by male speakers and 59% by female speakers. This outcome is very likely to be highly significant. The question remains: 59% out of what (Lass 2000)? This is an important issue because such an investigation only makes sense when we have an even number of tokens overall. But as Tables 6.2 has shown, we do not have even distribution of *like* across male and female speakers. An extreme example will make this point clear: Let us assume that men produced 100 quotatives and women 100,000. If both social groups were responsible for 5 *like*-tokens, a vertical statistical representation such as the above would end up showing a 50% *like*-frequency for each social category, men and women. Such an analysis makes it look as if men and women had an equal *like*-output, 50% each. Yet, the import of *like* within the quotative system of men and women is completely unlike: 5% of the hypothetical men's sample and only .005% of the hypothetical women's quotatives are tokens of *like*. If we calculate our frequencies vertically, the percentages for a quotative variant are effectively independent of the overall number of quotatives produced by the respective social category. The hypothetical example has shown that a numerical breakdown in this way can greatly falsify the statistical results. Consequently, it does not give accurate information about the build-up of the quotative system according to social categories such as gender, age etc. It follows that the skewedness of my data, and of probably

all non-normalized naturally occurring data, forbids a vertical statistical evaluation which does not take into account the overall sum of quotatives produced.<sup>9</sup>

It becomes clear that we need to have a stable denominator. We will have to measure the token-number of quotatives produced (*like, go, say, unframed, etc.*) against the variable N (number of all quotations produced) and we will have to do this for all social categories. In order to get at the proportional estimate of a frequency of quotative *x* by a social category *A*, we can use the formula

$$\frac{x(A)}{\sum (\text{quotatives})_A}$$

Accordingly, for the rest of this chapter, I will display frequency as rate of occurrence calculated with the above formula, in what amounts to a horizontal comparison. Henceforth, the frequency distribution (in %) of a type of quotative is understood to be the occurrence of this type of quotative per *all* quotative tokens produced.<sup>10</sup> This will become obvious in the next table:

**Table 6.6: Horizontal distribution of all quotatives per gender (in % and N)**

	like		go		say		think		unframed		tell		other		SUM
	%	N	%	N	%	N	%	N	%	N	%	N	%	N	
m.	6	89	8	119	44	665	8	123	25	369	2	33	7	101	1499
f.	7	125	12	225	45	863	9	173	20	391	2	32	7	125	1934

Table 6.6 restructures Table 6.5 and gives a breakdown of quotative use *per total number of quotations*. Every grey cell contains the percentage of occurrence of a particular quotative type used by the members of a particular social category (here male and female) out of *all* quotatives produced by this respective social category. The overall sum of quotations is given in the last column (in bold borders) on the right hand side of the table. Calculating ratios over the denominator, we end up with a split of the (say) variable in its composite variants/strategies for every demographic

<sup>9</sup> It needs to be stated at this point that I have refrained from setting up an index of N(tokens) / y(words) or z(seconds) for the individual quotatives. This is because such an index would not give any insights into the importance of social constraints on the statistical composition of the quotative *system* either. All it would indicate is, again, the *overall* number of quotatives produced by the individual social categories per so many words or seconds.

<sup>10</sup> Consider Dailey-O’Cain (2000:65) for what she calls “a ratio for each speaker which tells the numbers of times (...) quotative *like* occur[s] as compared with the number of times it was possible for [them] it to occur”. The formula provided treats the overall sum of quotatives as the possible times it could occur and hence puts this as the denominator.

category. For example, in the two corpora, *like* is used by male speakers to a frequency of 6% out of all quotatives produced by the male speakers. Put differently, in the quotative system of male speakers, 6% of tokens are *like* while *like* comprises 7% of the females' quotative system. The apparent 41% versus 59% difference between the sexes in Table 6.5. evaporates to a non-significant result ( $\chi^2(1): 2.563, p = .109$ ).<sup>11</sup>

Calculated this way, the numerical imbalance between the number of quotations as produced by male and female speakers does not skew the statistical evaluation. A horizontal break-down displays the distribution of variants accurately because frequencies of quotatives are not independent of the overall quotative production. On the contrary, because the numerical unequal outputs by female and male speakers constitute the denominator of the equation, the quotative frequency is a function of the variable overall production by the particular social category under investigation. It follows that such a horizontal subcategorisation is largely preferable to the vertical comparison. It evaluates "its proportional use .... within a given common domain (i.e. all quotatives) in the data set under investigation" (Tagliamonte and Hudson 1999:156).

The fact that only a horizontal evaluation makes sense in cases of naturally occurring data with non-normalized subsets and also because my data is not binary but consists of at least 5 dependent variables (cf. Labov 2001:96) precludes the use of a logistical package which has been widely used in sociolinguistic studies, VARBRUL (Guy 1988, Guy 1993, Pintzuk 1986, Young and Bailey 1996). VARBRUL comes in two versions, one for Macintosh, which works for 2 dependent variables and one for DOS environments, which can cope with 3 or more dependent variables. A higher number, if theoretically possible, is not practicable as the weighted values become small and therefore undifferentiable (Young and Bailey 1996). Hence, I reverted to the statistical package SPSS in order to perform the necessary cross-tabulations, frequency estimations and significance testing and all results reported in the rest of this chapter are based on this.

---

<sup>11</sup> The statistical analysis was run with a recoded set of variables, *like* versus else (variables are *like*, coded as a dummy variable 1, and all other quotatives, coded as variable 2). The chi square statistic was then run as a cross-tabulation with *like\_other* as the dependent and gender as the predictor variable (cf. Labov 2001).

### 6.3. A regression analysis

Let us now have a look at some comprehensive statistical results. In the following discussion, I will first look at US English and then at British English. In order to examine *like* and *go*'s social constraints, a regression analysis was run. Using social factors as predictors, such an analysis can account for the various strengths of the social factors involved. In so doing, we can find out which external constraints play a role in the patterning of a dependent variable and what the individual predictors' strength is. For the following tables, I have recoded all other quotative variants as a dummy variable. Its patterning was then compared against the variant we are interested in, in this case *like* or *go* (cf. Labov 2001). By calculating the statistical output for each independent variable while keeping the other predictors constant, we can measure the relative weight for every social predictor variable while still accounting for all others simultaneously. The following tables give the probabilities (p) for the respective social factors in the middle column. The  $r^2$  ratios in the right hand column indicate how much of the dependents variables' inherent variability can be accounted for by each independent variable (again, while keeping the others constant). Thus, for example, 2.8 % of *like*'s patterning can be accounted for by the factor age.

Furthermore, Guy (1993) and Young and Bailey (1996:273, cf. also R. Cameron 1998) have shown that log-likelihood can be used as an indicator of the goodness-of-fit of a model to the data. In this analysis, the category which shows a log-likelihood closest to zero represents the best account of the variability of the variant. In our specific case, we can do a run with all social categories included in order to test which social constraints are selected as most important with respect to the new quotatives' extralinguistic patterning. The combination of predictor variables which accounts best for the variability (the model with the best "fit") achieves the lowest log likelihood.

With respect to *like* and *go*'s patterning in the US, Table 6.7 plots each individual social factor's log-likelihood, p-value and  $r^2$  incrementation. The last row gives the model which has been chosen by the program as the best goodness of fit and the social factors involved in this particular model (with all factors in the run).

**Table 6.7: Logistic regression for *like* and *go* with the social categories age, gender, and class in US English (significant results in bold)**

	like			go		
	log likelihood	p <sup>12</sup>	r <sup>2</sup>	log likelihood	p	r <sup>2</sup>
age	779.360	<b>0.000</b>	0.028	601.166	<b>0.005</b>	0.005
class	809.714	<b>0.023</b>	0.006	609.044	0.2	0.000
gender	815.674	0.804	0.002	609.019	0.37	0.001
best goodness of fit						
age+class+gender	773.070	0.007	0.032	599.243	0.028	0.008

The log likelihood values indicate that the factor age alone is the best predictor for quotative *like* in the US. It also achieves the highest significance ( $p < .001$ ) and its  $r^2$  ratio shows us that the factor age alone accounts for 2.8% of *like*'s variance (cf. Labov 2001). While class is significant at the  $p < .05$  level, it can only account for a negligible 0.6% of *like*'s variance. The predictor gender is not significant and would only add 0.2% to the model of fit.

But while gender and class alone are not very good predictors, in combination with age, they nevertheless contribute to the explanatory model and add to its goodness of fit. The last line of the table shows that a model with all three variables was chosen for the best goodness of fit (lowest log-likelihood with 773.07).

The statistics for *go* are similar to those for *like* inasmuch as age is the most highly significant predictor ( $p = .005$ ). But in contrast to *like*, age only accounts for 0.5% of the variance for *go*. Class is not significant (with  $p = 0.2$ ) and accounts for only 0.1% of the variance. The variable gender is not at all significant at  $p = 0.37$ .

Note that while the factor age alone obviously achieves the log likelihood closest to zero (601.166), a combination with the other two predictors achieves the best goodness of fit. Again, a model which simultaneously takes into account gender, class, and age is the best model to account for *go*'s variance with a log likelihood of 599.243 ( $r^2$  ratio of .008).

Hence, in the US, the only factor that comes out as highly significant for *like* and *go*'s social patterning is age (with  $p = .000$  and  $.005$ ). *Like* also displays a slight class effect (with  $p = .023$ ), while *go* is completely unaffected by class. Note that,

<sup>12</sup> There and in the next probability testings, all three variables, age, class and gender, are entered for likelihood evaluation.

surprisingly, given the large focus of the sociolinguistic literature on the variable gender, this social factor does not play any role whatsoever in the patterning of the new quotatives in the US data. Overall, the social constraints on the new quotatives are far less important than we might have supposed.

Let us now look at the social situation across the Atlantic. Table 6.8 displays the importance of the factors age, gender and class with respect to the patterning of quotative *go* and *like* in British English.

**Table 6.8: Logistic regression for *like* and *go* with the social categories age, gender, and class in British English (significant results in bold)**

	like			go		
	log likelihood	p	r <sup>2</sup>	log likelihood	p	r <sup>2</sup>
age	697.121	<b>0.000</b>	0.020	1422.948	<b>0.000</b>	.073
class	758.173	0.851	0.000	1560.788	<b>0.002</b>	.009
gender	757.884	0.262	0.000	1566.288	<b>0.015</b>	.006
best goodness of fit						
age+class+gender	695.654	0.264	0.03	1406.02	0.732	0.08

Again, age is the single best predictor for quotative *like*. Its r<sup>2</sup> ratio indicates that it alone can account for 2% of *like*'s variance. When we look at the individual variables' goodness of fit as measured via the log likelihood in isolation, it becomes quite clear that the predictor age with a log likelihood of 697.121 shows the best fit. The combination of all three independent variables age, gender and class constitutes an even better model with a log likelihood of 695.654 and a slightly higher r<sup>2</sup>. Class and gender alone are not significant (p=0.851, 0.262).

Turning to *go*, we notice that a combinatory model which takes into account all three social variables again shows the best goodness of fit overall with a log likelihood of 1406.02 and an r<sup>2</sup> ratio of 8%. The individual independent variables also all achieve significance. Again, gender has the least significance (with p-values of .000, .002, and .015) and age is the single most significant factor involved. Note that with 7.3%, its r<sup>2</sup> is significantly higher than for *like*.

Thus, in the British system, quotative *like* only patterns with respect to age, whereas quotative *go* is subject to the social constraints of age as well as class and gender.

Overall, we can safely say that not much is going in terms of the new quotatives' patterning with respect to social factors. Contrary to the hype in the literature, we see that *like* does not display a strong social patterning with respect to social variables at all. The only factor that achieves significance in either variety is age (and a mild class effect for US English).

*Go*, while showing an effect with respect to all three variables, age, gender and class in British English, only patterns according to age in US English.

While the overall social patterning of *like* looks far less than exciting (and *go* is only showing strong social constraints in British English), I now will investigate in more detail what lies behind these general effects. A cross-correlation between social factors and comparison with other existing data-sets in real and apparent time will give some depth to the investigation.

## **6.4. US English**

This section will investigate the patterning of all quotative variants in the US, the variety where quotative *like* and *go* were first officially attested as new quotatives by Butters in 1980 and 1982. Consequently, *like* had had at least 6 years to spread across the social sphere by the time my corpus of USE was collected. I now will discuss their social patterning in my data from 1988-92 in the light of the earlier literature.

This section is subdivided into three parts. First, I will first discuss the (say) variable with respect to the social factor gender, then I will investigate the impact of the factor class. Finally, I will look more in-depth at the age effects across and within time.

### **6.4.1 Gender**

The factor gender has been a focus of previous research on quotatives. Studies either allocate *like* to the male or to the female speaker and only very few show no significant gender effect with respect to *like*. As can be seen from Table 6.9, which



depicts the patterning of all quotative variants in the US according to gender, *like* is slightly more common for female speakers but not significantly so.

**Table 6.9: Distribution of quotatives in US English per gender, significant differences bold**

	like		go		say		think		unframed		tell		other		SUM
	%	N	%	N	%	N	%	N	%	N	%	N	%	N	
m.	7	47	6	33	<b>45</b>	282	8	48	21	131	4	22	10	64	627
f.	10	74	6	47	<b>38</b>	280	10	73	22	167	3	20	11	82	743

In the sociolinguistic literature a female lead has sometimes been used as a diagnostic of language change in progress (cf. the literature cited above). But a chi square analysis shows that the 3% differential between men and women is not significant ( $\chi^2(1): 2.563, p = .109$ , cf. the regression analysis above). Consider in this vein Ferrara and Bell's (1995:285) results from a real time study where the female bias they found in the earlier data was diminishing in the later corpora as frequency was increasing and as *like* was generalising its intralinguistic functions.

According to their findings, while women are still in the forefront in 1990 in Texas, by 1992 this lead has levelled out. If this reflects a general trend for *like*, greater diffusion in any groups of speakers might be expected to go hand in hand with the levelling of the gender-effect. My data for the period 1988-1992 would slightly predate Ferrara and Bell's corpus but can be consolidated with their findings when we remember that *like* (first attested in the West and North East) arrived down south with a substantial time-lag and that the gender levelling would have occurred there later than in other parts of the country. But it is important to note that work by Cukor-Avila (2003) in Texas shows that *like* had arrived in the quotative system of Black as well as White speakers as early as 1988.

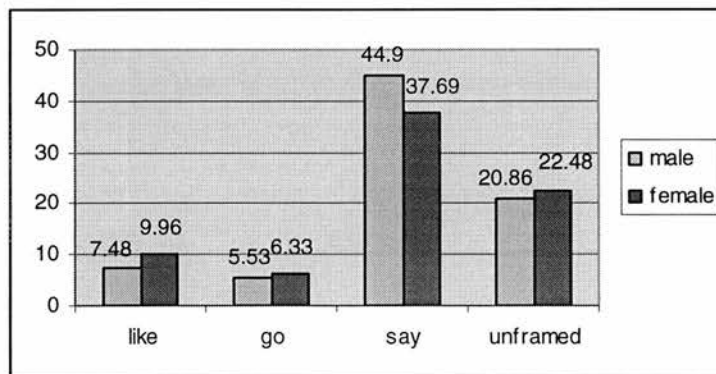
*Go*, far from sharply delineating male and female speech or at least showing significance, displays no gender difference (6% female and male). This result parallels findings by Ferrara and Bell (1995) and Blyth et al. (1991). Taken together, these results run counter to the prevailing and much cited stereotype that *go* is a "rough", male, blue collar feature (Blyth et al. 1990). Reality defies this stereotype, which has been around at least since *go* was first associated with male adolescents in North Carolina (Butters 1980). The results presented here on US English and

Tagliamonte and Hudson's data on Canadian and British English agree in the fact that *go* is not stratified by gender.

Unframed quotes pattern the same as *like* and *go*. The difference between female and male speakers (22% females 21% males) is at the level of chance, ( $p = .511$ ).

As Figure 6.2 shows, female speakers in the Switchboard corpus consistently favour vernacular ways of enquoting (*like*, *go* and unframed quotes). The preference of female speakers for non-canonical variants is directly reflected in their differential use of the canonical option *say* (7% gender difference 45% male versus 38% female). This production gap constitutes the biggest gender effect in the whole (*say*) variable and comes out statistically significant ( $\chi^2(1): 7.471, p < .05$ ). Figure 6.2 gives the break-down of canonical versus non-canonical forms in the US data:

**Figure 6.2: Distribution of canonical and non-canonical variants per gender**



While no significant gender effect holds for any one of the vernacular variants, the combined tendency canonical versus non-canonical just about reaches significance with  $p = .05$ . Even though the significance is not very strong, we can nevertheless confirm a tendency for women to favour non-canonical ways of quoting in general. Men are more prone to sticking to canonical *say*.

The finding that women use more non-traditional means to enquote past speech and thought events might be due to what Labov (1990:214) mentioned as a tendency of female speakers towards greater expressiveness (cf. also J. Milroy 1981:37). This higher degree of linguistic creativity manifests itself in the overall higher use of all

three non-traditional variants (*like*, *go* and unframed).<sup>13</sup> Note also that this balance is slightly different in British English as discussed in 6.5.1.

#### 6.4.2 Socio-economic Class <sup>14</sup>

The following paragraphs deal with the distribution of the members of the quotative system according to socio-economic class. While the literature gives comparatively little statistical evidence with respect to the patterning of quotatives in US English according to class, I will use impressionistic comments in the literature as well as from lay-people as a comparative foil whenever appropriate. Table 6.10 shows the distribution of all quotatives with respect to less (less) and more educated (edu.) speakers in my US corpus.

**Table 6.10: Distribution of quotatives in US English per educational level, significant differences in bold**

	like		go		say		think		unframed		tell		other		SUM
	%	N	%	N	%	N	%	N	%	N	%	N	%	N	
less	<b>11</b>	81	5	40	<b>38</b>	284	9	65	23	169	3	23	11	84	746
edu.	<b>6</b>	40	6	40	<b>45</b>	277	9	56	21	129	3	19	10	62	623

If, as claimed in Holmes (1992:174) and elsewhere (see Trudgill 1983) “vernacular forms express machismo”, we would expect *go*, stereotypically associated with working class speech (according to Blyth et al. (1990), used by “men like Rocky”), to carry the covert prestige of roughness and accordingly to be preferred by men. But the present investigation shows that *go* is neither a feature of male speech (Table 6.9), nor of less educated speakers. Rather, as we see in Table 6.10, there is no difference whatsoever in the distribution of quotative *go* by socio-economic status (with 5% and

<sup>13</sup> Cf. similar results on the use of quotatives by preadolescents by Levey (2003).

<sup>14</sup> The term “class”, amply used in sociolinguistic studies, is slightly awkward in general, and for an investigation in which the socio-economic categorization is based on education especially. In order to avoid the awkward long-windedness of the term ‘socio-economic standing as based on educational level’, I resorted to the admittedly less-than-perfect solution of speaking of socio-economic class (henceforth SEC) when investigating US English. As the British English data was collected from speakers belonging to two classes, I will refer to this term in the discussion of the Derby and Newcastle data. For information on how individual speaker were grouped with respect to class in the BrE data and socio-economic level based on education in the US data please consider Chapter 2.

6%,  $\chi^2(2): .754, p=.686$ ). This result belies the claims concerning its working class status.<sup>15</sup>

*Like*, on the other hand, does pattern according to socio-economic standing in US English. Contrary to the much-cited stereotype that it is used mainly by university students (Blyth et al. 1990), in my corpus, *like* is preferred by speakers with a lower educational level (with 11% versus 6%,  $\chi^2(1): 8.348, p<.05$ ). Note that Dailey-O'Cain's (2000) study and my study independently show that *like* is primarily a feature of the less educated speaker's habitus and, importantly, it is perceived as that in the US as well. The association with female university students might well stem from a persistent sampling bias: Most studies of the American English quotative system to date have been conducted only or mainly on university students and hence base their conclusions on a speaker selection which is not representative of the population as a whole.

To forestall possible objections that educational level is not independent of age and that, therefore, the preponderance of *like* among speakers of a lower socio-economic level might be a function of their being older, interaction effects between age and education were investigated. A factorial ANOVA between subjects rules out interaction as an explanatory parameter ( $F(1,1369) = 1.373, p= .242$ ). Thus, we can discount the possibility that *like*'s significant patterning with respect to educational level is due to the interdependence between the factors age and education. Rather, age and educational level are significant but independent of one another. The question of quotatives patterning according to age will be discussed in the next section.

*Say*, the canonical quotative, is used more by the higher educated speakers ( $\chi^2(1): 5.900 p< .05$ ). In the 1988-1992 US corpus, the educated speakers have a more conservative habitus generally. Again, a check for interaction effects with age shows no significance ( $F(1,1369) = 2.435, p= .119$ ).

---

<sup>15</sup> It has been suggested that *like* could have bled *go* for class. But an apparent time investigation reveals that *go*'s class affiliation stays stable across age. While it is true that *go* is used more by highly educated speakers in the youngest age group, the fact that all other age groups are unaffected by *like*-intrusion does not allow any conclusions regarding the interaction of *like* and *go* with respect to class.

To complete the discussion of the correlation between quotative choice and class, the patterning of the other quotative options proves to be unaffected by class and is non-significant (unframed:  $\chi^2(1): .783$   $p=.376$ , think  $\chi^2(1): .029$   $p=.865$ , tell  $\chi^2(1): .002$   $p=.967$ ).

We still need to investigate whether there are any interesting interactions between gender and class. Focusing on the variants *like*, *go*, *say* and unframed, I will now split up my US data according to those two predictor variables. Table 6.11 gives the break-down:

**Table 6.11: Distribution of *like*, *go*, *say* and unframed across gender and class (in %)**

	like	go	say	unframed
less educated men	7	3	47	21
less educated women	13	7	34	24
educated men	8	7	44	21
educated women	4	6	46	20

Table 6.11 shows that the group who mainly use the newcomer-quotative *like* are the less educated female speakers (13%). Educated women, in contrast, are the group who produces the least *like* with only 4%. Hence, with respect to *like*-use, the biggest gap in production is neither within the lower class (Cameron and Coates 1988) nor between male and female speakers (L. Milroy 1976, Milroy and Milroy 1978, J. Milroy 1980). Rather, the significant differential of *like*-use distinguishes less educated women from more educated women ( $\chi^2(1): 11.106$ ,  $p<.01$ , all other pairings n.s.), cf. also Eckert (2000), Labov (2001). A check for interaction effects (two-factorial ANOVA) revealed that age and educational level do indeed interact for the female speakers ( $F(1,742) = 9.467$ ,  $p=.002$ ) but not for the male speakers and not overall (cf. above).

Hence, to say the least, we can state that in the 1988-1992 US data, *like*-use divides females into more and less educated speakers. We need to modify the above statement that lower educated speakers are the heaviest *like*-users: rather, the less educated female speakers primarily use *like*, notably in contrast to higher educated women. The figures are too small to speak of an avoidance effect (as discussed by Bakht-Rofheart 2002). But there still seems to be some sort of dissociation going on, with educated women being less inclined to embrace *like*-use.

With respect to *go*, the differences in frequency are not very big; all speakers produce about 6% *go*, except for the less educated men, the very socio-economic category to which the literature has attributed *go*. A glance at British English will show that this result is consistent on both sides of the Atlantic (see Figure 6.15). While *go* is apparently perceived to be typical of lower class men, in reality they are precisely the speakers who produce it the least (with only 3%). This interesting finding shows how misleading sociolinguistic intuitions can be especially with respect to discourse, or in this case, quasi-discourse variables (Labov 2001). Furthermore, it underlines findings by Buchstaller (2003) about the non-coherence between perception and reality especially with respect to *go*.

Unframed quotes are again slightly favoured by less educated women (by 24%) whereas the rest of the speakers use them around 20% of the time.

Overall, women with less formal schooling prefer *like*, *go*, and unframed strategies, all of which are non-canonical strategies. Consequently, they have the lowest *say*-production overall with only 34% as compared to over 40% for all other social groups (lower class males 47%, educated females 46%, educated males 44%). While the results on the US English unframed and *say* variants corroborate the claim that the overall gender difference is bigger for the working class speakers (Cameron and Coates 1988), this does not tell the whole story. As I have shown above, it would be misleading to try to explain *like*'s patterning as a function of gender differences within class. Rather, the (say) variable patterns with respect to gender *and* class.

A final analysis by canonical and non-canonical variants shows that this subdivision is actually a good parameter for understanding the data:

**Table 6.12: Distribution canonical versus non-canonical quotatives split up by gender and educational level**

	non-canonical	canonical
less educated men	31	69
less educated women	43	58
educated men	35	66
educated women	31	69

While less educated women use non-canonical variants most, educated women use them least. Men pattern in-between. The cline for the use of non-canonical variants

can thus be represented as follows: Less educated women – educated men – less educated men – educated women, whereby both ends of the cline show significance but none of the intermediate stages does (less-educated men vs. less-educated women:  $\chi^2(1): 8.807, p=.004$ ).<sup>16</sup>

When looking at the use of the canonical variant, *say*, we notice a big gender divide amongst the less educated speakers. Here, women produce the smallest numbers and men prove to be more conservative *say*-users. But again the biggest overall effect is the difference between educated and less educated women in the use of the non-canonical variants.

What about the less educated men, the ones to whom Trudgill (1972, cf. also Labov et al 1968) ascribed the notion of “covert prestige” and who, according to Labov (2001:216), consider non-standard forms as “symbolic of masculinity or toughness; or as expressing a warmer, more personal, more human or more friendly approach to life”? As can be seen from Table 6.12, when it comes to quotatives, less educated men are the most conservative social group with respect to their linguistic habitus. Rather, it is the less educated women who have the most vernacular quotative system and produce the highest number of non-canonical quotatives (with 43%).

### 6.4.3. Age

The first accounts of quotative *like* allocated it to young, female university students from California, who were also the subjects of the first corpus-based studies. From then onwards, numerous papers commented on the upper age-limit of *like*-use in the US. Blyth et al note (1990:219) that *like* was completely absent from the quotative system of speakers older than 38, in 1995, Ferrara and Bell (1995:273) report with a corpus from 1992 that the oldest speaker who produced any tokens of *like* was 39. Singler (2001: 296-70) comes to the conclusion:

“The simplest view of the generational change in progress is that it is both straightforward and inexorable, with the prevalence of *like* growing as the

---

<sup>16</sup> Educated men vs. educated women:  $\chi^2(1): 1.385, p=.255$ , less-educated men vs. educated men:  $\chi^2(1): 1.128, p=.296$ , less-educated women vs. educated women:  $\chi^2(1): 9.459, p=.002$ .

individual users of *like* grow older and the percentage of primary *like*-users in the general population increases thereby. By that view, the sharp drop that separated the 27-33 group in the corpora from the 36-42 groups represents the temporal 'isogloss' of the change. This reasoning has it that in ten years the sharp drop will separate a 37-43 group from a 46-52 group”.

In short, with increasing time, *like* expands sociolinguistically. Its spread across age in real time is indicative of its newcomer status. But note that *like*-use also spreads upwards in age disproportionately to the ageing of the generation who first used it, as numerous studies have commented on *like*-borrowing by people who, according to their age, do not fall into the typical category of *like*-users.

Since non-attestation of a feature certainly is no proof for its absence in a particular, in this case social, category (consider the discussion in Chapter 5), the findings cited above call for a careful check of *like*'s patterning across age in the Switchboard corpus. The following paragraphs investigate the distribution of *like* as compared to the other members of the quotative pool with respect to the social variable age. I will first give a simple break-down of my speakers into older (38+) and younger (13-38) age. To counterbalance this very broad categorization, I will later align the speakers on an age-scale and correlate quotative production with numerical years of age.

**Table 6.13: Distribution of quotatives in US English per age-group, significant differences in bold**

	like		go		say		think		unframed		tell		other		SUM
	%	N	%	N	%	N	%	N	%	N	%	N	%	N	
young	<b>13</b>	100	<b>7</b>	58	<b>36</b>	285	7	55	23	181	3	23	10	81	783
old	<b>4</b>	21	<b>4</b>	22	<b>47</b>	277	11	66	20	117	3	19	11	65	587

A cross-correlation between age-group and *like*-ratio shows that younger speakers produce 100 tokens of *like*, which amounts to 13% in their system, whereas older speakers only use this new quotative with a frequency of 4% (N=21). This distribution across age is typical for changes in progress which are picked up by the younger speakers and do not find their way into the older speakers' more set habitus (Lenneberg 1967). *Like*'s age-patterned distribution is statistically significant at the .01 level (with  $\chi^2(1) = 34.086$ ).



To give some time depth to this study, I will now compare my findings with a later corpus of US English. The following table gives the remarkable numerical difference between the Switchboard (recorded in 1988-1992) and Singler's (1994-1999) NYU corpus. Note that I have included only speakers aged 15-51 from the Switchboard corpus in order to assure compatibility between the two corpora.

**Table 6.14: Comparison of quotative systems: *Like*-frequency in the NYU corpus (Singler 2001) and the SWB**

	SWB	NYU
young	13	60
old	4	12

Note the amazing difference in frequency between corpora which are not even 10 years apart. *Like*-frequencies soar from 4% and 13% in 1988-92 to 12% and 60% in 1994-99. But caution has to be exerted when comparing those two data sets. Firstly, they pertain to different genres. The NYU corpus consists of sociolinguistic interviews, which were mostly conducted between people who are friendly with or even family members of the interviewer. In contrast, the SWB data is made up of telephone conversations between strangers. Secondly, the populations the corpora are sampled from do not have the same regiolectal origin. The Singler corpus consists mainly of NYC speakers, whereas the Switchboard corpus represents a cross-cut through the 7 main dialectal areas of the US, Northern, South Midland, North Midland, West, New England, New York, South.

But due to the scarcity of comparable data, which obviously stems from the fact that this feature is still quite new, I nevertheless decided to compare the two corpora. The results are even more impressive when we consider the fact that New York is the dialectal area in the Switchboard corpus with the lowest *like*-frequency overall (5%). Such an investigation across real time shows the impact with which *like* is spreading across the quotative system.

In order to investigate *like*'s patterning across apparent time, we need to plot the occurrence of the variant on a linear age-scale (for a discussion of the limits and assets of apparent time studies see Bailey et al. 1991, Labov 2001). For the following discussion, I cross-correlated *like*-output (in % frequency) with numerical age. One data-point will henceforth represent the frequency of a variant as produced by any



York data is represented in my data as well. I will leave all speakers of the age range 20 in the calculation of *like* according to age brackets.

But let us now discuss the older speakers. Due to the high *like*-frequency amongst the 59-age group, the association *like* ratio-age would give a U-shaped curve because *like*-use increases again in the lowest age-group (consider the broken line in Figure 6.4). This would result in a patterning that is reminiscent of age-graded variables. Given the non-attestation of *like* in earlier accounts (prior to Butters 1982), this outcome is quite unlikely. Note again that I am not trying to argue that non-attestation implies non-existence. But two facts make it very implausible that *like* has been around long enough to have been in the quotative system of the speakers born in the 1930s ever since their youth: (i) neither the OED, nor any other dictionary before the middle of the nineties has picked up on *like* and (ii) the overwhelmingly unanimous impressionistic judgements by speech communities from both sides of the Pacific that *like*, as a quotative, is a newcomer. But obviously, only a full investigation of all diachronic evidence can lead to a scientifically founded claim about *like*'s apparent non-existence prior to 1982. However, any further diachronic investigation exceeds the frame of this thesis.

The question remains: what are we to do with the high *like*-frequency amongst the 59 year olds then? After checking the data, it became clear that the 8% *like* use for the oldest age-group is due to an outlier. The low numbers of overall tokens produced by the highest age groups are prone to distortion by single speakers who patterns atypically. The high *like*-frequency for speakers aged 59 is caused by one speaker who produced 2 tokens of *like* (13% of his only 15 quotatives). None of the other speakers in this age group (and the surrounding ones) produces any *like*-tokens at all. We do not know anything about this particular speaker's demographic profile as the Switchboard restricts the social information given to age, gender and educational level. But if we assume that this speaker has borrowed *like* as a lexical item from the younger generation (cf. Singler 2001) and hence eliminate him, the *like*-ratio for the oldest age-bracket falls to zero.<sup>17</sup> We end up with a very neat curve

---

<sup>17</sup> Obviously, there have been ample comments in the literature that outliers can teach us important facts about the behaviour of the community as a whole, namely what it is that they do differently. And indeed, it is an interesting fact to note that even older speakers acquire *like*, even if the use of this newcomer item is atypical for their age-group. But admittedly, singular cases can also detract from general trends and might skew the statistical analysis. Hence, having noted the outlier at age 59, I will henceforth eliminate this speaker and proceed to an investigation of the general patterning.

which shows a clear correlation of falling *like*-use with increasing age, as can be seen in Figure 6.4 (full line).

**Figure 6.4: Distribution of *like* –frequency over age-groups**

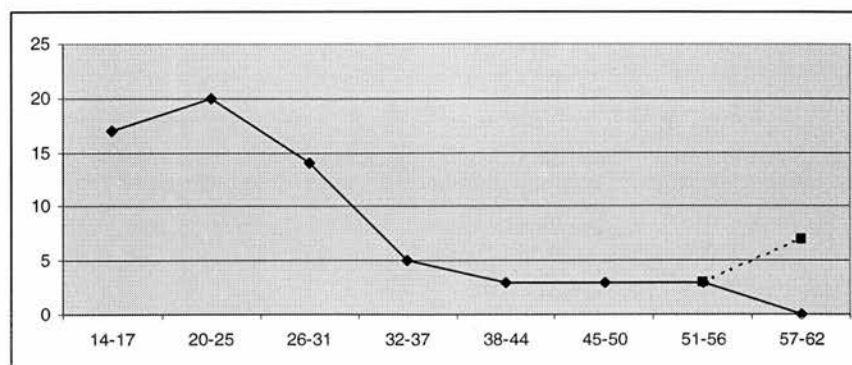
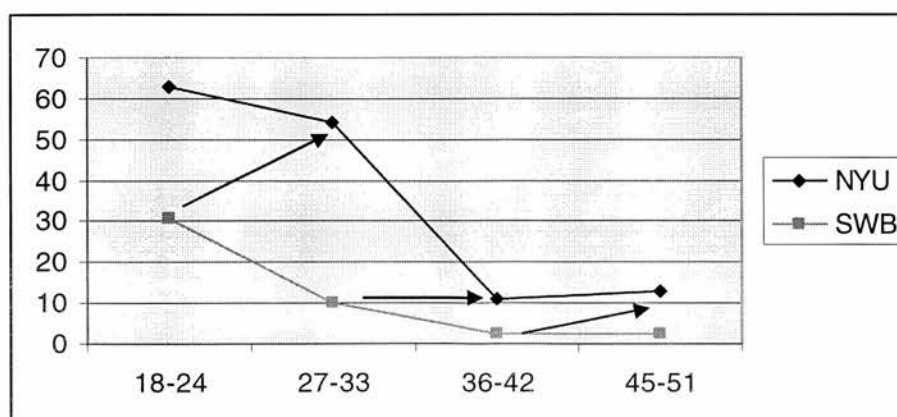


Figure 6.4 depicts this monotonic relationship between age and *like*-ratio. Once the one outlier is taken out (the dotted line), the *like*-curve (full line) runs smoothly except for the youngest speakers, who lag slightly (by 3%) behind in *like*-frequency, a pattern which seems typical of changes in progress (Eckert 2000, Guy 1990:52, Guy, Horvath, Vonwiller, Daisley and Rogers 1986:36). Without the outlier, the  $r^2$  goes from .027 ( $p = .000$ ) to an  $r^2$  of .031 ( $p = .000$ ). This means that 3% of *like*'s variance across speakers can be explained by the factor numerical age when we work with 8 age groups (cf. Table 6.7).

The following figure shows the possible development of *like* by age. I have plotted the Switchboard and the NYU corpora together into one graph in order to depict incremental *like*-frequencies across real time.

**Figure 6.5: Incrementation of *like*-frequencies in 7 years (as of a comparison between the Switchboard corpus and the Singler (2001) NYC corpus)**

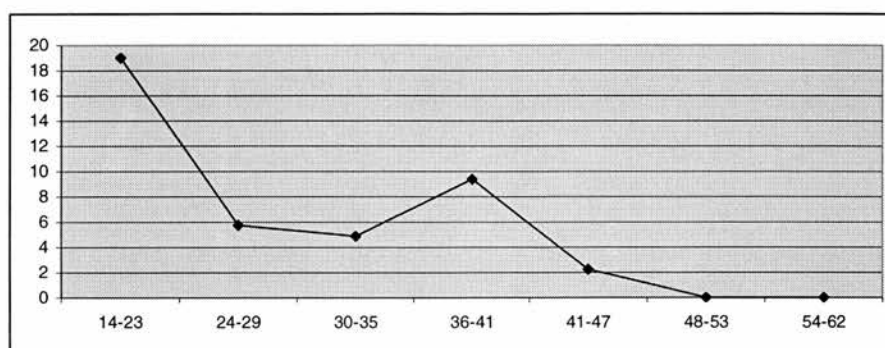


The NYU and the SWB corpus lie on average 7 years apart. The age brackets (as given by Singler 2001:277) are 9 years apart. The roughly equivalent difference in increments in numerical age and years of corpus collection make possible a direct comparison between an age bracket in the earlier SWB with the diachronically next one from the NYU corpus. Such a direct association of age-ranges with two diachronically separated corpora combines the merits of apparent time and real time investigations (Cedergren 1984). We can see the incremental frequencies from the youngest age group (a huge jump of 32%), while older speakers' linguistic habitus does not change much. The isogloss mentioned by Singler (2001) moves from the 27-33 year olds to the 36-42 year olds. Speakers beyond this cut-off point do not partake actively in the introduction of newcomer *like*, neither in 1988-92 (SWB) nor in 1994-99 (NYU).

Next, I will proceed to do the same statistical break-up across chronological age for *go*. Table 6.13 shows that in the USA, *go* is mainly a feature of younger people's speech with 7% as opposed to only 4% for the older speakers. This patterning is statistically significant at the .001 level ( $\chi^2(1): 7.519$ ). But note that when we applied chronological age as a predictor variable in a regression analysis, we only got a very low  $r^2$  value of 0.5% (cf. Table 6.7). By plotting the *go*-ratio on the dimension age rather than working with a binary category only, we will be able to see what underlies this age differential. Is the effect in US English due to age grading with falling values for increasing age and a prototypical tail for the oldest speakers? Or do we witness a generational change?

But note that with a proportion of only 5.8% of the (say) variable overall, *go* is not a very frequent variant in the United States. Due to the paucity of tokens, I have refrained from including a scatterplot by numerical age. Firstly, such a depiction would result in many zero age-ranges. Secondly, small token numbers would lead to a high risk of skewing results in the few age ranges that do display tokens. It seemed more meaningful to depict *go*'s patterning (in % frequency) across 7 age groups, which the following line graph shows.

Figure 6.6: Distribution of *go* across age groups



The line-graph we see in Figure 6.6 shows zero values for the lowest two age groups. Note that this outcome is in contrast to *go*'s patterning in British English, which I will argue is highly indicative of age-grading. But in US English, the oldest two age groups produce zero tokens of *go* (with  $\Sigma$  (all quotatives<sub>54-62</sub>) = 65 and  $\Sigma$  (all quotatives<sub>48-53</sub>) = 152). This non-occurrence in the repertoire of the oldest US speakers questions whether *go*, although recorded in British English since an early date (OED 1789), had ever gained any real frequency in the US before it became a trend at some point when the 41-47 year-old speakers were young. I will come back to this question during the discussion to follow. But first, I will discuss the statistical output in Figure 6.6 as a whole.

From the age-group 41-47 onwards, *go*-ratios rise with declining age and reach a peak in the age group 36-41 with close to 10%. Until here, the patterning looks like a typical change in progress: we notice zero tokens for older speakers, a start of the trend when the generation of 41-47 year old speakers were young, and rising ratios with progressively lower age.

But from the peak years (36-41) onwards, *go* ratios fall to 5% in the age bracket 30-35, reach a plateau for speakers 24-29 and rise again for the youngest speakers. This pattern is underlined by Blyth et al's (1991) findings that *go* is most frequent in the youngest age group (20-24, VARBRUL weight .799), not frequent in the middle age group (27-32, .243), which is exactly where Figure 6.6 shows dipping frequencies, and frequent again in the oldest (38-72, .439) age group.

Why, after *go* reaches its local high in the age group 36-41, do we see the line fall and stagnate with the younger speakers? This could be because *go* has reached its peak and is slowly dying away (see Labov 1994 for completed changes). But note

that this fall takes place in exactly the same age bracket where *like* comes to speed in the quotative system (consider Figure 6.4). This seems to indicate that, as Macaulay (2001) suggests, *go* became popular 25-30 years ago among younger speakers in the US and enjoyed quite some popularity afterwards (see the high values for the speakers in the age bracket 36-41). It was then challenged by a new item that hit the linguistic market, *like* (first mentioned as a trend that has been around for an indefinite amount of time in 1982). We can now assume that the oldest age-group that retreated from *go*-usage and took on *like* were the age group 30-35 (the then 20-25 years old) – note the dip in the curve in this age bracket. Hence, the fall in *go*-ratios for the 24-35 year olds can be accounted for by those speakers picking up the *like*-trend and changing their choice of non-canonical quotative from *go* to *like*. Note the consistency of this claim with Blyth et al.'s report in 1990 that in their corpus the absolute oldest speaker to use *like* was 39.

The change of preference from *go* to *like* is facilitated by the fact that both share the underlying feature [-canonical]. Speakers, when wanting to quote non-canonically, have had another option at their disposition since the early 1980ies. What could have been going on in the quotative system is the reallocation of a semantic feature [-canonical] from an old quotative (*go*) to an innovative variant (*like*). The newer option is thus taking away some of the vernacular ground that was previously covered by *go*. *Like* is intruding in *go*'s semantic-pragmatic sphere, which they now have to share.

Had *like* taken over and pressed *go* out of the system, we would witness a continuous fall in *go*-frequency from *like*'s introduction onwards. But, after a strong dip in *go*-ratios in the age brackets 24-29 and 30-35, we notice that the youngest age groups found renewed interest in the use of *go*. *Go*-frequencies even rise steeply for the youngest age group, with a climb up to 18%. Judging from Figure 6.6, the conjecture that *like* is going to replace *go* in American English (Bakht-Rofheart 2001, Macaulay 2001, Singler 2001) is not sustained by my data.

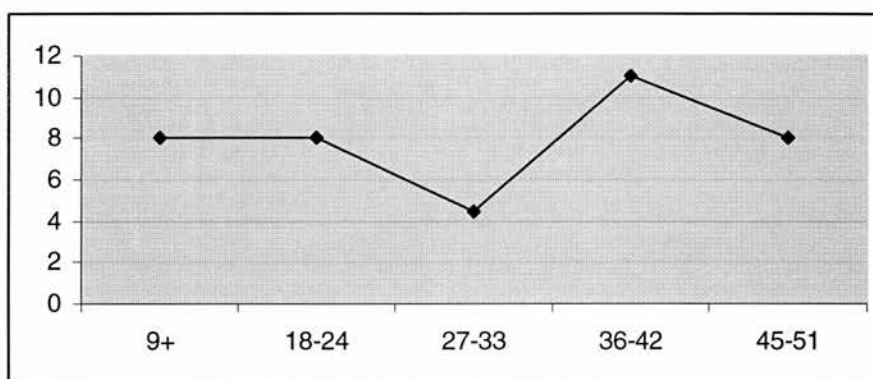
But the question presents itself: what do we make of the steep rise in the youngest age-bracket? For the 14-23 year old speakers, the *go*-ratio soars higher than any frequency it attains throughout the entire age-spectrum. I suggest that the boost in *go*-production amongst the youngest speakers might be the outcome of a

heightened awareness due to the more frequent use of non-standard quotatives, notably through newcomer *like* but also, more recently, *be all* (Bakht-Rofheart 2002, Singler 2001 etc.). It is very conceivable that the high frequency and salience (cf. the numerous comments and stereotypes cited in the literature and in the public domain such as the world wide web) of those non-canonical items have triggered a renewed wave of interest in *go* amongst certain sociolinguistic groups. *Go* is a variant that is operating well above the level of awareness. Such variants are prone to be adopted, enhanced or dropped across the linguistic community according to linguistic trends. I assume that it is not impossible that the vigorous new quotative *like* has thus sparked *go*-use through their common association with the underlying feature [-canonical] (the same can occur later for *be all*). Following this line of argumentation, the fall and resurgence of *go* can be interpreted as being due to a reallocation of [-canonical] from one surface item, the older *go* to its rising counterpart *like* and back.

Obviously, this study can only shed light on *go*'s patterning as it manifests itself during a very limited time frame. But a small diachronic window can be opened by comparing my Switchboard (1988-1992) data with the later NYU data (Singler 2001). Due to the aforementioned sampling differences and the small token numbers, it does not make sense to represent the two data sets in one graph. Also, in addition to the above mentioned differences, the NYC corpus is in many other ways not comparable to the Switchboard corpus: while the NYC operates with spaced-out age brackets, we find continuous age sampling in the Switchboard. Furthermore, the NYC corpus did not sample any speakers over 51. Also, the NYC corpus reveals much lower frequencies for *go* overall. But bearing those reservations in mind, it is noteworthy that we find the same apparent-time curve in the two corpora. Figure 6.7 represents the patterning of *go* in the NYU corpus.



Figure 6.7: *Go*'s patterning across age (data: Singler's NYC corpus)



The pattern noted in the Switchboard with a dip in *go*-ratios for speakers in their late twenties and early thirties with rising ratios for younger and older speakers found in 1988-1992 is replicated almost exactly 7 years later. Note that while we find a levelling out for the oldest speakers in 1988-92, in 1994-99 even the oldest speakers have 8% *go*-frequency. Given the fact that the NYC corpus is 7 years younger, the oldest age group in 1994-9 (45-51) represents the age slice where *go* boomed in (36-41 in 1988-90) and is thus exactly the generation in which we would expect high *go*-values. Hence, the peak in the 36-42 age group which we found in 1988-1992 stays in the diachronically later NYU corpus as well as the dip for the speakers slightly younger than the boom generation (27-33), which stems from higher *like*-ratios in the system of the younger speakers.

Hence, the whole line-graph (Figure 6.6) for *go* can be explicated as follows: speakers in the oldest age brackets did not take part in the *go*-trend when it was first introduced into the quotative system some 25 years ago and consequently show zero *go*-values today (age 48+). The speakers who espoused the variant when they were in their teens and twenties are middle aged today and continue to use *go* with high frequencies today (Lenneberg 1967). Among those speakers is the age-group where *go* is still peaking (36-41). Speakers born subsequently (age 35 and younger) have reduced *go*-ratios because they are using yet another, newer, quotative, *like*. *Go* bottoms out and stays stable at a very low level for an interval of 10 years (ages 24-35 at 5 and 6%). A glance at Figure 6.4 shows that it is exactly in this age-bracket that *like* was introduced into the quotative system in US English. It thus seems

appropriate to assume that *like* took away some of *go*'s ground, which results in lower *go*-ratios. But for the youngest speakers, *like* might have broadened the path for *go* by dragging it along in its current wave of popularity. Consequently, we find steeply rising *go*-ratios in the youngest age bracket in Figure 6.6. Singler's follow up study seven years later gives evidence that *go*-production stayed stable with the second youngest and the youngest generation producing *go* at about a rate of 8%.

Let us for a moment investigate the different possible scenarios that could have given rise to the patterning seen in Figure 6.6. If we disregard the recent developments connected with the rise of *like* and only consider *go*'s patterning in the age-groups from 36+ upwards, we see a steady decline in *go*-use with increasing age. When interpreting the patterning, we are faced with the following two unknowns:

(1) We lack the concrete diachronic evidence that would pinpoint the time when *go* arrived in the US quotative system.<sup>18</sup> Even though we can assume that quotative *go* is understood even by speakers who use it with zero frequency, there is no evidence that it was actually used in this variety until speakers in the age bracket 41-47 started the trend. Hence, *go*'s entry point could have been when the generation of 53 year-old or younger speakers were adolescents (note the zero-frequencies in the oldest age groups). Alternatively, given its earliest attestation, *go* could have been in use in the US for decades. Even though *go*'s first appearance in a source text in British English was after the War of Independence (1775-1783), we can nevertheless assume that there has been enough contact, face to face, and later especially via the media, for a variant such as *go* to spread.<sup>19</sup> Note in this context that Meyerhoff and Niedzielski (2003:26) contend that external influences on US English have largely been ignored in the sociolinguistic literature. They propose to compensate for this shortcoming by considering all cross-variety transfers as parts of one subordinate phenomenon. "If we conceive of the changes occurring across all varieties of English in terms of a broadening of the vernacular base, this avoids the (we think, highly

---

<sup>18</sup> This claim concerns quotative *go* per se. I am not considering the content of the quotative with respect to sound, voice effects or lexical content. All that matters for the present discussion, is the ability to frame quotes of some sort with the lexical item *go*.

<sup>19</sup> Of note also is of course also the steady stream of immigrants from the British Isles into the US even long after it became a separate nation.

problematic) claim that Americans are in any way trying to sound Scots (with *wee*) or British/Australian (with *wank(er)*”. Alternatively, it is also possible that quotative *go* had been in the system for some time before it was mentioned in the 1791 source document of the OED and that it had found its way into the system of US English with the settlers even before American independence. In both cases, we can assume that *go* has been a member of the US quotative system for centuries.

(2) As Hockett (1950) pointed out, an apparent time study investigates the “distribution of linguistic variables across age. If we discover a monotonic relationship between age and the linguistic variable, or a significant correlation between the two, then the issue is to decide whether we are dealing with a true change in progress or age grading”. On the same lines, Labov (2001:83) has shown that falling frequencies in apparent time can give rise to two interpretations: either the change is at the individual level where the community does not change - a situation of age grading. Or the community changes with individual speakers staying stable across their lifetime – a generational change.

The question is whether underlying *go*'s distribution there is the regular process whereby speakers embrace different quotative variants with age (Chambers 1995:189, Macaulay 1977) in successive generations. If this were the case, we would expect *go* to be a variant which is endorsed only by young speakers and regularly and predictably gets lost from their linguistic habitus during their involvement in the linguistic marketplace. More canonical quotatives such as *say* or even non-canonical ones such as the unframed strategy would then take over its role. Given how long ago *go* was attested in British English (OED 1791) relative to its non-attestation in US English as a full quotative until 1982<sup>20</sup>, do we have to assume, though unnoticed, it was always around as an age graded feature until it was finally officially “discovered” by Labov/Butters?

Alternatively, in the US, *go* could be considered a change in progress in the age group from 53 upwards until its spread was halted by a younger competitor, *like*. Note that research on the quotative system of African American vernacular speakers

---

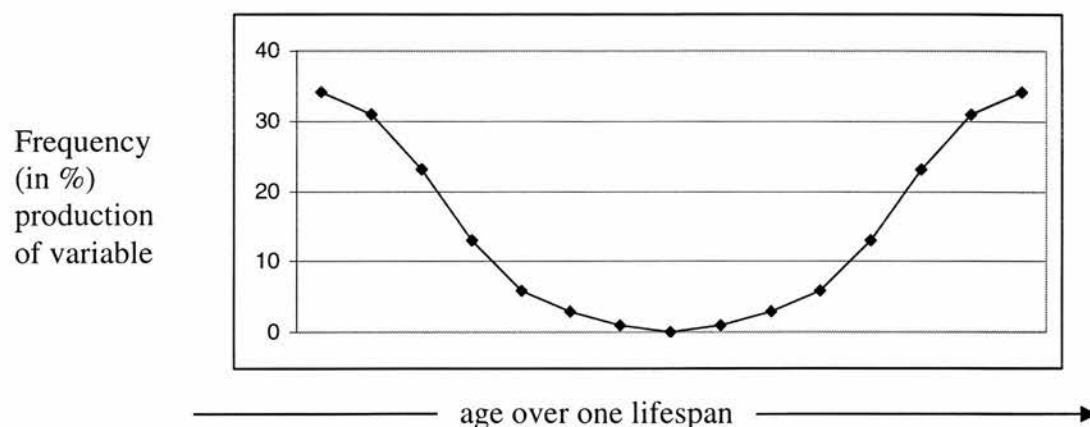
<sup>20</sup> The first mention of quotative *go* with the mimetic sound-rendering function that I know of is Labov (1972c:372) “In BEV, objects that do not speak but make noises are not said *to go X* but *to say X (...)* in White vernacular, *people go powww!* with their fists”.

in rural Texas underlines this hypothesis. In Cukor-Avila's (2002) data, only one of the pre-WWII speakers uses *go*, and his frequency of production is only 0.2%. When Cukor-Avila split up the data into three time periods, she found that this speaker actually used *go* only in the latest period, in 1997. This gives reason to assume that the few *go*-tokens in his system are lexical borrowings from younger speakers. As *go* had gained some frequency by 1997, we could imagine that the variant was sufficiently used in his surroundings to make such a borrowing possible. Of the post-1970s speakers, all (except for one) produce *go* with frequencies between 1.3 and 4.7%. Thus, again, it seems that *go* as a quotative has not been around as long in the US as it has in British English. Maybe Singler (2001:257) is right when he says that "the twentieth century saw the introduction of three quotatives into American English. First there was *go*, whose appearance appears to date at least as far back as the 1940s and 1950s (...)"? Note that the OED does not give any American English examples in the section on *go*'s quotative function but it does not specifically annotate the usage as British only (note in contrast the annotation *N. Amer. colloq* in the entry for *like*).

Bearing in mind these parameters, I will now consider *go*'s patterning across time in more detail. But note that the situation is further complicated by the fact that a number of competing variants interact within the (say) variable. In such a scenario, where we do not witness a replacement of one variant by another in a set of binary choices, it is much harder to know whether what we see synchronically is a change in progress (cf. Singler 2001). I will now discuss the possible scenarios that could have given rise to *go*'s patterning.

Let me first scrutinize the possibility that *go* is age-graded. Figure 6.8 displays the classical (roughly) symmetrical bi-modal pattern for age grading as discussed by Downes (1984) and (Holmes 1992), (see also Labov 2001, McMahan 1994).

Figure 6.8: Classical age grading pattern (Simplified)



When we look at *go*'s patterning across apparent time in Figure 6.6 we notice a steady decrease in *go*-frequencies with increasing age from the age-groups 36+ upwards. But the oldest age groups do not display any rising frequencies in *go*-production. This "tail" has been considered a typical feature of age-grading whereby older speakers are to some extent regaining the variables they have suppressed during their involvement in, what has been called the "marché linguistique" (Bourdieu and Boltanski 1975, Sankoff and Sankoff 1973). The reclaiming of nonstandard variants after retirement age has been considered an indicative cue for identifying a pattern as age-graded. Looking at *go*, we find that, in fact, speakers above the age of 48 do not produce a single token of it (out of 153 quotatives). But note that the absence of a tail is not a decisive criterion for the interpretation of an apparent-time pattern as age-graded. As Guy and Boyd (1990) have shown, not all cases of age grading actually display a tail. Their research on patterning of *-t,d* deletion in semi weak verbs across the age range 4-65 shows a regular decline in probability of deletion with increasing age.

Notice further that age-grading, which is "a regular change of linguistic behaviour with age that repeats in each generation" (Labov 2001:46), does not necessarily hinge on the fact that a variant has the underlying trait [vernacular], which leads to suppression and later endorsement of the particular variant, depending on the speaker's status with respect to the linguistic marketplace. Rather, there is a possibility that *go*'s falling frequency across age is due to other another underlying feature, which this variant carries and which is considered less appropriate for middle-aged speakers. Indeed, social attitudes research has revealed that *go* is

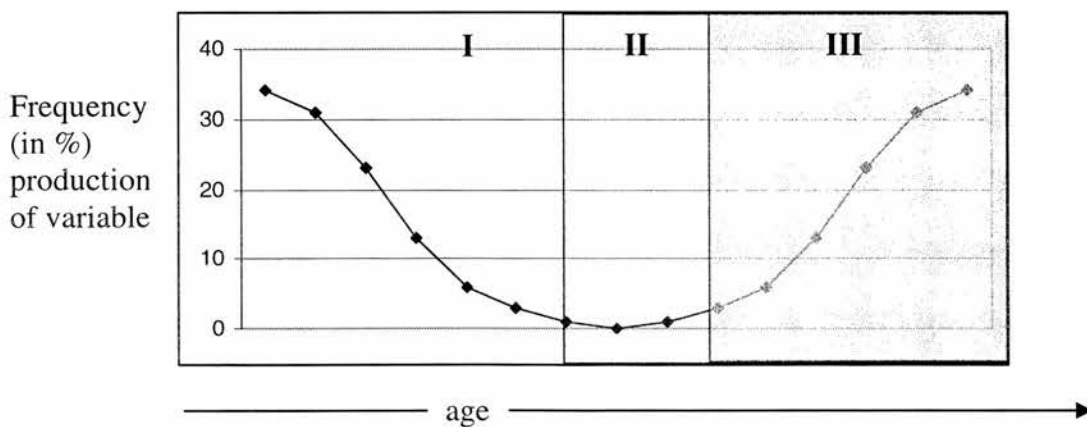
perceived as [+young] (Buchstaller 2003). Hence, it would be no surprise if older speakers discontinue a variant marked for young age once they reach middle age and that older speakers do not revert to *go*-use after their retirement.<sup>21</sup>

If we accept the age-grading scenario for a moment, we can assume the following diachronic scenario: quotative *go* has been around in British and US English (possibly since 1791, possibly before or after) and it patterns according to age. Note that also Singler’s apparent-time data could be interpreted as age grading with younger speakers embracing *go*, middle aged speakers reverting to other variants (such as *like*, albeit only the speakers ages 27-33), and the oldest speakers again showing higher frequencies.

But again, we do not have any evidence of *go*’s existence in the US before 1972, or, alternatively, in line with the concept of apparent time, when the speakers aged 53 or younger were young. Consequently, if we interpret the falling frequencies in the apparent-time data as an indication of a generational change, we can assume that *go* is a change in progress which started about 45 years ago (note again the zero frequencies for the 53+ year olds).

The problem with this interpretation is that we still cannot account for the attestation of *go* in 1791 (nor for much anecdotal evidence that *go* has indeed existed at least since the 50s in the US). In order to get around this dilemma, let us for a moment assume that there is a spread-out curve underlying the chronological development of *go*, as depicted in Figure 6.9.

**Figure 6.9: Generational change captured at two stages**



<sup>21</sup> Notice that this argument does not hold generally. In British English, *go* can also be used by the oldest age group (consider the patterning in British English as in Figure 6.23).

Figure 6.9 depicts a generational change, stable at the individual level and unstable at the community level. The unshaded area (I) spans one lifetime. The shaded areas on the u-shaped curve depict the zones which are inaccessible to the analyst in an apparent time study, such as the present one, because the speakers who would feature in these age brackets have already died. The curve represented in I depicts *go*'s patterning as we know it from Figure 6.6. But I suggest that we should ask whether zero production by the oldest speakers necessarily means that this variant has never been in use before. Let us for a moment play with the question of what would happen if there was a "wave" underlying *go*'s chronological development and that this curve was too widely spread to be captured full cycle by one single life-span. In this case, the cut-off point would occur at the lowest point of the curve. The information needed to show whether frequencies stay at zero would not be accessible to us from an apparent time study. This is exactly the situation we are in with respect to *go* at the moment: we need real time evidence from earlier time slices. If we can be sure that we do not have any previous attestation of the linguistic variant (such as *like*), we can fairly safely assume that a monotonically falling line is evidence of a change in progress. In cases such as *go*, though, where earlier investigation has captured the variant, the question arises whether non-attestation in older age groups is not simply an artefact of the fact that our sampling does not span the whole spread of the curve. In the current situation, we simply cannot safely exclude the possibility that we might not have the situation depicted by the shaded areas (II and III) in Figure 6.9. It might be the case that a sampling error due to a spread-out u-curve falsifies our results within an apparent time study only. If we could have sampled the 120 year old speakers, we would have been able to see a rise in frequency. And indeed, *go*'s patterning in apparent time in British English looks like II.

This effectively means that what underlies *go*'s patterning could be a change in progress but with a very rapid ebb and flow between the variants. There is a possibility that underlying the chronological development of *go* is a repetitive u-shaped curve. In this case *go*-production might have been going up and down in waves.

Figure 6.10: Chronological development over several generations

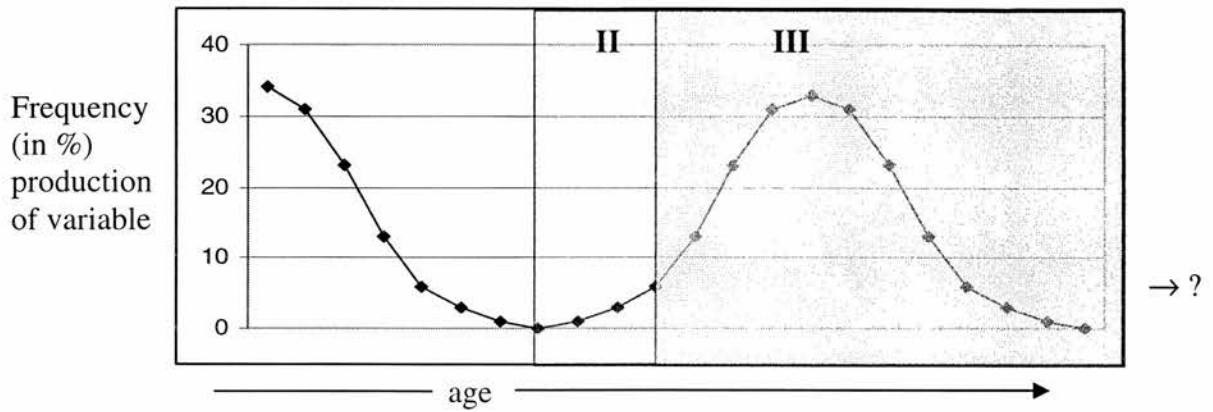


Figure 6.10 depicts a (hypothetical) situation whereby, across time, *go*-production comes and goes. The wave-like status of variables is a feature that was discussed recently at the fourth Language Variation and Change conference with respect to spread of the labiodental variants [f] [v] for interdental /θ/ and /ð/. In this particular case, Kerswill’s (2003) data suggests a slow diffusion from London starting in 1850 and accelerating towards the north and west, reaching the big towns first via the major rail lines (Bristol 1880, Reading 1950, Norwich 1960, Hull 1970, Newcastle and Glasgow 1980 etc.). However, the fact that the development has been mentioned at various times and in several places before the main diffusion pattern from London leads to the speculation that some natural changes come in waves. In the case of the labialization of fricatives, it seems that localised wave-like developments were superseded by a nation-wide spread from London outwards. Once the variant became stereotyped in London in the 1880s and acquired a critical weight, it spread to the big cities. Similar points have been raised by R. Cameron (2000), and during the discussion following Wolfram’s comment concerning the ebb and flow of linguistic variables at the NWAVE 2003.

It might well be that a more thorough investigation into *go*’s real time patterning would also show a steady presence across the centuries. Hence, if we accept this scenario, we would interpret the data as a change in behavioural patterns in which a variant that has been latently around for ages is picked up and highly favoured by a generation of speakers – a fad. There is indeed reason to believe that *go* has been latently present in the system, peaking and declining across time (Joan Beal, p.c.). Some generations pick up on it and make it a “trend”. This change in



variant preference is brought to completion when a certain saturation level (viz. a certain frequency, cf. Labov 2001) is reached, whereupon the variant is disfavoured by subsequent generations of speakers and replaced or superseded by some newer variants (such as *like* in our case).

If we sampled across the population during such a trend, we might notice a pattern across apparent time which shows all the features of age grading or, alternatively, of a change in progress, depending on the spread of the curve of the wave and the window we employ (cf. Figure 6.10). The important feature in this diachronic scenario, though, is that the variant (e.g. *go*, labialized fricatives) stays latent in the linguistic repertoire of the community, never fully dies out and can potentially be picked up as a fashion again.

Two important pieces of evidence speak in favour of such an interpretation: The first is the fact that the youngest speakers use *go* again with higher frequencies after a significant drop in the curve (cf. Figure 6.6). The high frequencies amongst the youngest age groups could then be interpreted as another peak of the wave-like development. Indeed, seven years later, Singler's NYC corpus provides evidence that the curve might already be levelling out again (Figure 6.7). Hence, *go*'s development starts to look like a change in progress but with a very rapid ebb and flow of variants. The second piece of evidence is the differential distribution of *go* across age in British English and in US English (see 6.5.3), which could be seen as two stages in the continuous wave-pattern, with BrE representing the darker shaded pattern and the US representing the lighter shaded pattern.

But obviously, the question of whether *go* as discovered by Butters (1980) really first started to appear as a quotative item in US English can only be fully answered on the basis of diachronic evidence. Only a larger time frame than one lifespan can show that what we are witnessing in 1988-92 is age grading, a change in progress or a wave-like development. In order to develop a complete sociolinguistic history of *go*, a real time study with a diachronic corpus spanning the time since its first attestation would have to be conducted. Unfortunately, due to the limited scope of this thesis, only educated guesses can be made about what might lie behind the

pattern we see in the apparent time study on the basis of synchronic data and of the diachronic evidence cited.

As the following discussion will reveal, we are actually not telling the whole story if we are only picking up on the patterning of one or two new or age graded variants. In order to do full justice to a situation of language variation or change, we need to describe not only the new form, but all variants that make up the variable. This is especially the case when we are faced with a situation like the quotative pool, where there is a whole set of variants. The following paragraphs will show first the interesting interplay between the new variants and the older non-canonical option, unframed quotes. Finally, I will concentrate on the interdependence between the non-canonical and canonical variants.

The intimate interplay between *like* and *go* raises the question in what way the non-lexical non-canonical option is affected by the patterning of its lexical variants. After all, if the frequencies of one variant go up, some other variants' frequencies must go down. Obviously, only a study going back centuries could give conclusive evidence of what happened once *go* was introduced into the quotative system (if, as I assumed earlier on, *go* has been around longer than a lifespan in US English). Bearing this shortcoming in mind, I will now conduct an apparent-time analysis on the patterning of the unframed variant. The questions to answer in this respect are the following: what is the interaction between the three non-canonical variants, *like*, *go*, and unframed quotes? Do *like* and *go* take over some of the territory previously held by unframed quotes or do they intrude into the *say*-space?

When we look at the unframed quotes' patterning in the US, we find that younger speakers use this variant significantly more with 23% versus older speakers with 20% (n.s.).

Above, I have discussed how *like* and *go*'s patterning across apparent time can be explained by assuming that the underlying feature [-canonical] has been transferred back and forth. We saw that because those variants share an underlying semantic element, *go* got some impetus from *like*'s high frequency among the speakers in the youngest age-group. But notice that there is reason to believe that the quotative option which allowed the feature transfer in the first place is the non-

lexical variant, unframed. Indeed, I would argue that it was through the reallocation of [-canonical], first from unframed to *go* (when *go* first arose as a variant), and consequently from *go* to *like* (at *like*'s introduction into the pool of quotatives), that the non-traditional lexical quotatives made their entry into the quotative system in the first place.

Obviously, the OED does not give any information with respect to zero items. Also, non-canonical variants do not appear in writing as frequently as standard ones and often with a significant time lag (Romaine and Lange 1991, Singler and Woods 2002). Consequently, the claim that unframed quotation was diachronically prior at least to *like* and *go* cannot be sustained by data from the OED but is based on 2 arguments: Firstly, we can rely on occurrence in written fiction. Page (1973:26) shows that in *David Copperfield* “2 very short sentences are not explicitly attributed to a speaker”. Note that I do not have historical evidence for unframed quotes' diachronic primacy that goes beyond 1848-1850 but in any case it predates the first mention of *like* by far. Evidence of the existence of unframed quotes in the nineteenth century literature also refutes Macaulay (2001:16), who argued for a developmental process along the lines say > say like > be like > be > zero.

The second line of argumentation relies on arguments made in the fields of grammaticalisation and creolistics. Buchstaller (1999), Couper-Kuhlen (1998), Hopper (1987, 1998), Mann and Thompson (1986) and Thumm (2000, in preparation) and many others underline the fact that linguistic information is usually highly contextualized and implicitly given by the textual environment. There is evidence that phylogenetically, as well as ontogenetically such information is generally not explicitly expressed using specific lexical means but only slowly emerges out of regular association between contextualized notions and corresponding surface expressions. Morphological matter regularly or habitually found in certain contexts becomes the surface expression of these underlying notions (cf. the notion of “emergent grammar” Hopper 1987, 1998). Indeed, the genesis of pidgin and creole languages (Hopper and Traugott 1993, Thomason and Kaufmann 1988), gives ample evidence for the process whereby linguistic structure boils out of discourse (cf. also Buchstaller 1999, Bickerton (ms): 2ff, Sankoff 1980b, Thumm 1999).

Unframed quotes, i.e. quotes not marked by lexico-syntactic means, are highly contextualized and rely on other means rather than graphic introducers. Contextualization cues (Gumperz 1989, 1992), such as “a greater expressive power of the human voice” (cf. Tannen 1986:323 for the concept of a continuum of quotative options) indicate to the interlocutor a change in “footing” (Goffman 1981) and open an “embedded mental space” (Fauconnier and Sweetser 1996). For a discussion of the means speakers have at their disposition in order to signal embedding into the space “quote from same or other” and how conversely interlocutors can interpret speech as quoted, see Fonagy (1986), Klewitz and Couper-Kuhlen (1999) and Kvavik (1986).

While I do not have any literary evidence which would underline a claim that unframed quotes predated even the oldest, canonical lexical quotative *say*, such a scenario would typologically be conceivable and in line with the above argumentation. For the present purposes, I will assume that unframed quotes, the first non-canonical quotative option and variant to standard *say*, predates quotative *go*, which is attested as early as 1791 as well as *like*. If we accept this scenario, we can suppose that the feature [-canonical], present already in the form of non-lexical, unframed quotatives, got transferred onto lexical *go* as it arose as another non-traditional variant.

Figure 6.11 gives the line graph of the patterning of unframed quotatives across numerical age.

**Figure 6.11: Distribution of unframed quotes across age**

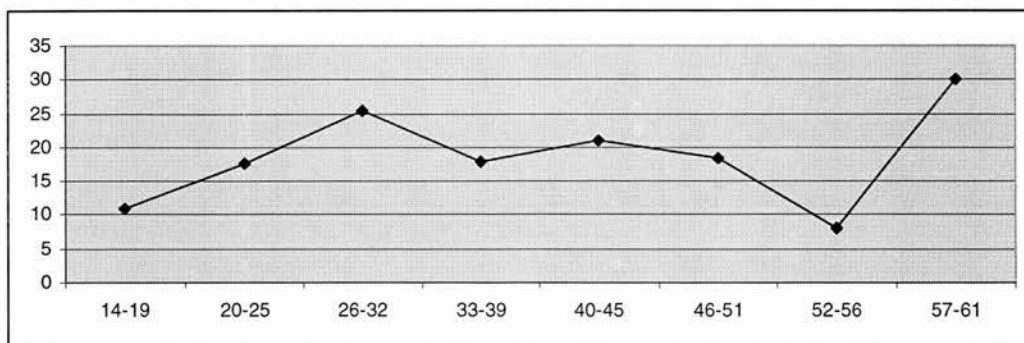


Figure 6.11 shows, starting from the left, there is a rise in non-canonical unframed quotes with increasing age up to the ages 26-32. After this, the frequency falls until

we witness a steep increase until 30% in unframed quotes amongst the oldest speakers.

This fall in vernacular variants from a certain age onwards and a sudden jump for the over 55 year-olds has been shown to be typical for age-graded variables (Downes 1984, Holmes 1992). The obvious tail in the graph can hence be explained in part by the fact that as older speakers are relieved of the pressures to conform created by the job market, vernacular forms start to creep in again. Also, in the case of the competing (say) variants, we know that the oldest speakers do not yet have *like* as a non-canonical variant at their disposition. And *go* seems to have been a low frequency variant during their formative years (cf. Figure 6.6). Hence, speakers over 42 do not have any noteworthy alternative non-canonical variants in their linguistic repertoire of quotative devices. When using vernacular quotative options, all their non-canonical output will be unframed quotes. It is because the unframed strategy has the competitor variants *like* and *go* amongst the younger speakers that the patterning of non-canonical strategies shows the highest frequencies in the oldest age group (cf. Downes 1984: 191).

This situation of competing variants is directly derivable from the graph: Notice the first dip in the unframed curve amongst ages 33-39. A comparison with the *go*-curve (Figure 6.6) reveals that this is exactly the age-range in which *go* peaks. Furthermore, from the 25% peak in the group 26-32 onwards, the ratio of unframed quotes falls until it reaches only 11% with the youngest age group. A direct comparison with the curves for other non-canonical quotatives, *like* (Figure 6.4) and *go* (Figure 6.6), shows that this downwards trend is concomitant with their rise in frequency. Consequently, there is reason to believe that the advent of *like* and *go* has led to the reduced frequencies for the unframed strategy that we witness amongst those youngest speakers.

We saw earlier on that when *like* first reaches full force amongst the 20-35 year-olds, it takes over some of the ground of quotative *go*, which subsequently drops by 3.5%. Then the *like*-trend gives a boost to *go*, which experiences a steep rise with the youngest age group. In this group, *like* loses 2% but stays high. These high frequencies of *like* and *go* combined exert pressure on the unframed variant which sags by 14 percentage points to 11% in the youngest age group. In conclusion, with the intrusion of 2 new members into the semantic-pragmatic field hitherto held

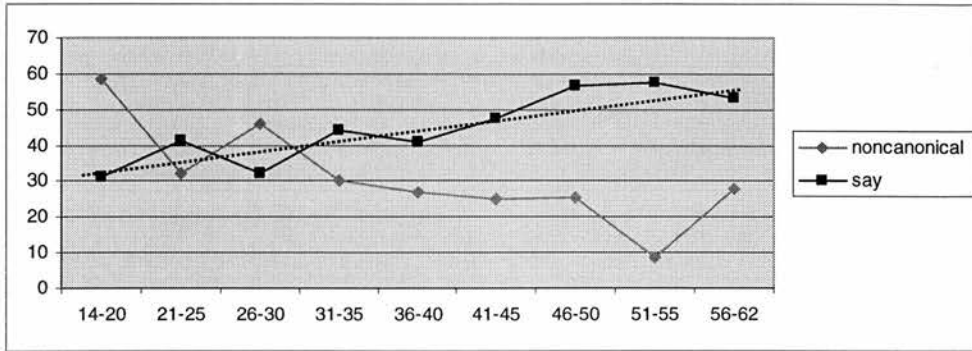
by the strategy unframed alone, the old non-lexical variant has lost ground. This, as we know, is possible due to a reallocation of the feature [-canonical] amongst the speakers in the youngest age-groups.

The question is still open as to whether this development has implications for *say* or whether the distributive reallocation is only occurring amongst the [-canonical] variants. Overall, as Table 6.13 shows, the proportion of *say*-use is significantly higher amongst the older speakers (36% versus 47%,  $\chi^2(1): 16.146, p = .000$ ). But this differential could be merely due to the much-cited fact that younger speakers use more vernacular variants (*like*, *go* and unframed) whereas older speakers generally use more standard variants. In order to see whether the frequency of *say* is at all affected by the trend towards *like* and *go*, we will have to monitor *say*-ratios across age. Is there a noticeable dip in *say*-frequencies at the point in time when the new quotative *like* gains full speed? Or is there one when *go* starts being used with incremental frequencies by the youngest age group?

I will now contrast the patterning of non-canonical quotatives as a sum of *go*, *like* and unframed combined with the distribution of *say* across numerical age. I have chosen a binary depiction [+/- canonical] because it enables us to see whether newcomer variants interact mainly within their own semantic field or whether they have an influence on *say*.

If *say*-frequencies across time prove to be affected, it is an easy step to find out which of the vernacular variants show fluctuations in patterning in this particular age bracket. If on the other hand *say* stays relatively unaffected, we will have to assume that the reorganisation within the quotative pool has taken place within the vernacular sub-set.

**Figure 6.12: Contrastive patterning of non-canonical and canonical (*say*) quotatives across age**



We can see that when we contrast [+/- canonical variants], the resultant pattern is an x-shaped curve. As expected, vernacular variants (amongst which are the innovative ones) are increasing with falling age whereas older speakers increasingly prefer canonical variants. As older speakers tend to have a more conservative habitus (Guy et al. 1986, Holmes 1992, Labov 1994, Maling 2002, Romaine 1984, Zilles 2003), the graph shows falling *say*-frequencies with decreasing age. Note also the jump in vernacular quotatives in the oldest age-group due to age-grading (cf. Holmes 1992, Downes 1984). The incremental frequencies can be put down to higher unframed rates (consider Figure 6.11). But higher non-canonical rates in the oldest age bracket do not turn out to be at the expense of *say*, which is only slightly reduced for the age group 56-62 (by 4%). Other variants, which have not been considered in this discussion, must come into the picture here.

We notice an intimate interplay between +/- canonical especially in the younger age brackets where *say*-frequencies are reacting to higher production of non-canonical variants. Most importantly for our purposes, the *say* and the non-canonical lines display a mirror patterning with *say* dipping at ages 26-30 and peaking during ages 21-25. Does that have anything to do with the introduction of the new quotatives *like* and *go*? This brings us back to the question of whether there is any indication that incoming new variants push out canonical *say*. In order to answer this, we will have to see whether the point in time at which *like* became vigorous in the quotative system (from age 26 onwards, cf. Figure 6.4) or the peak in *go*-frequency in the 36-41 year old and the youngest age groups (cf. Figure 6.6) corresponds to a decline in *say*-ratios.

Let us first discuss ages 26-30. Figure 6.4 shows that *like* is rising to high frequencies (while *go*-frequencies are low in this age bracket). But the peak in the non-canonical curve (and corresponding dip in *say*-ratios) also matches up exactly with high unframed frequencies. Hence, the low in canonical frequencies is not due to trendy *like* alone. On the other hand, the high in *say*-frequencies occurs exactly in the age bracket when *like*-frequencies soar and *go*-frequencies are ascending. Pulling these pieces of information together, we have to contend that the addition of two variants to the quotative repertoire of the younger age groups does not have any catastrophic effect on the *say*-curve. *Like* and *go*, while certainly interacting with *say*, have not intruded much into *say*'s territory.<sup>22</sup>

In the field of grammaticalisation. Bybee, Perkins and Pagliuca (1994) have shown that sometimes newly developing forms can restrict the linguistic contexts in which older forms persist. When we transfer this notion to variationist sociolinguistics, we find that many accounts of quotatives include some discussion about certain variants driving out others (Ferrara and Bell 1995) or intruding in some other variant's space (Cukor-Avila 2002) etc. Indeed, most of the sociolinguistic literature is concerned with the ensuing consequences for the traditional quotative, *say*. Or else, the literature wonders, will *like* drive out quotative *go*? (Cukor-Avila 2002, Bakht-Rofheart 2002, Tagliamonte and Hudson 1999). In Chapter 3 of this thesis, I have given linguistic evidence that an expulsion scenario is rather unlikely given the functional equilibrium of all members of the (say) variable. Now we see that on the social plane as well there is interaction but not expulsion. Earlier in this chapter, I showed that far from driving out quotative *go*, *like* seems to have given it fuel amongst the youngest generation. Rather, there is evidence that it is the third non-canonical variant, unframed quotes, which covered some of the ground now held by *like* and *go*. Indeed, looking at Figure 6.11, we noticed a dip in the unframed-curve during the age range where *go* peaks and a fall just at the time when the new vigorous change in progress, *like*, reached an all-time high of 20%. I have not yet

---

<sup>22</sup> When we compare the deviations of the *say*-curve with a perfect falling line (the broken line), we notice that the *say*-patterning deviates only slightly from it. Hence, even if *like* and *go*'s peaks alone were responsible for the dips in the *say*-curve, their influence is minor and does not exceed 8% (in 26-30).



come across any study that considered this possibility of a restructuring amongst the non-canonical set.

Also, we found that the advent of the new quotatives has not had any major impact on the traditional quotative *say*. Fully in tune with the much cited fact that younger speakers have a less conservative system, *like* and *go* have taken away some of *say*'s ground in the youngest age frame. But it appears that *go* and *like* are the major competitors for unframed quotes. As younger speakers' choice of variants is wider than the older speakers', there are potentially three strategies that can divide their [-canonical] quotative category, which was occupied by unframed quotes alone before the arrival of *go* and *like*.<sup>23</sup> With the introduction of two new members into this semantic space, the balance amongst the quotative options which inhabit this space is in the process of restructuring. We do not expect the variant unframed to be pushed out of the system, though (as alluded to in Tagliamonte and Hudson 1995). As the only non-lexical option, it bears stylistic options that cannot be fulfilled by either *go* or *like*, amongst which are the non-allocation to a speaker role, the non-determinateness with respect to time, and the closer proximity of message and quotation (Yule and Mathis 1992).

In conclusion, it is only when we look at the quotative system as a whole that we see the intimate interplay of the competing variants with one another. In the US system, we have seen that far from ousting the unframed or *say*-variant, *like* and *go* add options within the vernacular category. The next few sections will show whether this finding can be extended to British English.

---

<sup>23</sup> Note that this is obviously not the case for any individual speaker.

## 6.5 British English

The following paragraphs investigate the (say) variable in British English. The data for this investigation was collected in 1994 in Derby and in 1994/5 in Newcastle. It will be shown that the social constraints operating in the quotative system that we have found in the US are not the same in British English. Hence, it makes sense to treat the two varieties as separate and coherent systems in their own right.

I will now discuss the possible ways in which the linguistic feature *like* could have spread from the USA into the habitus of British English speakers. For UK residents, US English is ever present, directly through interpersonal contact on British soil or abroad, facilitated by intercontinental travel, as well as indirectly, through the written and spoken media (interviews with stars, soap operas, music channels such as MTV etc.). We know that many US-stars are heavy *like*-users, as is exemplified by two stereotypical examples in the excerpts below:

So how long have Britney and Justin been an item? They first met years ago, when both were cast members of the Mickey Mouse Club. "we were like 12 or 13 and it was like 'Hi do you like me?'" says Spears, smashing the existing record for the number of likes in a sentence of 15 words or less. (The Guardian, June 15, 2001)

You know, he actually surprised me when we first started getting to know each other," she remembered. "He would be singing the hip-hop songs. He loves Ja Rule, he loves Nelly, he loves Snoop. He would know all the lyrics to these songs. He'd just start singing, and I would be like, 'What is that?' And he's like, 'That's Snoop from back in the day.' I'd be like, 'What?' (Jennifer Lopez on Ben Affleck on MTV news, [http://www.mtv.com/news/articles/1458399/20021029/lopez\\_jennifer.jhtml?headlines=true](http://www.mtv.com/news/articles/1458399/20021029/lopez_jennifer.jhtml?headlines=true), 29. October, 2002)

The following passage from a teen weblog, reporting on a series of "Friends", shows that characters in the show are also perceived as being high *like*-users:

rachel was in labour for 47 hours [lol] and rachel and ross kissed. and phoebe was like "why the hell aren't you two together?" and ross (...) was like "noo, i explained it already.." and phoebs was like "right right. so who does the baby look like?" and ross was like "are you kidding? emma is gorgeous, just like rachel..." (source: [http://qtc.blogspot.com/2002\\_05\\_12\\_qtc\\_archive.html](http://qtc.blogspot.com/2002_05_12_qtc_archive.html))

Trudgill (1983), while giving motivations as to why British celebrities should take on US features, shows that the US variety is indeed influential on Britain's rock stars (with respect to (t)-flapping, postvocalic (r) and other linguistic features). The following snippets show that the British media-darlings have picked up on the *like*-trend as well. The utterances are chosen from two women who we can safely assume to be major role models for British youth when it comes to fashion and music, Ms Dynamite and Victoria Beckham.

He listened to Boo and there are about four bars on there of me singing and he was like "have you thought about singing?" and I was like "not a whole song"... He told me to try writing a song, so I did. (Ms Dynamite in Interview on BBC1, [http://www.bbc.co.uk/radio1/urbannelson/msdynamite\\_1\\_20020429.shtml](http://www.bbc.co.uk/radio1/urbannelson/msdynamite_1_20020429.shtml))

I think it is great that the fans want to see a little picture of Brooklyn so I thought, what better way to do it than on my official site where I can control what goes out. David has just seen the site for the first time and he was like, "Wow that's amazing". It's very cool. (Victoria Beckham on <http://www.spicenews.com/iviews/view91.html>)

It would seem reasonable to expect that, along with clothing style (V. Beckham: "I think that people are really interested in how I do my hair and the clothes I wear and how my make up artist Karin does my make-up" (<http://www.spicenews.com/iviewsiview91.html>)), and aspects of behaviour, certain features of the stars' linguistic behaviour will be mimicked by the crowd of fans.<sup>24</sup>

Within the field of linguistics, several studies recently have asked whether and, if yes, how much effect television has on the linguistic behaviour of its consumers (Davies 2002, Meyerhoff and Niedzielski 2002, Stuart-Smith 2001, Tagliamonte and Hudson 1999, Walters 2002). Hudson (1996) rejects the specific causal role of TV in systemic language change, (cf. also Milroy and Milroy 1985, Trudgill 1986, 1988).<sup>25</sup> Work by Androutsopoulos (2001), Foulkes and Docherty (2000), Strang (1970) and Williams and Kerswill (1999) has been attributing a more important role to the media as a probable source for change in language attitudes or spread of innovations.

---

<sup>24</sup> In an interview for Remix, Ms. Dynamite shows that she takes being a role model seriously "I'm really aware of the fact that, as a young black woman in Britain, there are not really any other black women that are given the position that I am, where I can be conscious and speak about whatever's on my mind". ([http://remixmag.com/ar/remix\\_ms\\_dynamite/](http://remixmag.com/ar/remix_ms_dynamite/)).

<sup>25</sup> See Parameswaran's critique of the role of the mass media which, she says, is one of the "key sites" in which ideologies of globalisation and post-colonial attitudes are "repeatedly manufactured and distributed" (2002:312). This supports the causal role of the media in transmission of attitudes.

Stuart-Smith (2001) reports that while studies have shown the influence of TV on non-linguistic social behaviour (Strasburger 1995: 13, APA (1993) in McQuail (2000: 436) the evidence for the effect of the media on linguistic behaviour is inconclusive to say the least. Stuart-Smith issues a call for a more differentiated approach to investigating the influence of TV on systemic language change.

Most authors acknowledge that TV creates knowledge, which can contribute to a change in attitudes. Furthermore, it can raise awareness about language varieties or linguistic variants. Why should we assume that these new states of current knowledge or disposition could not lead to new patterns of behaviour, in linguistics as much as elsewhere? Indeed, reports such as the one by the British Heart Foundation (2000: 4) that “watching TV is the most popular sedentary activity for children of all ages, with nearly three-quarters of 11-16 year-olds watching for 2 hours/day, and over a quarter watching for more than 4 hours/day” question the claim that face-to-face interaction is a necessary condition for the transmission of linguistic and other variables. Notice that Trudgill (1986:40) argues against transmission via the media claiming that if the media was so strong a factor, then the whole of Britain would be influenced by a particular innovation simultaneously. But this is exactly the case with *like*, which is attested in many varieties in the UK, in London (Andersen 1996), in York (Tagliamonte and Hudson 1999), in Derby and Newcastle (this thesis) and in the Shetland Isles (Scobbie, p.c.).

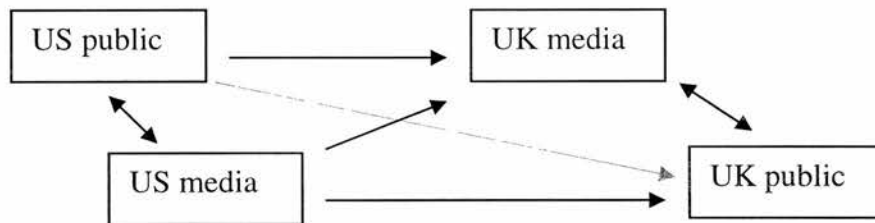
If adolescents spend more time watching TV than any other leisure activity (Strasburger 1995: 2, cf. also Giddens 1997:368), it seems to be reasonable to assume that this activity creates virtual communities such as fans of series X and Y-buffs. This is in vein with Noble’s (1975 in McQuail 2000: 406) claim that “these [TV] characters serve as something akin to a screen community with whom the viewer regularly talks and interacts ... this serves for many as an extended kin grouping”.

Work by Cutler (1996, 2002) shows that young hip-hop fans display their affiliation with the hip hop culture through various stylistic choices such as clothing, hair style, forms of walking, gesturing, etc. By copying such traits from the stars, solidarity within the community of fans is created. Cutler’s work identifies, amongst other traits, also linguistic features (such as (r)-deletion, (t/θ) and (d/ð) alternation

etc.) which are shared by the community of hip-hoppers. If it is the case that the stars are one of the important linguistic models for rap-fans and that the fans use certain linguistic variants as a means to affiliate themselves with the world of rap, there is reason to assume that a variant such as *like*, which started being used and consequently became associated with a prestigious social group, the world of pop and soap stars, could be transmitted from the role models to the group of fans.

Hence, as Figure 6.13 shows, we have a two (or three) way adoption scenario for the British public. The influence on the British public could be exerted (i) by US movies, television series (*Friends*, *Buffy the Vampire Slayer*, *Sex in the City*), as well as American showbiz and glam programmes which are readily and regularly available in the UK and (ii) directly via the UK media, and probably to a varying extent by both.<sup>26</sup>

**Figure 6.13: A possible scenario for contact-induced *like*-adoption**



On both sides of the Atlantic, young speakers pick up features from film and music stars and use them to perform “acts of identity”<sup>27</sup> (Le Page and Tabouret-Keller 1985) signalling their adherence to groups of fans. *Like* was first picked up as such a marker of in-group status by early adopters and was subsequently spread to a larger adolescent public.

Note also the two-sided arrows between public and media. While the above discussion has mainly been concerned with the influence of the media on the public,

<sup>26</sup> Notice that Figure 6.13 contains a weaker coloured link between the US public and the UK public. I assume direct face-to-face influence to be quantitatively relatively small because the main introducers of *like* in the UK, speakers in their late teens and early twenties, do not usually have extensive intercontinental contact. An investigation into the qualitative strength of such transatlantic interpersonal ties between US and UK youth and the impact of such contacts is still to be made.

<sup>27</sup> Le Page and Tabouret-Keller give two explanations for the adoptions of new linguistic forms: 1) the attribution of positively evaluated traits with in-group status 2) the association of a linguistic form with that status. Both attitudes seem to prevail when *like* is interpreted as *Friends*- and *Sex in the City*-speak and is taken on board by the target group of these programmes, adolescents and young adults.

it seems more than evident that the general public has an influence on the media in the sense that they are “in” it. Real people are interviewed or depicted in some way in television programmes. Also, fictional programmes try to represent “real” people. Hence, we find a reciprocal influence between media and public. In this vein, Foulkes and Docherty (1999) note the trend towards vernacular regional features in the popular media: (i) pop song lyrics (Oasis, Blur, All Saints etc.) are often replete with regional variants (ii) stars speak regional dialects in interviews and (iii) television soaps and dramas are frequently set in lower class urban environments (East Enders, Coronation Street, Teachers). This might have the effect of making vernacular features more commonplace in the population as a whole, which in turn again heightens the probability that they will be portrayed in the media and so on.

We know that by 1994 *like* had found its way in the use of London teenagers (Andersen 1996). I would claim that contact with a *like*-variant has had the following linguistic repercussions on British English speakers’ habitus: In a situation of contact with US English, in which the variant *like* is already present, the feature [- canonical] could have been reallocated from already existing variants to a new incoming one. Hitherto primarily carried by the unframed quotes and to a lesser extent by *go*, [- canonical] is now also carried by the newcomer to the quotative frame, *like*. We expect this to have happened considerably faster with British pop stars and other early innovators. The majority of *like* users followed in due course. Hence, in line with Britain’s (1997, 2002) findings about the reallocation of linguistic variables in a situation of dialect contact, I suggest that it is not unreasonable to claim that the (say) variable in Britain could be restructured to accommodate the newcomer. The reallocation hypothesis can help to explain the synchronic facts of *like*’s global diffusion: the feature [-canonical], carried hitherto by older variants only, is carried over to the newly adopted *like*. As a consequence, in Britain as well, [-canonical] quotatives now have a competitor variant, *like*. This section investigates how the quotative pool reacts to *like*’s intrusion. The following pages will give evidence of how redistribution in the quotative pool of British English emerges on the social plane, post *like*.

### 6.5.1 Gender

Table 6.15 (below) shows that at the point in time of *like*'s introduction into the British English quotative cohort, there is no significant gender bias. In 1994/5 we find 5% *like*-use for male and 4% for female speakers.

**Table 6.15: Distribution of quotatives per gender**

	like		go		say		think		unframed		tell		other		SUM
	%	N	%	N	%	N	%	N	%	N	%	N	%	N	
m.	5	42	<b>10</b>	86	<b>44</b>	383	9	75	27	238	1	11	4	37	872
f.	4	51	<b>15</b>	177	<b>49</b>	583	8	100	19	224	1	12	4	43	1190

Table 6.15 demonstrates that *like* in the Derby and Newcastle corpus is a trans-gender phenomenon. This result is especially interesting given the overwhelming preponderance with which the literature on quotatives ascribes *like* either to the female or the male speaker. With respect to British varieties, Macaulay (2001) and Tagliamonte and Hudson (1999) report that *like* achieves higher frequencies with female speakers.

Quotative *go*, on the other hand, has been around long enough to become stratified by gender. In my corpus, *go* is predominantly used by female speakers (15% versus 10% for men). This result, which is significant at the .01 level ( $\chi^2(1):11.205$ ), is contrary to some much-cited stereotypes. But note that these stereotypes are reported from the US. I have not found any comment in the literature on social stereotypes with respect to *go* in British English. In Buchstaller (2003) I have shown that results from a social attitudes questionnaire and from a matched guise test reveal that middle class British informants do not seem to have strong stereotypes towards *go*.

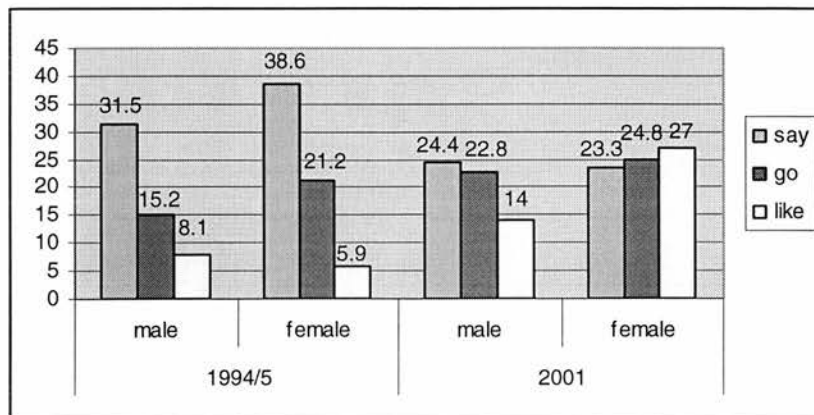
My finding converges with Macaulay's Glasgow data (17% male versus 23% female), as well as with Winter's (2002) results from Australia, where females favour *go*, albeit working class females. Hence, quotative *go*, a variant which had hitherto been associated with the feature [+male], is produced with higher frequencies by female speakers in Derby, Newcastle, Australia and Glasgow and has no gender effect whatsoever in my US corpus (cf. Table 6.9).

Numerous studies have argued that women use fewer vernacular or non-standard variants because “females show greater sensitivity to socially evaluative linguistic features than males” (Labov 1969/1972: 76, see also Cameron and Coates 1988, Labov 1972a, 1990, 2001, Mougeon et al. 1988, Romaine 1978, 1984, Trudgill 1983, Wolfram and Fasold 1974). Data from 3 British varieties (Derby, Newcastle, and Glasgow) and from Australia show that such generalizations do not hold with respect to the non-categorical members of the quotative system. Indeed, with respect to the (say) variable, women consistently use more non-canonical variants than men (see also Chambers 1995, Labov 1990, 2001:71). But note that the quotative system in British English also shows that women favour more standard variants (cf. the significant results for the use of say (men 44% and women 49%,  $p = .024$ ). It seems that the patterning of the variable cannot be explained by simply subsuming the variants under the two headings ‘standard’ and ‘non-standard’. A system with more than one standard and non-standard variants calls for a more differentiated subcategorisation of the variable. When we implement a two-way differentiation of the system as canonical versus non-canonical and old versus new, we can state that women use more of the old canonical variant *say*. But women also pick up incoming non-canonical variants with higher probabilities. Men prefer the older [-canonical] variant, unframed.

To give some time-depth to this investigation, Figure 6.14, below, depicts the comparative histogram of the 2002 York data with my corpus split up by gender. As the categories unframed and *think* were a matter of inconsistency, I will only discuss the variants *like*, *go* and *say* in the two corpora.



**Figure 6.14: Real-time comparison between *say*, *like*, *go* in 1994/5 and 2001 (data extrapolated from Johnson 2001) for the youngest age-group**



If we compare the Derby and Newcastle corpus from 1994/5 with the York 2001 corpus, it becomes evident that while *say* is reducing in frequency, it is losing its gender bias as well. In the 1994/5 data, we notice a bias towards female speakers for the whole population (with 49% versus 44%, cf. Table 6.15) as well as for adolescent speakers (with 32% versus 39%, cf. Figure 6.14), both of which are significant at the .05 level. This bias levels off by 2001, where Johnson reports no statistical significance.

Gender also plays an important role in the development of both non-canonical quotatives. While in 1994/5 the difference between young male and female speakers for *like* is only 2.2% ( $p=.138$ ), by 2001, this has grown to a significant 13%. *Like* develops a gender effect. I interpret this result as an indication that *like*, while only an incipient change in 1994/5 with no gender differentiation, has picked up momentum in 2001. This finding is concurrent with Labov's (2001:186) claim that "incipient changes ... do not have strong enough correlations with age and social factors to be informative. A clear view of the social location of a change in progress appears only when its age coefficient becomes significant, reflecting the fact that it is approaching the middle section of its s-shaped curve" (see also Eckert 1989).

Moreover, Chambers (1995) points out that it is more geographical and social mobility and access to information which determine our use of innovative features. I will assume that during *like*'s initial stage of entry into the British quotative system, potentially transmitted via TV, it is available to speakers of both genders (but cf. Nordberg and Sundgren 1998). We see from the data that in Great Britain younger speakers of both genders acquire an incoming variant in equal numbers. It is initially

picked up at equivalent rates. In subsequent years, this innovative feature can be further exploited by specific subgroups and can become a stereotype or even a marker of this group. This seems to have been the case with respect to *like*-use in British English: By 2001, *like* is used more by young female speakers.

Note that the Tagliamonte and Hudson (1999) data from 1995-6 in York already shows a significant gender skew for the distribution of *like* with a lead for the females (VARBRUL weights .67 versus .36, they unfortunately do not give any raw numbers). Whether *like* has taken on the trait [+female] within the short time between 1994/5-1995/6 (collection of my data and collection of the Tagliamonte and Hudson York corpus) or whether these results are a case of regionally conditioned variability remains a question. It might even be the case that *like* (for a reason yet to be determined) was introduced earlier into the York variety and thus had more time to develop a gender skew. But from my Derby and Newcastle data it seems reasonable to assume that in its initial stages of entry, the social variant *like* does not show any gender effect (which supports Labov 2001 as well as Chambers 1995). Once *like*'s introduction into the quotative system is vigorous by 1996 in York, it is carried by the typical innovators, young women,

These findings show that the neutralisation of the gender bias, which Ferrara and Bell observed for their US data over the 4-year time period of their investigation, cannot be confirmed in the BrE corpus. On the contrary, my findings fully support the claim made by Tagliamonte and Hudson (1999:167) that “the more diffused *be like* is, the more likely it is to differentiate male and female speech”. But note that BrE *like*, even in 2001, is still a young phenomenon with a life-span of approximately only 6 years, as compared to 12 years for the 1994 corpus in which Ferrara and Bell found a levelling of the gender-bias. Initially, in 1990, 8 years after *like*'s official “discovery” by Butters in the US, Ferrara and Bell, too, found a significant gender skew in their US data. Whether the gender-bias in BrE as attested in Figure 6.14 will level off to follow the American trajectory or whether it will remain or even become stronger as predicted by Tagliamonte and Hudson (1999) is a matter for future investigation.

*Go*, on the contrary, seems to lose its gender constraint. In 1994/5, it has an overall gender bias (cf. Table 6.15) and is clearly favoured by the female speakers (15% versus 21%,  $p = .006$ ). A comparative examination with newer corpora shows that the percentage gap of 6% in 1994/5 shrivels to a 2 % difference in Johnson's data (2001). Already the Tagliamonte and Hudson (1999) data from 1996 shows no significance for gender (.46 versus .53 VARBRUL weights, not significant). Hence, from Derby / Newcastle to the 2001 York corpus, *go* loses its gender effect and increases in frequency.

The fact that *go* does not have any gender associations in the US data whereas it is clearly associated with the feature [+female] in British English in 1994/5 but later loses this gender effect proves again that social meaning is indeed arbitrarily assigned and can be associated differently in different varieties and at different times.

At this point, it needs to be stressed that the figures given in Table 6.15 are only the average total numbers from the Derby and Newcastle corpus. If we split the figures up by variety (cf. Table 6.16), we can see that within Britain, different localities show different tendencies.

**Table 6.16: Distribution of *like* per gender and variety (in % frequency)**

	Derby	Newcastle
male	4	6
female	1	8
N	22	71

While both varieties participate in the new *like*-trend, Table 6.16 shows that in Newcastle, where *like* is overall more frequent, men as well as women are the driving forces behind the introduction of this new quotative. Here, women are slightly at the forefront with 8% versus 6%. Those numbers look as if they could be an earlier stage of the Johnson (2001) and Tagliamonte and Hudson (1999) statistics. We can imagine a scenario in which the 2% difference in Newcastle in 1994/5 has grown to 13% by 2001 (Johnson 2001).

But Derby speakers, in spite of being the later corpus and having had more time to adopt the variant *like*, produce far fewer tokens of it than their Newcastle

counterparts (N=22 versus 71).<sup>28</sup> Derby men slightly lag behind Newcastle men (4% versus 6%). But the Derby women, who hardly use *like* at all, mainly account for the big regional difference. Indeed, while the Newcastle grid looks indicative of a prototypical change in progress with women as the innovators and men lagging behind (Aikio 1992, Feagin 1980, Haeri 1994, L. Milroy 1999), the Derby women do not participate in this innovative feature. Their 1% *like*-usage runs clearly counter to the many statements in the literature which point to female speakers as the principal carriers of the recent changes in the quotative system (Ferrara and Bell 1995, Igoe, Lamb, Gilman and Kim 1999, Johnson 2001). Rather, the Derby women in my corpus do not participate in a supra-local innovation. While it is still a conjecture at this point, there is a possibility that the avoidance phenomenon that Bakht-Rofheart (2002) has detected for her Long Island teenagers could provide an explanation for the behaviour of the young Derby women. In not using the innovation, they would then distinguish themselves simultaneously from (i) their male counterparts, who adopt *like*, (ii) as well as from other women (in their country as well as abroad) who have been pinpointed as being the principal *like*-users. This finding is fully coherent with Eckert's (1989:253) claim that "gender does not have a uniform effect on linguistic behaviour for the community as a whole".

Now let us see whether Derby women generally do not participate in the use of the non-canonical quotatives. The next table shows the distribution of quotative *go* split up by gender for both varieties, Derby and Newcastle:

**Table 6.17: Distribution of *go* with respect to gender per variety (in % frequency)**

	Derby	Newcastle
male	13	7
female	14	17
N	135	128

<sup>28</sup> How much difference, if any, in the diffusion of such rapidly spreading features we would expect from one year's difference in data collection is a matter of conjecture. However, given the speed with which *like* appeared in most major varieties of English and the vigour with which it spread both across the sociological spectrum and within the linguistic system, we cannot easily dismiss one year as insignificant.

In Newcastle, we witness the same effect as in Table 6.16: female speakers favour the non-canonical quotative (by 17% versus 7%,  $\chi^2(1):20.618$ ,  $p<.01$ ). Hence, again, Newcastle speakers display a gender gap with women predominantly using the non-canonical form. But in Derby, there is a roughly even frequency distribution between men and women (13% versus 14%, not significant). Derby women are not lagging behind as with *like*, but they still do not produce (significantly) more of the vernacular variant. Thus, while Table 6.17 shows that overall women use *go* more frequently than men in British English, we now see again that the distribution by gender is also conditioned by locality: while Newcastle women are in the lead for both variants, Derby shows no significant gender-effect because conservative female ratios and high male production cause their output to meet in the statistical middle.

Let us now reconsider Foulkes and Docherty's (1999) claim that female speakers use more supra-local forms. Indeed, the above presented data shows that *go* is favoured by women of all ages, in Derby (albeit with a slim margin) as well as in Newcastle. *Go* is also a supra-local form. It is noted in London English (Andersen 1997), in Glasgow English (Macaulay 2001), in Edinburgh (Miller and Weinert 1995) and multiple other varieties of English (US English, Tannen 1986, Tagliamonte and Hudson 1999, Australian English, Winter 2002). Thus, in line with Foulkes and Docherty, women in my data from 1994/5 favour this global variant, across all classes (see below) and in both varieties.

But we also find that quotative *like*, another supra-local form, does not yet have any gender effect in 1994/5. But an investigation in real time demonstrates that it rapidly acquires a significant gender-differential with women in the lead. A split up by variety showed that Newcastle women are in tune with Foulkes and Docherty's generalisation. The fact that Derby women do not participate in the innovation in 1994/5 might be due to an avoidance behaviour (cf. above).

But note that the category supra-local vernacular forms also includes bald unframed quotations. Table 6.15 shows the great difference in frequency differential with which male speakers prefer this option (27% male versus 19% female,  $p=0.000$ ,  $\chi^2(1):20.371$ ). This result certainly runs counter to Foulkes and Docherty's claim. However, as the hypothesis was borne out with respect to supra-local vernacular variants which have a surface form, we could restate it as follows: amongst the (say)

variants, women favour supra-local variants that have a surface form. The null variant behaves differently.

### 6.5.2. Socio-economic Class

Most of the literature on the new quotatives discusses data from one social sphere only, the middle class. In the three large-scale studies on southern British English so far, Tagliamonte and Hudson (1999) only consider middle class speakers. Johnston (2001) does not mention socio-economic standing at all. Only Macaulay's (2001) investigation in Glasgow includes class as a social variable. Speakers are "drawn from two areas of the city, representing broadly urban working-class and suburban middle-class areas" (2001:7). Macaulay's figures show that in Glasgow Scots, *like* is a middle class feature (with an approximate 13% lead)<sup>29</sup>, whereas *go* is more a working class feature (with a lead of roughly 10%). I will now investigate the Derby and Newcastle corpus in the light of the social stratification within the quotative system. Table 6.18 gives an overall break-down of the patterning of all quotative strategies in the data according to class.

**Table 6.18: Distribution of quotative strategies according to class**

	like		go		say		think		unframed		tell		other		SUM
	%	N	%	N	%	N	%	N	%	N	%	N	%	N	
WC	4	50	<b>10</b>	115	48	549	7	76	<b>25</b>	292	1	14	5	52	1148
MC	5	43	<b>16</b>	149	46	418	<b>11</b>	99	<b>19</b>	170	1	9	3	28	916

Table 6.18 shows that *like* is used by speakers of both social categories to the same degree (with 4% and 5%, n.s.). This finding is sustained by a break-down according to locality (Derby working class 2%, middle class 2%; Newcastle working class 7%, middle class 8%). But notice that it is in contrast to Macaulay's Scottish speakers, where *like* is clearly a middle class feature. It also runs counter to the many reports of US English according to which *like* is carried by middle class speakers. Note again that this result is consistent with Labov's (2001) findings on incipient changes. Given *like*'s budding stage in the British English corpus, we would have expected social non-distinctiveness, which we indeed see here as well as in Table 6.15. Recall that in

<sup>29</sup> No exact numbers were available in Macaulay's article. I therefore had to estimate the frequencies from the tables he gives.

my US data, where *like* had been around for at least 7 years by the time of data collection, it patterns according to class. Whether *like* in Southern British English develops a class effect in parallel with a gender-bias – which more recent corpora have found it to acquire – and whether it will subsequently match the stratification found by Macaulay for Glasgow Scots is a matter for future investigation.

Quotative *go*, contrary to most findings in the literature and notwithstanding stereotypes on both sides of the Atlantic (cf. Buchstaller 2003) is a middle class feature in Derby and Newcastle (16 % versus 10 %,  $\chi^2(1): 17.835$ ,  $p < .001$ ). This result is again confirmed by its stability across varieties: In both the Derby and the Newcastle data we find a clear predilection for *go* amongst the middle class speakers (17% and 15% versus 10% and 11% respectively). Note that this result again runs counter to Macaulay's findings for Glaswegian speakers, where it is unanimously a working class feature.

The frequency patterning of *say* looks very much the same in both social strata (48% and 46%). The results for *think* seem to confirm the age-old stereotype of middle class thinkers; it is used less by the working class speakers (7%) and by more by the middle class speakers (10% of the time). However, the question needs to be raised of how much we can build on an average 4% difference, even if the difference does show statistical significance with  $p = .01$  and  $\chi^2(1): 10.98$  (note in this respect also that, in the US, we do not find any difference in *think*-frequencies).

Table 6.18 clearly shows that the non-lexical quotative, unframed quotes, is a strategy of working class speakers, who use it 25% of the time (as opposed to only 19% for middle class speakers,  $\chi^2(1): 13.876$ ,  $p < .001$ ). Hence, when wanting to quote non-canonically, while the middle class reverts to *go*, the working class chooses unframed quotes as the vernacular strategy.

The broader effects given in Table 6.18 will now be explored in more detail in a cross-categorisation of the category class with the orthogonal factor gender. Figures 6.15 and 6.16 depict a cross-correlation of class and gender for the quotatives *like* and *go*:

Figure 6.15: Distribution of *go* across class and gender

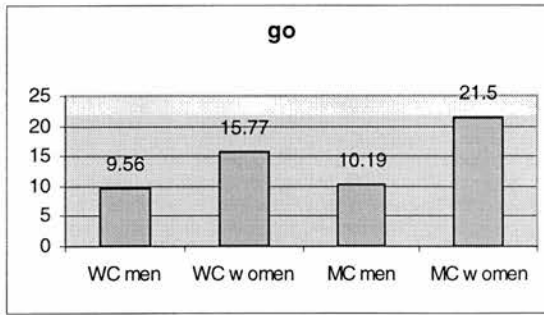
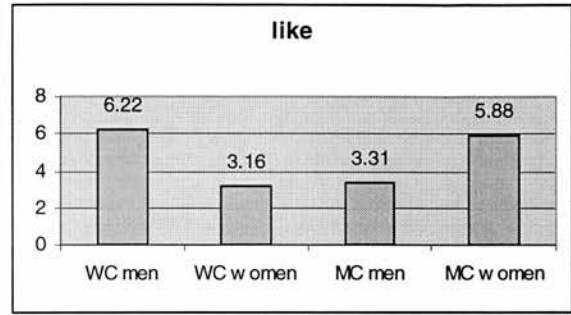


Figure 6.16: Distribution of *like* across class and gender



Let me first discuss quotative *go* (Figure 6.15). In the Derby/Newcastle corpus, at the intersection of class, the gender effect for *go* remains constant. Following earlier studies that have shown the patterned relationship between gender and linguistic variables (Cameron and Coates 1988:13, Chambers 1995:112, Gal 1980, Holmquist 1988), the recent sociolinguistic literature has pointed out that any generalising claims about the link between gender and use of language features are highly problematic (Holmes and Meyerhoff 2003, Labov 2001:265 et passim, Romaine 2003). Linguistic behaviour, just as any kind of social behaviour, needs to be situated in the local contexts and the history of the speech community in which the data is gathered (Eckert and McConnet-Ginet 2003, Fader forthcoming, Milroy, Milroy and Hartley 1994).

In the Derby and Newcastle data, we find that gender by no means has a uniform effect on the output of the new quotatives (cf. also the US data). Figure 6.15 shows that women use higher ratios of *go* in both classes. And the highest ratios of non-canonical *go* are produced by female middle class speakers. Note also that, contrary to the much-cited trend (L. Milroy 1980, cf. also Chambers 1995: 128 et passim) whereby less closely-knit networks for the middle classes lead to a weakening of effects on linguistic variables, in the Derby and Newcastle data the frequency difference for *go* is highly significant for the middle class ( $p < .01$ ) but not for the working class ( $p = .763$ ). Overall, the distribution of quotative *go* has shown that it does not conform to the general patterning outlined for stable situations by Labov (2001), whereby stigmatised features are mostly carried by lower class, younger and male speakers. Rather, in Derby and Newcastle, *go* is mainly used by younger speakers of both genders, by mainly middle class speakers and most of all



by middle class women.<sup>30</sup> An investigation into lay perceptions of this variant (Buchstaller 2003) has also shown that, contrary to many claims in the literature which were based on US informants, there are no strong attitudes attached to *go* in the UK.

But notice that it is *like*, the innovative variant, which is characteristic of young working class men's speech. This is in accordance with Chambers (1995) and Trudgill's (1980:97-8) claim that "changes away from the prestige norm...will have working class (..) men in the vanguard". Table 6.18 has given evidence that we do not have a stable class effect. Table 6.15 has revealed that there is no stable gender effect either. A deconstruction of the overall figures for *like* by the two orthogonal factors shows that its distribution is the result of its complementary distribution across gender and class. In *like*-use, working class women lag behind the corresponding males and middle class men lag behind their female counterparts (with 3.06% and 2.57% difference respectively). But while the working class difference still achieves significance with  $p = .017$  ( $\chi^2(1): 5.477$ ) (cf. L. Milroy 1980), the middle class slightly misses the cut-off point with ( $\chi^2(1): 2.792$ ,  $p = .084$ ). Note that this pattern, a change carried by working class men and middle class women, is exactly opposite to the pattern detected by Labov (2001) for changes that are well underway. Here, working class men and middle class women fall behind. Note also that the overall trend summarised by Foulkes and Docherty (1999) cannot account for *like*'s social reality: In the case of *like*, a situation of incipient language change, working class men (6.22%) and middle class women (5.88%) are the primary innovating forces.

In the case of the (say) variable, the use of the newest variant *like* and the old, canonical *say* are not stratified by class. *Like*'s non-significant patterning can be accounted for by the fact that it so new (Labov 2001). The social stability of *say*, as the most unmarked, neutral run-of-the-mill quote introductory item is not surprising either. Rather, the situation presents itself such that the quotatives that do show social stratification, *go* and unframed quotes, revolve around the stable axis of *say*. Middle

---

<sup>30</sup> A comparison with Macaulay's (2001) Glasgow data shows that also further north, we cannot speak of *go* as a working class phenomenon generally. In Glaswegian English, *go* is primarily produced by the middle class boys (43%) followed by the working class girls with still roughly 29%.

class speakers use more non-canonical *go* whereas working class speakers use more unframed quotes. This, again, points to the usefulness of the reallocation principle in explaining synchronic facts: while working class speakers use primarily unframed quotes as the [-canonical] quotative, middle class speakers revert to *go*. The underlying feature stays the same, the surface output looks different according to class.

### 6.5.3. Age

We know that by 1993, *like* has found its way into the speech of London teenagers because Andersen (1996) points it out in the COLT. Having looked at the rest of the British National Corpus (BNC), I can safely state that I find zero tokens of full quotative *like* in British English before 1993 (cf. Chapter 4 for the discussion of ambiguous cases). Hence, we are in the position to claim with relative certainty that my data represents the very point in time when *like* was first introduced into the British quotative system. Looking at the data in apparent time, we expect to find a frequency distribution that shows this change in progress.

My data from 1994/5 gives time depth to the York studies of 1996 and 2001. We are thus given the chance to investigate the real time development of the restructuring of the British quotative system as a whole. In contrast to trend studies such as Trudgill (1988), which add a more recent time slice to an existing apparent-time study, this study proceeds in the opposite direction and adds an earlier time slice, the 1994/5 data, to the 1996 and 2001 corpora. This methodology gives us the chance to catch a variant in its incipient stage and to monitor its further interaction with other members of the quotative pool at  $t_{1994-1996}=1$  and  $t_{1994-2001}=6$  years.

Quotative *go*, far from new, has been around as a variant for centuries. The section on US English has provided some interesting results with respect to *go*'s patterning across real and apparent time. I will now compare these findings with the British English data.

Consider first Table 6.19 below for an overall distribution of the (say) variable.

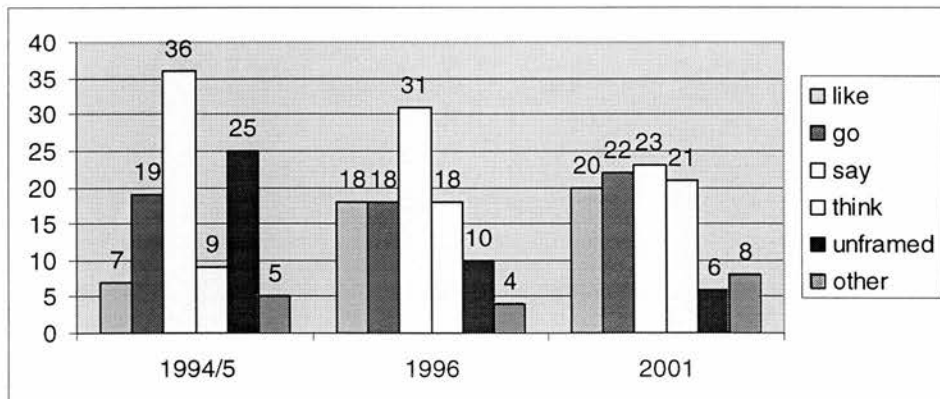
**Table 6.19: Distribution of quotatives per age-group**

	like		go		say		think		unframed		tell		other		SUM
	%	N	%	N	%	N	%	N	%	N	%	N	%	N	
young	7	90	19	250	36	476	9	120	25	326	1	11	4	54	1327
old	1	3	2	14	67	491	7	55	18	136	2	12	4	26	737

Looking at *like*'s patterning, two things already mentioned come to mind: the older speakers do not seem to use *like* at all (1% occurrence versus 7% for the younger speakers,  $\chi^2(1): 43.288, p < .001$ ). Notice also that *like* accounts for only a small number of the whole cohort (7% and 1%, overall 4.5%). This can be explained by its short age at the time of recording (cf. 8% in Australian English in 1997-99 as reported by Winter 2002). This result further confirms my earlier claim that the Derby and Newcastle corpora (1994/5) reflect the point in time at which *like* emerged in the inventory of BrE quotatives. But we know from comparison with more recent corpora that it has picked up speed within only a few years.

Figure 6.17 plots quotative production in different time slices. We get an approximate idea of *like*'s real-time spread in frequency from my 1994/5 Derby and Newcastle corpus to Tagliamonte and Hudson's (1999) and Johnson's (2001) data from York. In order to achieve general comparability, I have singled out the youngest age groups in the data.<sup>31</sup>

**Figure 6.17: Comparison of younger speakers' production of quotatives in 1994/5 (Derby and Newcastle), and 1996 and 2001 (York)**



<sup>31</sup> The Tagliamonte and Hudson (1999) corpus was taken from young university students only, the social group that has been pointed out as being the heaviest *like*-users. Johnson's (2001) corpus consisted of only 30% older speakers.

We notice a rapid numerical increase for *like*. From 7% in 1994/5 its frequency of occurrence rises to 18% in 1996 and furthermore to 20% in 2001. Quotative *go*, on the other hand, stays relatively stable across these 5 years of comparative evidence with a slight rise in frequencies during the last interval.

A comparison with Table 6.19 reveals that *say* is still overall the most frequent quotative in 1994/5. While it occurs 36% of the time with the youngest speakers (for similar results see Cameron 1998, Ferrara and Bell 1995, Johnson 2001, Tannen 1986), it makes up 67% of the older speakers' quotative repertoire. But Figure 6.17 shows that amongst the youngest speakers, *say*-frequencies shrink from 36% to 31% in 1996 and further to only 23% by 2001. By that time, *say*-frequency levels out with that of *like*, *go* and *think*. I will discuss this fall in *say*-use later.

Note also the differential output of unframed quotes in Figure 6.17. Unframed quotes account for 15% fewer tokens in the 1996 York corpus than in the 1994/5 Derby and Newcastle corpus<sup>32</sup> in spite of the one-year difference - and they show even lower frequencies in 2001 in York. This could be due to a rapid decline in the older, non-lexical non-canonical option as *like*-frequencies rise (via reallocation).

But a comparison with other corpora points to another possible conclusion: The unframed variant was noted independently by Mathis and Yule (1994), Cameron (1998), Tannen (1986) and Winter (2002) for their US English, Puerto Rican Spanish, Greek, and Australian English corpora. In all these data sets, freestanding quotes account for roughly comparable numbers, between 26% and 18%. Tagliamonte and Hudson (1999) also found 20% unframed quotes in their Canadian English corpus, cf. also Tagliamonte and D'Arcy to appear. It seems that a frequency-count of roughly 20% unframed quotes is cross-varietally the norm and that the two York corpora, for whatever reason, behave differently. The differential unframed output could be seen as an indication that the dialectal difference between York and other varieties is quite extensive with respect to unframed quotes, a function of different narrative styles (Tagliamonte and Hudson 1999). Alternatively, the numerical differences could be a function of different coding criteria in the two York studies, viz. what counts as an

---

<sup>32</sup> In order to ensure consistency of coding, I have compared the frequency distribution across varieties. My results are validated by the fact that there are equal ratios in the two regional varieties of my corpus, 20.7% in Newcastle and 18.6% in Derby, which puts them into the normal distribution that has been attested cross-linguistically.

unframed quote. Chapter 2 lays out in detail how I define unframed quote for this investigation.

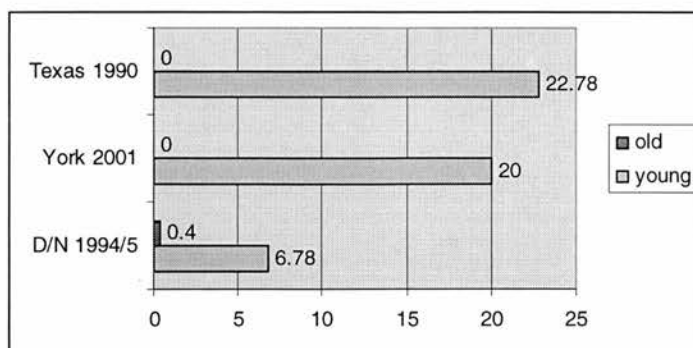
The different proportion of unframed quotes in my data versus the York corpora skews the investigation and makes comparison harder. But note that another factor which complicates comparative evaluation is the much higher frequency of *think* in the York corpora. As I have outlined in earlier sections (Chapter 2), *think*-framed quotes are often ambiguous with respect to their status as quotes. As an inclusion of large numbers of indirect *think*-framed quotes would significantly skew results, I have consequently excluded all tokens of ambiguous *think*. I can only speculate that an explanation for the different *think*-ratios in my corpus and the other two York corpora might be that the York corpora applied different criteria for the inclusion of this category in the data.

These numerical inconsistencies might threaten the claim of comparability of the two data-sets. But a closer look at the table shows that if we combine the frequencies of *think* and unframed quotes across the corpora, we arrive at approximately equivalent values: 31% in 1994/5, 28% in 1996 and 27% in 2001. Hence, the total remainder of variants is constant across corpora. This numerical stability ensures comparability with respect to *like*, *go* and *say* frequencies within the quotative system.

To sum up results for the variants *like*, *go* and *say*: *like*-frequencies monotonically increase with time, whereas *say*-frequencies decrease from 36% to 31% and 23%. *Go* stays remarkably stable through time with 19%, 18% and 22% respectively. Given the fact that *go* is an old variant, we would not expect otherwise, unless it is currently spreading. I will discuss the interaction amongst the non-canonical variants in more detail below.

The following table investigates further the *like*-spread across real time: Figure 6.18 plots the 1994/5 Derby and Newcastle data (D/N) against Johnson's 2001 York data and the data from Ferrara and Bell's (1995) 1990 Texas corpus.

**Figure 6.18: Distribution of *like* among young and old speakers in the Texas, York and Derby/Newcastle (D/N) corpus**



Within British English, *like*-frequencies amongst younger speakers spread by 13 percentage points from Derby/Newcastle to York by 2001.<sup>33</sup> They now parallel the distribution found by Ferrara and Bell (1995) in US English, where in 1990 an average of 22.78% of all quotatives were *like* (29% of the young females' quotatives and 15% of the young males'). Hence, *like*-ratios have achieved a roughly comparable frequency distribution in Texas, at least 8 years after *like* intruded into the US (say) variable, and in York 7 years after *like* was first noticed in British English. This result seems to indicate that *like* is spreading faster in Great Britain. This might be a function of the fact that British English speakers perceive already high ratios in the US variety and model their behaviour towards these frequencies. Note also that in all three corpora older speakers do not participate in this innovative feature. The fact that the Derby/Newcastle corpus shows a 0.4% frequency of *like* for the older speakers and the other corpora do not might stem from the fact that the borderline of what counts as "old" is quite low in Derby/Newcastle (over 38) in comparison to Texas (40+) and especially York (over 70). The speakers over 40 in my corpus produce 3 tokens out of 691 quotatives.

*Like*'s frequency distribution and patterning in real time point to a change in progress. But splitting up speakers into only two age-groups entails the loss of an important amount of information. What about the spread of *like*'s distribution across those age-groups? Where does *like* peak? Only if we ask these questions can we find

<sup>33</sup> Note that this result is paralleled by Macaulay's findings for Glasgow (2001:11). In a 1997 corpus he found 21.5% *like* use for younger speakers in contrast to only 0.2% for the older ones. The exact date when *like* was introduced into Scottish English is not certain. In any case, Miller and Weinert (1995) show that there is no attestation of quotative *like* before 1980.

out whether the initial split-up that was imposed upon the data actually makes sense in the investigation of this particular feature. A division of speakers by chronological age will provide insights into the patterning of this innovative feature across apparent time. We can now check for a monotonic relationship between age and *like*-use.

When looking at the scattergram in Figure 6.19, which depicts *like*-frequency per numerical age bracket, we have to consider the sampling method used for the British English data (as described in the data section in Chapter 2). Speaker selection was done so that the break-down by age results in a bi-polar speaker clustering with relatively few speakers between ages 28-45.

Figure 6.19: *Like*-ratio use per chronological age

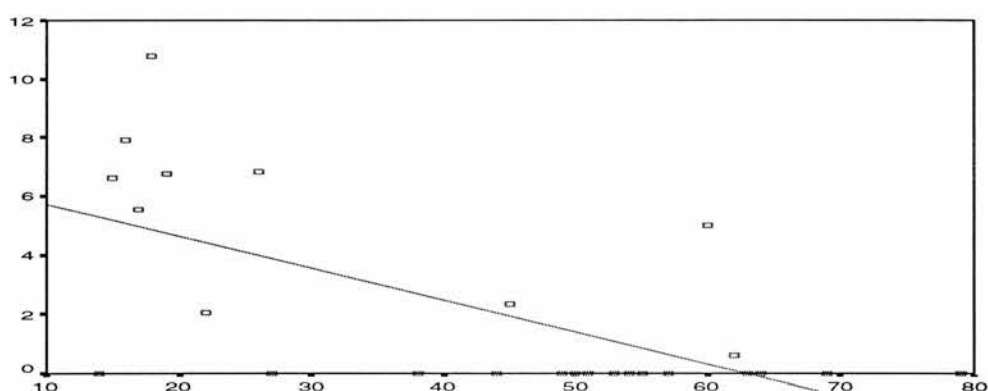
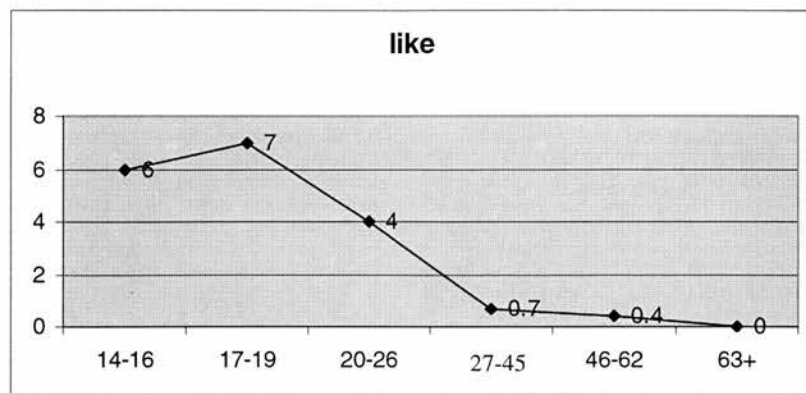


Figure 6.19 shows a clear clustering of high *like*-ratios in younger years (between age 18 and 28). From then onwards we notice an abrupt fall towards older speakers. The  $r^2$  ratio of 0.029 means that 3% of *like*'s variance can be explained by numerical years. However, the slope of the regression line does not fall as steeply as one might have expected given the many zero tokens from the age of 38 onwards ( $r = -0.108$ ). This is because of one surprisingly high *like*-ratio at age 57 (with 5% frequency). Otherwise, we notice zero-values from age 26 onwards (not considering a 2.2% at age 46 and a 0.6% frequency at age 62, which should not be disturbing given their values close to the regression line). When investigating the source of the 5% *like*-ratio at age 57, it becomes apparent that it stems from one speaker who produced only 20 tokens, one of which was *like*. Calculated with such a low denominator, this singleton token of *like* sticks out. Notice in contrast the consistency of zero tokens in neighbouring age-groups in spite of high token numbers overall ( $\Sigma(\text{all verbs})_{\text{ages}50+}$

550). When we exclude the outlier and group *like*-use in several age-brackets in which *like*-ratios average out, Figure 6.20 ensues:

**Figure 6.20: Distribution of *like* per chronological age by brackets.**



The line graph with a consistent downwards slope is typical for an apparent-time pattern representing a change in progress (see also McMahon 1994 and for a variety of sound changes consider Labov 1994, 2001). As we have seen for the US data, the innovative feature *like* peaks in adolescence (cf. Cedergren 1973, Eckert 2000, Labov 2001). From the highest point of the curve, the 17-19 age-group (with 7%), we notice a steep fall towards older speakers. *Like*-frequencies sag to an intermediate 4% among the 20-26 year olds and subsequently to frequencies below 1%. Overall, the graph shows a consistent demarcation by age. Note in this respect that Blyth et al. (1990) found that nobody over 39 used *like* in the US. Singler (2001) found what he calls the temporal isogloss separating speakers age 27-33 from speakers age 36-42. Younger speakers, in my data under the age of 27, use the incipient variant whereas everyone above this age-limit - with the exception of 3 speakers who each produce one *like*-token - stays unaffected by the new form. Hence, given the assured innovativeness of the feature *like*, we can safely affirm that the pattern across age gives evidence of a change in progress.

Because of its novelty, we might want to check for *like*'s distribution across varieties. If both Derby and Newcastle show the above bias, we have reason to claim that we are already looking at a general tendency in 1994/5. Consider Table 6.20.



**Table 6.20: Distribution of *like* with respect to age per variety (in %frequency)**

	Derby	Newcastle
young	4	10
old	0	1

Indeed, the *like*-trend in British English is confirmed by both Derby and Newcastle: *like*-use is restricted to the younger speakers. While Derby is lagging behind (consider Table 6.16), it is evident that all of the few tokens produced come from young speakers. In Newcastle, *like* is already more advanced. The quotative pool of younger speakers consists of almost 10% *like*. The older speakers' quotative system contains slightly less than one percent of the variant, which can be accounted for by borrowing from the younger generation. Cross-generational borrowing has been pointed out for phonological items by Labov (1994), and for music and fashion by Katz and Lazarsfeld (1955). In conclusion, *like*'s distributive evidence in both varieties points independently to a patterning according to age. It also shows that the advent of newcomer *like* in 1994/5 is a supra-local trend in British English.

One possible scenario is that British English speakers first react to the non-canonical innovation *like* with a higher production of variants which have the same underlying [-canonical] feature while still being reluctant to adopt the surface item. Only in a further step do they adopt the lexical variant and reallocate the feature [-canonical] within their own system from the old variants to the new variant *like*. We would expect such a development to have left its imprints on the distribution of unframed and *go*-framed quotes in apparent time. In the following paragraphs, I will discuss the patterning of unframed, *like*, *go*, and canonical *say* across age. Other questions that have to be answered are the following: does *go* stay stagnant while *like* is being introduced into the quotative pool? Do we have a balance in total number of non-canonical quoting devices or does newcomer *like* intrude in *say*'s traditional territory?

The following paragraphs discuss the distribution of *go*. Table 6.19 has shown that the younger generation uses *go* (19% of the time), whereas for older speakers it accounts for only 2% of the quotative system. This is the strongest age-effect in the distribution of the whole quotative system and unsurprisingly very

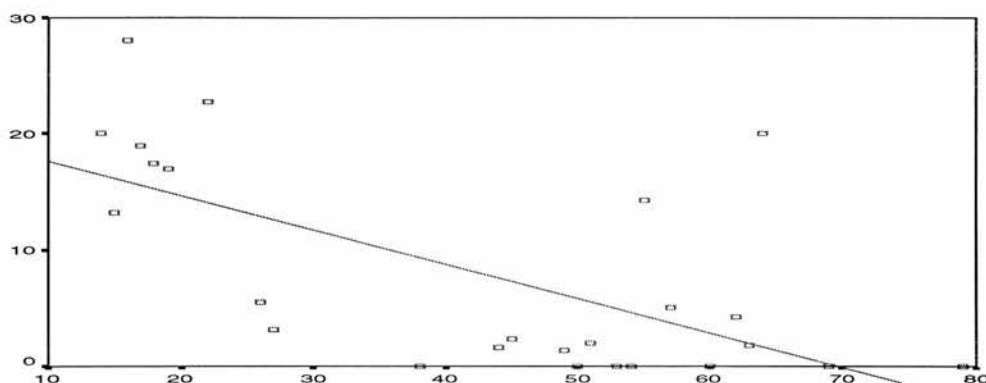
significant ( $\chi^2(1):120.383$ ). The age-gap persists across both varieties (young<sub>Derby</sub>= 21%, young<sub>Newcastle</sub> =17% versus old<sub>Derby</sub> = 1% and old<sub>Newcastle</sub> = 3%) as well as across gender (young men = 15%, young women = 21%, old men = 2%, old women= 2%).

A cross-correlation with class reveals that young middle class women lead the way (with 25%). This is in accordance with Ferrara and Bell's (1995) finding that "upwardly mobile young females – the very group reported in studies on gender to have a high level of awareness of socially correct norms of language use ..... college-aged females are advancing nonstandard forms". Young middle class men also have high *go*-ratios, albeit to a lesser degree (18%). The third highest frequencies are produced by the working class men (with 13%). Hence, *go* in Derby and Newcastle is carried by younger speakers of both genders, mainly middle class speakers and most of all by the young middle class women.

*Go*'s early attestation (OED, 1791) rules out the interpretation of its distribution as a change in progress (contrary claims about *go*'s newcomer status such as Macaulay 2001). The question arises as to why, if this quotative has been around for a while, do we see such a huge age effect? The following paragraphs will discuss the patterning of *go* in my data with respect to various possible scenarios. *Go* could show an age graded pattern – note that an apparent time investigation in the US did not yield conclusive evidence as frequencies showed no "tail". At the same time, we might witness the revitalization of an old, stable variant, which is variably spread across the population. Recall in this vein the US data where I argued that quotative *go* started to extend due to the impetus given by the advent of another non-canonical quotative, *like*. Alternatively, some researchers have claimed that *go* could be pushed out of the quotative system by newcomer *like* (Bakht-Rofheart 2002, Cukor-Avila 2002, Ferrara and Bell 1995) In this case, we would witness a dip in *go*-production from the intrusion of *like* into the (say) variable onwards. The following paragraphs discuss the data with respect to those competing hypotheses (age-graded, higher frequencies for the youngest speakers, *like* pushing out *go*) in real and apparent time.

The York corpora (cf. Figure 6.17) add later time-slices to the investigation. While overall, *go*-frequencies stay remarkably stable, we notice a slightly higher *go*-ratio both with the younger and with the older speakers in 2001 in York (19% (young) and 2% (old) in 1994/5 in Derby/Newcastle versus 22% (young) and 6% (old) in 2001 in York).<sup>34</sup> While, again, comparisons across regions have to be taken with a grain of salt, this 3% and 4% differential over 6 years could give grounds for the assumption that *go*, profiting from the drive of the new quotative *like*, is also gaining in frequency. A fine-differentiation of the category age can reveal *go*-use across time. The scatterplot in Figure 6.21 displays *go*'s patterning in apparent time.

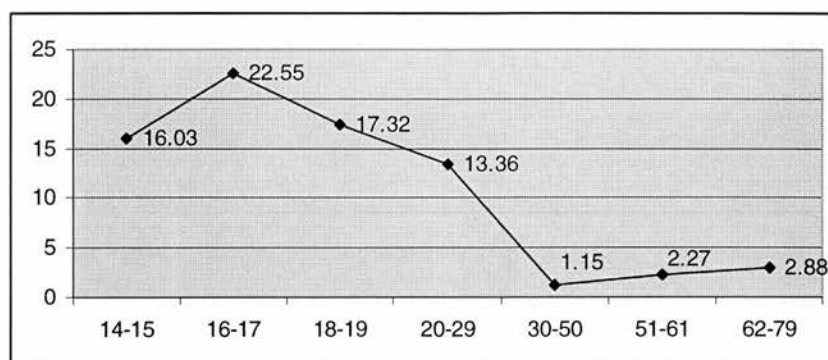
**Figure 6.21: Distribution of *go* by chronological age**



We notice a general association between *go*-production and age. Falling frequencies correlate with increasing age ( $r^2 = .073$ ,  $p < .01$ ,  $r = -0.296$ ). But the scattergram shows that this is not the whole picture. *Go*-ratios cluster as bi-modal with a peak around age 17 and a later, more diffuse, secondary peak amongst the oldest age-groups. In order to smooth out the scatterplot, I have reduced scalar numerical age into 7 age-groups, which brings out the age curve better. Consider Figure 6.22:

<sup>34</sup> Judging from Macaulay's (2001) Glasgow data, it seems that Scottish English has a higher *go*-ratio overall with 25% for the young speakers versus 11% for the older speakers in 1997.

Figure 6.22: Distribution of *go*-frequency by chronological age<sup>35</sup>



*Go*'s patterning as represented in the above graph is as follows: it peaks in the age-group 16-17 at 22.55%. We notice a swift fall for younger as well as for older speakers. The downwards trend with increasing numerical age is continued in the age group 30-50 with only 1.15%. After this steadily falling curve through the ages, *go* finally rises again among speakers aged 51-61, 62-79 and increases to 2.88% in the oldest age group. Note in contrast *go*'s patterning in the US, where we did not have any such 'tail' for the higher age groups. I will now discuss the patterning of *go* in the UK in light of the discussion in Chapter 6.4.3.

As we know, *go* has been attested in British English for centuries and, according to native speakers' judgments, has been above the level of awareness for a long time. Perception data seems to speak in favour of an age graded status of *go*: The typical reaction of middle aged people is illustrated by our computer support person Michael Bennett's comment (p.c.) "reminds me of my youth". Note also the following posting on a web discussion group:

"Subject: Re: 'Stereotype' thread – evolved: ... I knew a (then) girl who spoke the exact same way, "he goes, she goes," back in the late 70s, and who endured much harassment for it from her elders. She grew up to use language just fine, and is now a professional writer of excellent quality" (<http://www.raycomm.com/techwhirl/archives/0011/techwhirl-0011-00530.html>, January 24<sup>th</sup>, 2004).

<sup>35</sup> Note that the speaker sampling amongst the older speakers is sufficiently spread (consider Appendix 1). There are zero speakers in the age-brackets 29-38, 40-44, 51-54 as well as for the ages 47, 48, 59, 60 and so on. Those gaps in the sampling put a notable constraint on the grouping according to age: In order to avoid empty cells, I had to keep speaker numbers in the age groups roughly constant. Thus, comparability was the major motivating factor for splitting up numerical age in brackets in Figure 6.23.

Hence, it looks as if the patterning of quotative *go* in the UK is due to unstable behaviour at the individual level and stable behaviour at the community level: it seems to be relatively stably age graded. But let us first look at the patterning amongst the younger age groups: The fact that the youngest age-group displays a swift fall of 6.5% versus the peak at age 16-17 goes counter to the hypothesis that a reallocation of [-canonical] from *like* to *go* could lead to incremental *go*-frequencies for the youngest speakers (as it has done in the US data). Note that frequencies of the newcomer variant *like* stay high for the youngest speakers but do not rise. Dipping frequencies for *go* cannot therefore be due to increasing *like*. I will discuss the patterning of the other contenders, unframed and *say*-quotes, later on.

How can we interpret the lower frequencies for *go* amongst the youngest speakers? Research by Eckert (2000, 2003), Bucholtz (1996, 1999) and Thorne (1993) has shown that the importance of the life-stage “adolescence” for the construction and maintenance of one’s personality. During this age period, which has been termed “social hothouse” (Eckert 2000), fashion, music, substance use, activities as well as linguistic variables are increasingly exploited and become semiotic resources for the creation of distinction (Irvine 2001, Mendoza-Denton 1996). In the same vein, Labov (2001: 515 - 516) points out that adolescents who enter the adult world are engaged in a struggle for status, friends, work etc. He argues that conforming and nonconforming to societal norms are a strategy to these ends via the linguistic channel. He then formulates the “Nonconformity Principle: *Ongoing linguistic changes are emblematic of nonconformity to established social norms of appropriate behavior, and are generated in the social milieu that most consistently defies those norms.*” (Labov 2001: 513). Consequently, we would expect vernacular linguistic features which are important for the creation and negotiation of identity and which signal that the speakers are distancing themselves from societal norms to be most prominent in the mid-teen age groups. Quotative *go* is potentially such a variant. Accordingly, we see the line in Figure 6.22 peak in the age group 16-17.

How then do we explain the 3% *go*-differential in the younger and even the 4% differential in the older age group in real time between 1994/5 and 2001? One

possible explanation is that we simply cannot compare the York and the Derby /Newcastle data.

The alternative hypothesis is that, through time, all speakers (possibly excepting the younger ones, who primarily carry the change in progress, *like*), increase their *go*-frequencies and thereby participate in the augmented *go*-ratios. We would thereby witness a linguistic situation which is unstable at the community level; the whole population increases its use of the variant. The underlying assumption for this scenario is that across the whole population, the introduction of the new, [-canonical] variant *like* leads to renewed interest in *go*. Hence, a new wave of language change could have triggered renewed prestige for *go* amongst certain sociolinguistic groups. Reallocation could then have taken place due to the fact that *go* and the new, salient quotative *like* both carry the underlying feature [-canonical]. *Go* could subsequently have gained in frequency, especially with the older speakers who, due to their more stable linguistic habitus (Chambers 1995:185, Labov 2001) are much less likely to take on a new lexical item. But by using more of another variant with the same underlying feature [-canonical], they still participate in the trend without participating in the variant itself.

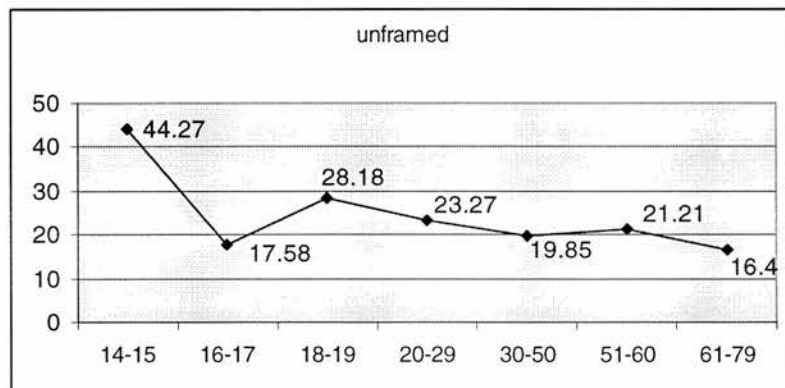
This hypothesis, which is supported by findings by Britain (1997) and Trudgill (1985), shows that the concept of reallocation can be useful for situations in which older speakers are reluctant to take on new variants but still take part in language change by associating a feature adherent to the new variant with an item already in their repertoire and vice versa. The high *like*-frequencies for younger speakers may thus be mapped across time onto correspondingly higher *go*-frequencies for people beyond the critical age. The reallocation hypothesis thus provides thus possible explanation which can account for *go*'s distribution in 1994/5, as shown in the line graph in Figure 6.22, while still being able to project higher numbers across time, as seen in Figure 6.17. An explanation of this type would also tie in with the discussion in Chapter 6.4.3.

To recapitulate, in the US we do not find any evidence of a "tail", whereas British English displays a very typical pattern for an age-graded distribution. How we can explain the different results from British and US English with respect to *go*'s patterning in apparent time? One tentative conclusion to a cross-varietal description in the spirit of Meyerhoff and Niedzielski (2003) is to consider *go* as both age-graded

and changing. The patterning in apparent time depends on the window we employ and the behaviour of individual and community at this particular moment in time. Looking at it from within a lifetime, the frequency distribution could look like a change in progress (US data) but from a cross-generational perspective, we would have to assume that it is at least to some extent age-graded (cf. the British English data). While *go* admittedly does exist with varying frequencies across generations, space and time, it has nevertheless played a role within the quotative system as a variant at the very least since the mid eighteenth century. Due to its latent presence in the linguistic repertoire of the community, *go* can be picked up either by younger speakers as a trend or by older speakers by way of reallocation.

Having discussed the two lexical non-canonical quotatives, I will now compare their patterning with the other vernacular strategy of the (say) variable, unframed quotes. Table 6.19 has given evidence that younger speakers in Britain highly favour this quotative option with  $p=.002$  ( $\chi^2(1):9.657$ ). While this result is contrary to what we find in the US data, Cameron notes for Puerto Rican Spanish that unframed strategies “are more closely connected to youth” (1998:70).<sup>36</sup> Figure 6.23 gives a line-graph of this variant’s patterning across chronological age.

**Figure 6.23: Distribution of unframed as a ratio by chronological age**



<sup>36</sup> Notice in this respect that both Cameron’s strategies *Y NP* (and *NP*) and *freestanding* are subsumed under my category *unframed* quotes. In Puerto Rican Spanish both are disfavoured by speakers over 20 whereas teens and preteens favour *Y NP* (and to a lesser extent the teens also *freestanding*).

Unframed-values slowly and steadily fall from the youngest to the oldest speakers with the only exception of a deep valley for the second youngest age group. This line suggests that frequencies for unframed quotes ascend smoothly to ages 16-17, where frequencies abruptly bottom out at 17.58% and then regain ground. The overall rising trend for unframed quotes is in line with findings that younger speakers use more vernacular variants generally. But we have to explain the incident that resulted in the indented line in the second youngest age group.

When we compare Figure 6.23 with the corresponding graphs for *like* (Figure 6.19) and *go* (Figure 6.22), we notice that the dip for unframed quotatives is synchronous with the peak in the other [-canonical] quotatives' frequency. The second youngest age group reverts to new (*like*) and highly salient (*go*) as non-canonical options. This choice of preference is to the detriment of the third choice within the vernacular category, unframed. The youngest speakers, who have lower frequencies for *like* and *go*, produce higher proportions of the unframed variant. It is still to be conclusively investigated whether the high frequencies of the unframed variant during the early teens is due to developmental reasons, due to the speakers' early pubescent status, or whether the *like*-trend has triggered higher unframed frequencies.

Even if the total differential between the ages 16-17 and the surrounding two age groups is not fully accounted for by the comparative highs in *like* and *go*'s frequencies, it seems clear that the three non-canonical variants are interacting with one another. In the age ranges where *like* and *go* gain ground, they intrude into the territory of unframed quotes. When unframed quotes rise again amongst the youngest speakers, *go*-frequencies fall. Again, reallocation between variants with the same underlying feature seems a likely explanation for this intimate interplay.

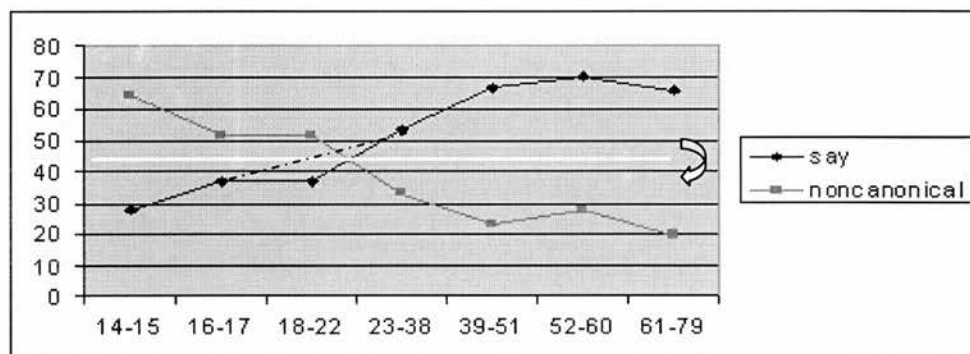
It remains to be discussed how the standard quotative *say* patterns across age and whether its frequency distribution shows signs of interaction with any of the non-canonical options. Given the fact that there is a close correspondence between the frequency patterns of all three non-canonical variants, we would not expect the change in progress *like* to have much influence on the *say*-output. Rather, it seems likely that *like*, *go* and unframed quotes compete with one another within their synsemantic field, with little or no effect on *say*.



Older speakers' preference for the older, standard quotative has been pointed out by Johnson (2001), Ferrara and Bell (1999) and numerous other authors. This is in line with general statements about the more conservative habitus of older speakers as well as with the fact that older speakers show more resistance to adopting new linguistic variants, such as *like* (Blyth et al. 1990, Singler 2001). This tendency is underlined by my data, where the age-difference with respect to *say*-use is highly significant (with  $\chi^2(1):171.71, p < .01$ ).

The following line graph depicts *say*-frequencies contrastively with the patterning of non-canonical variants (*go, like* and unframed quotes combined). This display enables us to see whether there is a direct inverse relationship between *say* and non-canonical frequencies.

**Figure 6.24: Patterning of non-canonical and *say*-variants across apparent time**



Note: The broken line depicts the hypothetical perfectly monotonous fall for *say* across age. The solid line shows the actual patterning with a marked disturbance in the age range 18-22, probably due to high *like* and *go* ratios in this particular bracket.

Figure 6.24 depicts – in corroboration of the discussion above – a falling line for *say* with decreasing age. This fall in the canonical variant corresponds quite neatly with a rising line for the non-canonical options. Holmes' prediction that “the use of standard or prestige forms peaks between the ages of 30 and 55 when people experience maximum societal pressure to conform” (1992:186) is borne out by this diagram. Here, it is the speakers from 52-60 who produce the highest *say*-frequencies (cf. also Downes 1984:191). The lines seem to exactly mirror one another (centering around

the axis depicted by the white bar).<sup>37</sup> The relationship between canonical/vernacular options is directly converse.

To get back to the initial question: does the intrusion of *like* into the pool of quotatives have any effect on *say*'s patterning in British English at all? The only time when the slope of the smoothly falling/rising lines of the converse vernacular/canonical pair (r-coefficients *say*–noncanonical:  $-.959$ ,  $p=.001$ ) is notably disturbed is in the age group 18-22. This is an age when *like* as well as *go* ratios are high. But note that both curves level out after that. Rather than falling due to the intrusion of newcomer *like*, *say* ratios recuperate and continue falling with their usual slope for the last two age brackets. This fact underlines again that *say*, at the initial stages of the introduction of *like* into British English, is relatively unaffected by the introduction of a new quotative option into the pool of quotative devices.<sup>38</sup> *Say*-ratios react to the corresponding peaks amongst other, noncanonical variants in 18-22 the age-group. But Figure 6.24 shows that the differential is not even 10% (consider the difference between the broken and the solid line). Thus, in 1994/5, the newcomer *like*, while certainly interacting with canonical *say*, only impinges on it to a very limited degree. As a [-canonical] quotative, *like* rather interacts with other options (such as *go* and unframed quotes) within its synsemantic field and only minimally spread into the ground covered by quotative *say*.

## 6.6. Conclusion

In tune with Weinreich, Labov and Herzog's (1968:188) claim that "all change involves variability and heterogeneity", the (*say*) variable as a whole has reacted to newcomers to the pool. We consequently find orderly variation in the quotative system which can be represented as statistically significant co-variance between linguistic variables and social predictors. Overall, the factor that plays the most

---

<sup>37</sup> Note in this respect that 3 variants of the (*say*) variable are not taken into consideration for this graph: *think*, *tell* and other quotatives, up to 14% overall. This leads me to conclude that the remaining variants do not partake much in the intimate interplay between the option *say* and noncanonical variants. The only time the curves do not perfectly mirror one another is in the 52-60 age bracket, where we can consequently expect some interaction with the remaining quotative options.

<sup>38</sup> But note in this respect Figure 17. From the later York data it seems that, over time, *like* does actually intrude into the ground covered by *say*.

important role is the age of the speaker. With respect to *like* we witness the same trend in US and British English in both varieties, the typical line-graph of a change in progress (Labov 1994, 2001). While *go* also patterns by age, diachronic information suggests that we have to assume a more complicated picture.

But overall, the extralinguistic constraints are only moderately exciting. This applies especially to gender (which accounts for the fact that in the literature we find that some studies claim that *like* is typically male and some claim that it is typically female) but also to class. The finding that social effects are not very robust in the (say) variable is in tune with the studies that have investigated both intralinguistic and extralinguistic constraints (Bell 1984, Eckert 2000, Labov 1994, Meyerhoff 1999, 2000, Preston 1991, Rickford and McNair-Knox 1996). In situations of ordered variability the linguistic constraints are often more important than the social ones. Accordingly, as I have shown in Chapter 3, variation at the system internal level (hypotheticality level, priming effects, mimesis) mainly governs *like*'s and *go*'s distribution. This chapter demonstrates that the effect that social factors have on the variability of the (say) variable is only very mild.

The discussion of the (say) variable has shown that the quotative systems in US and British English have reacted in an idiosyncratic manner to intruding *like*. They constitute internally coherent, self-sustained systems with their individual constraints and accommodate a new variant in their own system-specific ways. Consequently, we notice that, after the introduction of a newcomer variant into a delimited set of variants, social reallocation takes place within the respective systems, British English as well as US English. Hence, social constraints on the (say) variable are far from consistent across varieties. We cannot generalise social facts across the Atlantic. It is this system-specificity which led me to treat the two varieties separately in the first place.

Furthermore, quotative *like*, far from being transferred with all the social meaning it bears in US English, seems to lead a very independent life in the UK. Even if it has been borrowed from American English (which is, given the naturalness of the development, only one possibility), it can only have been taken over as the

surface item. *Like's* social meaning, which is by definition high context information and arbitrarily assigned, has not been transferred from the donor variety into the borrowing variety. Consequently, after *like's* introduction into the pool of quotatives in British English, new external, social meaning must have been attributed to it (cf. findings reported in Johnson 2001, Tagliamonte and Hudson 1999). This finding underlines claims that the mass media is an insufficient transmission channel for the whole variable.

It could be argued that this chapter provides comparative evidence of how the British English system with incipient *like could* look in a few years' time. But firstly such a comparative view merely presents one of a multiplicity of possible developments the (say) variable could take. Secondly, in the light of the functional autonomy of British English as demonstrated in Chapter 5 and the independent patterning on the external plane demonstrated above, it is clear that the development of the earlier US system might not necessarily be an indication of the future development of the British English (say) variable.

The historical precedent for competition among dialogue introducers indicates that the arrival of a new form warrants scrutiny in order to determine pragmatic function, processes of grammaticalization, forces of obsolescence and social distribution. (Ferrara and Bell 1995: 269)

## Chapter 7

### Conclusion and Outlook

The advent of *like* as a variant within the (say) variable has provided an exceptionally interesting research site. This is mainly for two reasons: (i) as Labov (2001) has pointed out, we know relatively little about syntactic or discourse changes in progress. The on-going transformation and refunctionalisation within the field of quotatives offers sociolinguists the opportunity to investigate such a change in progress. And (ii) because *like* intrusion imbalances the interaction amongst a range of variants within the (say) variable. A look at the entire quotative system reveals the intimate interplay of competing choices amongst the variants in the pool. An investigation of the (say) variable system presents a unique opportunity to study the restructuring of a whole system during the insertion of an innovative member.

This thesis is a variationist investigation of the rule-governed variability of innovative features and their rival variants (*say, tell, think, cry, ...*) in order to show the constraints which govern the quotative system as a whole. I have used a synthetic approach in order to investigate the advent of new vernacular options within the (say) variable under pragmatic, syntactic, discursal, and social aspects. I have attempted to (i) explain variation at all levels (linguistic as well as social) within the variable in order to account for as much of the variability as possible and (ii) investigate the systematic interrelation between all competing variants within the quotative pool. But a large-scale quantitative study, which only considers general trends, tends to simplify many phenomena. This is because it cannot account for the single case, which is especially important in cases of incipient language change. Qualitative analyses on the other hand tend to defy generalisation and often do not allow the drawing of general conclusions. Hence, in line with Wodak and Benke (1997), I have opted for a “combination of

methods, a multimethod approach in which different aspects of the object under investigation are grasped by different quantitative and qualitative methods which complement and do not exclude each other". In this present study, I have conducted a qualitative, discourse analytic and a quantitative, variationist sociolinguistic investigation of the pool of quotative devices.

After having defined my data and motivated my variable (Chapter 2), I first investigated the intralinguistic factors in the quotative frame. Earlier research had shown that quotative introducers can be used with reported speech and inner monologue. The problem with these accounts was that the question of what constitutes speech and what constitutes thought was not very well resolved. While most studies showed that they were aware of the problem (Butters 1982, Ferrara and Bell 1995, Romaine and Lange 1991, Tagliamonte and Hudson 1999, Tannen 1986) the question of how to resolve it was not fully investigated. In this thesis, I have given an indication of how we can fruitfully parse the epistemic continuum with respect to reported speech. I have demonstrated that we need to define our categories before we use them and suggested a clear contextually rooted pragmatic description of my categories. Setting up a hypotheticality continuum, I have shown in Chapter 3 that all quotatives, including the newcomers *like* and *go*, pattern in their own way with respect to epistemic stances. Due to the fact that they all have their own functional niche, and by association with this particular function, all quotative variants fulfil their specific task: *say* and *go* introduce real occurring quotes, *think* and *like* introduce hypothetical quotes and *like* also introduces situational quotes. I have also provided evidence that *like* is the most indeterminate in this respect and that it parallels the non-lexical variant unframed quotes in this functionally flexible patterning. This functional versatility might be one of the reasons why *like* has become so frequent in the lect of many young speakers.

I have also presented an examination of the "delivery aspects" (Clark and Gerrig 1990) of quotative variants, their co-occurrence with inner state markers or "response cries" (Goffman 1981), and "birds of a feather" priming effects (Scherre and Naro 1991, 1992). Overall, an investigation into the intralinguistic patterning of all quotative variants has shown that the quotative pool has reacted to the intrusion of two incoming items. There has been a redistribution and refunctionalisation of all quotatives as the

new ones have come in. As a consequence, the functional load is neatly distributed between all members of the quotative cohort.

The fact that the new quotatives pattern quite independently of the rest of the pool underlines their functional load. Hence, I have argued that *like* and *go* are not vacuous but indeed have their justified place amongst quotative devices. This claim can be justified by the fact that they have taken on quite novel functions with respect to mimetic enactments, the marking of epistemicity, speaker role demarcation etc.

Chapter 4 has provided a motivation of *like* and *go*'s recent functional extension to quotative introductory items. I have given evidence that the quotative introducers *like* (Buchstaller 2001 a,b) and *go* (Buchstaller, in preparation) are polyfunctional and build up a whole network of relations in conversational English. I have shown that only when we take into account all instances of *like* and *go*, and only if we allow for the case-wise ambiguity and multifunctionality any individual cases may display in real occurring speech, can we attempt to give a coherent account of how the quotative function can be linked up with other functions. I have proposed a solution to how *like*'s (Chapter 4.3) and *go*'s (Chapter 4.4) synchronic functions can be conceptualized as motivated and interrelated. Later functions, such as the quotative use, can be motivated as functional extensions from a common core and are connected via cognitive links in a radial structure (Lakoff 1987). These structures highlight the multifunctionality and the overlap of the various functions associated with *like* and *go* that are evidenced in my data. We can thus explain how transfers in the linguistic system result in newer, conceptually connected, functions. Language can be shown to be the reflection of conceptual manipulation.

*Like* and *go*'s widespread attestation has led researchers to compare their patterning with respect to social and linguistic factors in different places. A comparative investigation presupposes the variants' general linguistic comparability across varieties. However, sociolinguistics has shown in numerous studies that regionality has an impact on language structure. It seems in order to conduct a cross-variety assessment (Chapter 5) so as to be able to answer the important question of how much we can generalise the findings reported thus far. Do locally separate systems handle the situation in the same way? It seems that to assume that superficial resemblance (in this case the same surface

form) indicates overall equivalent linguistic constraints in various varieties is to run the risk of comparing apples with oranges. I suggest that it is important to recognise that while the underlying processes are the same, the surface output differs. Indeed, a comparison of the relevant pragmatic and syntactic factors in two varieties of English revealed some interesting differences in *go*'s and *like*'s development i.e. *go* occurs with indirect quotations in British English but not in US English (cf. Butters 1980).

This kind of cross-dialectal evaluation provides an important foundation for further research on the very nature of grammaticalisation. Is it meaningful, for instance, to consider the grammaticalisation of *like* and *go* as quotatives in British English and US English as constituting "the same" phenomenon? As Therese Lindstrom (p.c.) pointed out to me, one could talk about the grammaticalisations of *go* and *like* within English (according to variety or other sociolinguistic variables) and show what they have in common and what differs. This would certainly have the positive effect of making it possible to work without the simplistic notion of one grammaticalisation channel per language while trying to search for common ground amongst different varieties. As I noted in Chapter 5, these findings fuel discussion of the importance of a social dimension in grammaticalisation.

Moreover, I have demonstrated that the combination of the concepts of grammaticalisation and variationist sociolinguistics proves to be very fruitful. Not only does it make sense to turn to grammaticalisation in the search for generally valid principles of language change but a grammaticalisation approach can also help to define what a sociolinguistic variant is. Chapter 5 has shown how sociolinguists can take on board grammaticalisation as a means of assessing whether our variants generally function within the linguistic system in the same way. Applied this way, grammaticalisation helps ratify and rationalise our decision about whether or not a variant is actually "the same". We can thus ascertain general comparability across varieties, registers and other lects. In our specific case, I argued that only when we are in a position to state functional equivalence do we have justifiable reason to claim that a potential variant is behaving fundamentally as the same thing. And only once this general sameness is asserted can we investigate how this variant patterns with respect to extralinguistic or intralinguistic factors such as variety.



Having ratified my variables in this way, I presented a systematic study of the extralinguistic constraints on the quotative system in Chapter 6. On the basis of my data, I examined the social constraints in the two varieties and investigated to what extent they hold generally. In cases such as the rapid language change phenomenon *like*, it is important to explore which social groups introduce the newcomer into the quotative system in the two varieties. I have found orderly variation in the quotative system on the social plane and have given evidence that the (say) variable has reacted to the new addition to the pool. But while previous chapters have underlined *like* and *go*'s important intralinguistic constraints, Chapter 6 gives evidence that their extralinguistic constraints are much less remarkable. This finding constitutes an important argument for a synthetic approach to variationist sociolinguistics which, in explaining the variability of the variable, takes linguistic and social factors into account. My findings complement a growing body of research which has productively investigated the import of linguistic factors in situations of competing grammars and which has often found functional constraints to be much more important and robust than social factors. Accordingly, variation at the system internal level (hypotheticality level, priming effects, mimesis, ... ) mainly governs *like*'s and *go*'s distribution. But in order to understand the constraints that have a bearing on a given variable as fully as possible, we need to account for all factors, social as well as linguistic.

The question of whether *like*'s attestation in a multiplicity of varieties constitutes a borrowing from the US (where it was first noted) or an independent linguistic development cannot be easily resolved. With respect to discourse marker *like*, Fleischman and Yaguello write of the francophone *comme*, that it might "be a loan translation of *like*, an independent development, or an independent development influenced by *like*". On the basis of the evidence presented in this thesis, I would argue in favour of a combinatory approach. Given the naturalness of the underlying process (Chapter 4), the development of *like*'s quotative function in British English and other varieties might well have been independent. But we have to assume that the process of *like*'s functional extension in the UK was at least influenced by US English, a variety in which the grammaticalisation of quotative *like* was attested at least seven years earlier.

The very social group which introduces *like* into the linguistic system is also the one we can assume to be most likely to pick up a new lexical variant through soaps, talkshows, etc. But Chapters 5 and 6 have shown that the introduction of *like* into British English, while very probably sparked off or helped by its already further advanced status in US English, does not mirror the American functional and social trajectory. The quotative systems in US and British English have reacted in a locally idiosyncratic manner to the intrusion of newcomer *like*. They constitute internally coherent, self-sustained organisms with their individual constraints and accommodate a new variant in their own system-specific way, on the intralinguistic as well as the extralinguistic level.

This finding further underlines the claim that a linguistic sign consists of the surface item, the functional value and the social value. While a surface item can get diffused quite straightforwardly, the constraints this variant acquires once it is borrowed are determined by two aspects:<sup>1</sup> The first is the functional equilibrium within the accommodating variable. Obviously, we will expect the distribution in which the system had previously stabilised, the established allotment of older variants into functional niches, to have a certain amount of influence on the functional reallocation that occurs once *like* is introduced. The second is the social facts that have a bearing on the accommodating system. Given the fact that the allocation of social meaning is arbitrary and highly culture specific, we would expect an idiosyncratic local patterning of sociolinguistic variables with respect to social constraints. The social associations which newcomer variants acquire are a matter of convention and are attributed by the borrowing community itself. The data presented in Chapter 6 has indeed revealed that *like's* social meaning, which is by definition high context information (von Hippel 1994), has not been transferred straightforwardly from the donor variety into the borrowing variety.

Note that this finding also further corroborates claims that the mass media is an insufficient transmission channel for the whole variant (surface item, functional value and social value). If *like* was transmitted from the USA to the UK, the social information adherent to it in the donor variety was not taken over but seems to have

---

<sup>1</sup> A third, orthogonal, factor that plays into the acquisition of functions is the criterion of naturalness, which hinges on underlying and persistent semantic-pragmatic traits as I discussed in Chapter 4. I will not elaborate on this issue here.

been re-created in the borrowing variety, as has been suggested by Meyerhoff and Niedzielski (2002) and Buchstaller (2003). What seems to be taken over is the surface form but not the systematic functional and social load. Far from paralleling the US development, *like* on this side of the Atlantic takes on an individual profile, both social and linguistic. The transfer of a surface form goes hand in hand with reallocation and reappropriation in the borrowing variety.

In line with Meillet's (1926:17-18) claims that social and linguistic structure correspond and that general changes in one are translated into changes in the other, the data presented in this thesis show that after the introduction of a newcomer variant into the (say) variable, social and functional reallocation takes place within the respective systems, British English and US English. This finding has an important bearing on research on globalisation phenomena and so-called linguistic "mega-trends" (Buchstaller 2003, Kerswill 2003, Meyerhoff and Niedzielski 2003, Tagliamonte and Hudson 1999, Trudgill 1983, 1994, and many others).

The findings reported in this thesis are hence important for a whole range of reasons: the question of globalisation of innovative variants, of competition within the same socio-pragmatic field, of functional restructuring with numerous competing variants within one variable, the role of grammaticalisation in sociolinguistic research and the role of extralinguistic and intralinguistic constraints.

One of the questions that remains open is the future of the quotative system, post *like*. There has been lengthy speculation in the sociolinguistic literature about the possible outcomes of this linguistic change. Will the system stabilise in its richer form? Or, alternatively, will the advent of *like* drive out one of the older members of the quotative cohort? As Ferrara and Bell (1995) have pointed out, the introduction of a new quotative, *tellan* did indeed once lead to another member of the quotative complex, *cweðan*, being pushed out during the Old English period (even though *quoth* continued to exist in literary styles and numerous dialects until the Modern English period.). It remains to be seen whether the introduction of *like* into the pool of quotatives will result in a reduction of the set to the original number or whether the neat task-sharing within the quotative complex will maintain its stability.

The question remains: will *like* survive? Some researchers are not convinced. Dailey-O’Cain (2000) shows that *like* is stigmatized by society at large. Bailey (in Cukor-Avila 2002: 22) sees it more as a sociolectal adjustment (Chambers 1995) of speakers who try to fit in with the latest trends in society. Others, such as Linda Coleman, director of the Freshman Writing Program at the University of Maryland, College Park suggest that “‘like’ is something young people will grow out of. Once students establish more stable relationships, ‘like’ (...) will be employed less often”. (<http://www.inform.umd.edu/News/Diamondback/archives/2002/09/16/news5.html>). But most researchers are careful not to declare their position. For good reason because why, if *like*-users sound “like an airhead” (Guy in Cukor-Avila 2003: 22) does *like* pop up in most varieties around the world and maintain its grip on the quotative system of speakers as they grow older (Singler 2001)? It has even been introduced into the written media (Romaine and Lange 1991, Singler and Woods 2002).

There is, to date, no evidence that *like* will fade away. Even though the first *like*-users are only in their 20s and 30s, as the generation who started the change in progress grows older, the age limit in which *like* is found is continuously reported to move upwards as well (Singler 2001). Also, from a structural point of view, the loss of *like* seems highly unlikely. As one quotative option amongst many reported dialogue introducers, *like* certainly has a big advantage over the traditional variants: it can be used for internal and external speech, approximate or exact renderings, speech, sounds, voice effects and gestures. As Ferrara (1992) has pointed out, on-line speech production requires a multiplicity of fast choices. The flexibility of *like* makes it a ready-made item, available in multifold discourse planning situations. Furthermore, Chapter 4 has given evidence that *like*’s quotative function is cross-linguistically sustained. The development of an item meaning ‘similar to’ to an introducer of reported speech and thought is a very natural and frequently attested process. Well stabilised through its underlying semantic-pragmatic field, quotative *like* has the potential to stay.

But only a real time study can reveal whether the generation which uses *like* now will continue or whether they will eventually stop using it as they grow older. Singler (2001) is right when he points out that the trait “young” is perceptually attached to *like*, which might, in the long run, lead to a stably age-graded distribution. But the last few years have seen the rise of the phenomenon “youth culture” (Bucholtz 2000, Pujolar

2000). Google produced 250,000 hits for the phrase “youth culture” on February 7<sup>th</sup>, 2004, as opposed to only 108,000 hits on November 17, 2003. Much more than formerly (Eckert 2003), the formative teenage years are being seen as more than merely a step towards adulthood but are considered as a way-of-life. Society at large has been reifying this age-group, fuelled by the commercial industry which has long ago discovered its market potential. With *like* and *go* being so highly associated with youth culture and its expression via American soaps, it remains to be seen which influence this non-canonical [+youth] feature might exert on older speakers. It will have to be explored whether speakers drop their *likes*, just as I have shown them to drop their *goes* in British English, once they have matured out of their ‘Friends’ and ‘Sex in the City’ phase. Without real time evidence, it is impossible to tell whether *like* will only persist as a fad of the generation that uses it today, whether it will result as an age-graded feature, or whether its spread will continue. We will need data from later time slices in order to establish whether or not age-stratified patterns of variation in the data actually reflect a change in progress in the long term.

It is true that the comparative evidence presented in this thesis gives time-depth to the investigation. It could be argued that US English provides an indication of how the British English system with incipient *like* could look in a few years. In this case we could treat the two corpora as an embryonic and a developed stage. We would consider the British data as the patterning of a new quotative item at its very emergence and treat the US data as if it was an advanced stage of British English. But in the light of the functional autonomy of British English as demonstrated in Chapter 5 and its independent patterning on the external plane demonstrated in Chapter 6, it is clear that the development of the US system might not necessarily be an indication of the future development of the British English (say) variable. Rather, British English constitutes an independent development where *like*, from the very beginning, invents itself. This is why I have treated the two systems as self-contained and hermetic entities. We would need real-time data from the respective varieties to see how the individual systems evolve in their own right.

Furthermore, newer quotative options were reported directly subsequent to *like*’s earliest attestation. Stein (1990) documents *I’m sittin’ there* as a new quotative. *Be all*

was first mentioned by Lamb, Igoe, and Gilman (1999) and has consequently been pointed out in numerous varieties. While my data only contain very small token numbers of this quotative (US English 1 token, British English 4 tokens), its persistence and robustness in California led Labov to issue a general call for the investigation of this quotative at the NWave in 2003. Other newcomers, ephemeral as well as longer lived ones, are surely in the making. Indeed, it has been claimed in the literature that *go* and *like* will be driven out in a chain-like fashion by yet newer quotatives such as *be all* (Bakht-Rofheart 2002, Cukor-Avila 2002, Ferrara and Bell 1995). But only a more recent corpus, post-*be all*, can show the future composition of the quotative system after this newest robust arrival. An exhaustive investigation of the relationships the newcomer quotatives develop between one another in the language external and internal plane remains a project for the future.

Finally, the results of this study bring us a step nearer to answering the fundamental questions raised by Ferrara and Bell (1995:269): What factors cause certain forms to prevail while others die out? Are those processes cyclic? And why, if there is such a variety of quotatives readily available, is there a need for another quotative?

An integrated approach such as the one presented in this thesis can provide an answer as to why some forms prevail while others die out. An underlying semantic pragmatic field which is cross-linguistically sustained provides an indication that persistence is at least sustained by a highly natural underlying process. Also, if we look at the quotative system post *like*, we notice that the introduction of a newcomer variant has resulted in an important restructuring of the quotative system and that this reorganisation produced a situation of stable and orderly variability depending on internal as well as external factors. Such a firmly organised system, where every variant retains its own functional and social niche, has the potential for diachronic stability. Cukor-Avila (2003) equates the spread of *like* to the spread of *yall* outside the South. As Tillery, Wikle and Bailey (2000) have shown, *yall* also fulfils important grammatical and pragmatic functions, filling the systematic gap in the paradigm of personal pronouns for the second person plural. But whereas *yall*'s morphological shape mirrors already existing forms, which gives it an easy entry path into the system, *like* and *go* do not have any predecessors with formal similarities. Yet, Schourup (1982) for *go* and

Cukor-Avila (2002) following him with *like* make the point that they disambiguate indirect from direct quotes. In Chapter 3 of this thesis I have shown *like* and *go*'s functional load in a larger spectrum (with respect to mimesis, epistemic stances, speaker disambiguation etc.). Hence, we are in a position to assert that a defined and delimited functional niche for the variants as well as a balanced distribution within the variable are factors that favour persistence. The absence of these factors will probably not cause innovative forms to die out but will offer them a less accommodating environment.

As for whether such processes are cyclic, it needs to be pointed out that this study is not diachronic. But even from a synchronic point of view, we come across a number of attestations of newcomer variants in the quotative pool. *Go* as an introducer for sounds and voice seems to have made its way into the quotative system in 1789 and has enlarged its potential to the introduction of all quotes at some point before 1980. Dubois (1989) mentions what she calls the "pseudoquotation" *Hey* which takes the form 'Hey, she didn't really say it'. A little later, Butters first mentions *like* and Igoe, Kim and Gilman draw our attention to *all*. Stein (1990:303) further points out *I'm sitting here*. Numerous other, more short-lived innovations have surely come and gone in the meantime without having been picked up in the literature or without having caught my attention. While none of these introductions have led to the clear-cut ousting scenario described by Ferrara and Bell (1995) for *cweðan* in Old English, we can nevertheless conclude that the quotative frame seems to be a site for repeated innovative language use. Below, I will argue that this is no surprise given the highly expressive genre, narratives, in which quotations tend to occur and given the fact that quotation constitutes highly stylised theatrical rendering of past events.

Why then is there a need for newcomer quotatives? The question can be answered by pointing to three particularities of quotations: (i) the genre and the place in the narrative sequence in which they tend to be found (ii) their important pragmatic status as highly salient introducers of theatrical instances of animating others' voices and (iii) the social profile of the most frequent users of quotation.

The most prototypical genre of occurrence of quotation, narrative, is very frequent and largely informal, a factor which favours innovative language use. We can furthermore assume that quotations occur at crucial points during the narrative sequence (Labov and Waletzky 1967), where we would expect a "breakthrough into

performance” (Hymes 1975) and where speakers expectedly talk at their most vernacular. It is at this site that linguistic creativity is most freely exerted and that narrators are under the most pressure to perform captivating and attention-catching output in order to keep the narrative floor. Theatrical mimicking of others’ voices bears an enormous expressive potential. These pragmatic and discorsal factors favour the use or creation of linguistic innovations.

Furthermore, as Nordberg (1980) has pointed out, quotation is an important linguistic style marker for adolescents. We know that the social group who use most quotation, and where *like*-frequencies peak, adolescents, tends to exploit vernacular linguistic features (innovative ones as well as older ones). Distancing themselves from societal norms (Labov’s ‘Nonconformity Principle’, 2001: 513) and engaged in the creation and negotiation of personal style via linguistic and other ends, adolescents lead all other age groups in innovative language use. Consequently, sociolinguistic research has shown on numerous occasions that creative use of new variants peak in adolescence. Hence, it should come as no surprise that adolescents have come up with and continue to fashion ingenious new variants (*like*, *all*) for an old variable. This is all the more expectable as the prototypical context of occurrence of the variable calls for effect and dramatic heightening. In conclusion, the quotative frame seems to be a perfect site for newcomer variants.

This thesis has shed some light on the social repercussions and functional distribution of one innovative variant, *like*. I will now pass the ball to whichever researcher answers to the need for an investigation of *be all* and other newer variants.



## Appendix 1: Sociolinguistic Information

### I. Sociolinguistic Distribution in the British English data

#### Derby

**Table 1: Matrix of Derby speakers**

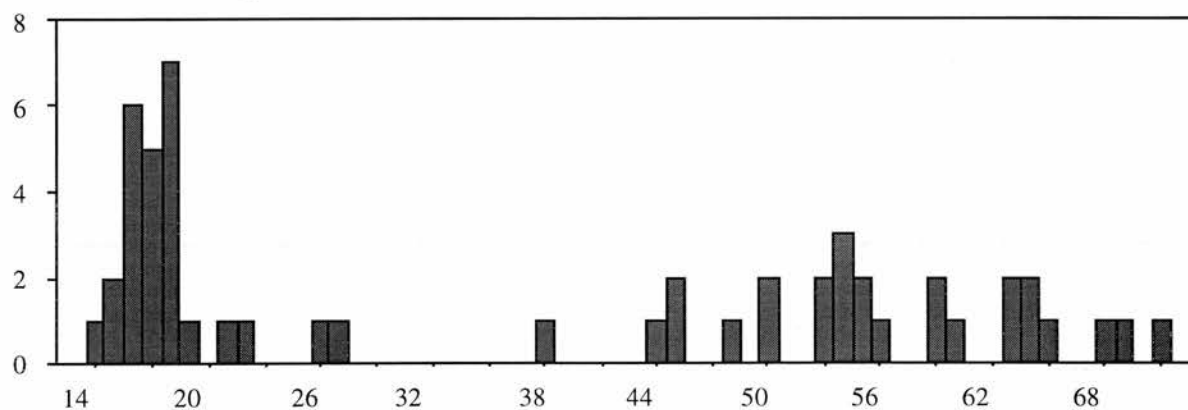
	MC female	MC male	WC female	WC male
old	4	4	4	4
young	4	4	4	4

#### Newcastle

**Table 2: Matrix of Newcastle speakers**

	MC female	MC male	WC female	WC male
old	4	4	4	4
young	4	4	4	4

**Figure 1: Distribution of British English speakers across chronological age**

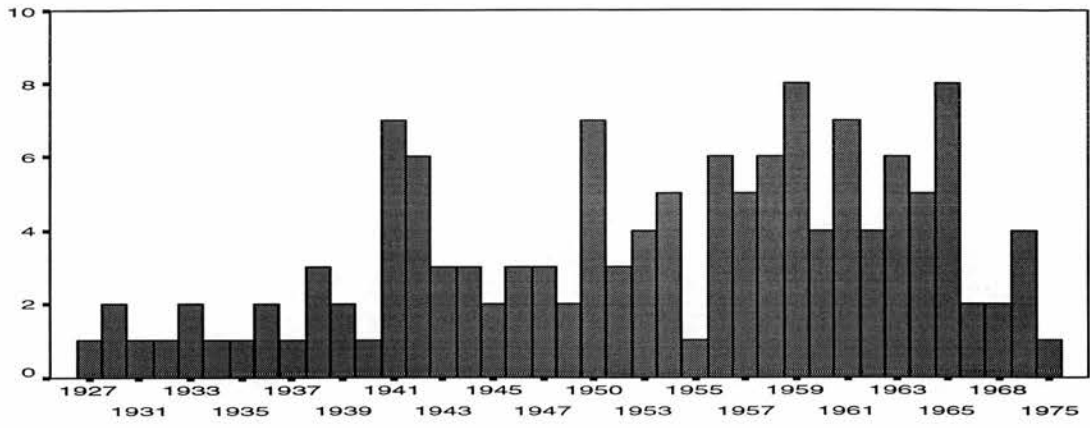


### II. Sociolinguistic distribution in the US data

**Table 3: Matrix of US speakers**

		young	old
lower	male	17	17
	female	17	17
educated	male	17	17
	female	17	17

Figure 2: Distribution of US speakers by year of birth



## Appendix 2: Transcription Conventions

carriage return	intonation unit
[ ]	overlap
=	quick, immediate connection of new turns or single units
(.)	micro-pause
(-), (--)	short, middle pause
:::	lengthening, according to its duration
?	high rise, appeal intonation
,	mid rise, continuing intonation
.	low fall, final intonation
( )	unintelligible passage, according to its duration
<u>accent</u>	primary or main accent
!ac!cent	extra strong accent
↑	pitch step up
↓	pitch step down
‘ ‘	signals for start and end of quote
@	laughter

## References

- Abdel-Jawad, Hassan. (1981). *Phonological and social variation in Arabic in Amman*. University of Pennsylvania dissertation.
- Aijmer, Karin. (2001). *I Think* as a Marker of Discourse Style in Argumentative Student Writing. *Gothenburg Studies in English* 81: 247-257.
- Aikio, Marjut. (1992). Are women innovators in a shift to a second language? A case study of reindeer Sami women and men. *International Journal of the Sociology of Language* 94.
- Akatsuka, Noriko. (1986). Conditionals and epistemic scale. *Language* 61(3): 625-639.
- The American Heritage College Dictionary. (1993). Houghton Mifflin Company.
- Andersen, Elaine, Brizuela, Maquela, Dupuy, Beatrice and Gonnermann, Laura. (1999). Cross-linguistic evidence for the early acquisition of discourse markers as register variables. *Journal of Pragmatics* 31: 1339-1351.
- Andersen, Gisle. (1996). They like wanna see like how we talk and all that: The use of like as a discourse marker in London teenage speech. *Corpus-Based Studies in English: Papers from the 17th International Conference on English Language Research on Computerized Corpora (ICAME 17)*. Stockholm.
- Andersen, Gisle. (1997). *I goes you hang it up in your shower innit? He goes yeah. The use and development of invariant tags and follow-ups in London teenage speech*. Paper presented at the 1st UK Language Variation Workshop. University of Reading.
- Andersen, Gisle. (1998). The pragmatic marker *like* from a Relevance-theoretic Perspective. In A. Jucker & Y. Ziv (Eds.), *Discourse Markers: Description and Theory*. Amsterdam, John Benjamins: 147-170.
- Andersen, Henning. (1989). Understanding linguistic innovations. In L.E. Breigik & E. H. Jahr (Eds.), *Language Change: Contributions to the study of Causes*. Berlin, Mouton de Gruyter: 5-27.
- Armbruster, Charles. (1960). *Donolese Nubian: A grammar*. Cambridge, Cambridge University Press.
- Auer, Peter. (1992). Introduction: John Gumperz' approach to contextualization. In P. Auer & A. di Luzio (Eds.), *The Contextualization of Language*. Amsterdam, Benjamins: 1-38.
- Auer, Peter. (1996). On the Prosody and Syntax of Turn-Continuations. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation: Interactional Studies*. Cambridge, Cambridge University Press: 57-100.

- Aukrust, Vibeke G. & Snow, Caroline E. (1998). Narratives and explanations in Norwegian and American mealtime conversations. *Language and Society*, 27: 221-246.
- Bailey, Charles-James. (1973). The patterning of language variation. In R.W. Bailey & J.L. Robinson (Eds.), *Varieties of present-day English*. London, Collier-Macmillan: 159-86.
- Bailey, Guy, Wikle, Tom, Tillery, Jan and Sand, Lori. (1991). The apparent time construct. *Language Variation and Change* 3: 241-64.
- Bakht-Rofheart, Maryam. (2002). *Avoidance of a New Standard: Quotative Use among Long Island Teenagers*. Paper presented at NWAVE 31, Stanford University.
- Bakhtin, Mikhail M. (1981). *The dialogic imagination: four essays*. Austin, Texas. University of Austin Press.
- Bakhtin, Mikhail M. (1984). *Problems of Dostoevsky's Poetics*. 1929. Trans. by Caryl Emerson. Manchester and New York, Manchester University Press.
- Bakhtin, Mikhail M. (1986). *Speech Genres and Other Late Essays*. Translation by Vern W. McGee. Austin, Texas, University of Texas Press.
- Bakhtin, Mikhail M. / Pavel N. Medvedev. (1978). *The Formal Method in Literary Scholarship*. 1928. Trans. by Albert J. Wehrle. Baltimore and London, The John Hopkins University Press.
- Banfield, Ann (1982). *Unspeakable sentences: Narration and representation in the language of fiction*. Boston, Routledge and Kegan Paul.
- Barth-Weingarten, Dagmar and Couper-Kuhlen, Elizabeth. (2002). On the development of final *though*: A case of grammaticalization? In I. Wischer & G. Diewald (Eds.), *New Reflections on Grammaticalization: Proceedings from the International Symposium on Grammaticalization*, Potsdam, Amsterdam, Benjamins: 363-78.
- Bauer, Robert S. (1982). *Lexical Diffusion in Hong Kong Cantonese: "Five" leads the way*. Paper presented at the 8<sup>th</sup> annual BLS meeting.
- Bauer, Robert S. (1986). The Microhistory of a Sound Change in Progress. *Journal of Chinese Linguistics* 14: 3-41.
- Baugh, John. (1979). *Linguistic style-shifting in Black English*. University of Pennsylvania dissertation.
- Baumann, Richard. (1986). *Story, Performance, and Event*. Cambridge: Cambridge University Press.

- Becker, A.L. (1988). Language in particular. An Essay. In D. Tannen (Ed.), *Linguistics in Context*. Norwood, NJ, Ablex: 17-35.
- Beheny, Richard. (1998). Interface Economy and Focus. In V. Rouchota & A.H. Jucker (Eds.), *Current Issues in Relevance Theory*. Amsterdam, Benjamins: 105-139.
- Bell, Allan. (1984). Language style as audience design. *Language in Society* 13: 145-204.
- Berckmans, Paul.(1993). The Quantifier Theory of *Even*. *Linguistics and Philosophy*, 16 (6):589-611.
- Berlin, Brent. (1992). *Ethnobiological Classification*. Princetown University Press.
- Bickerton, Derek. (ms). Experimental Creation of a Natural Language. Unpublished and undated manuscript. Honolulu, University of Hawaii at Manoa.
- Blank, Andreas. (1999). Co-Presence and Succession: A Cognitive Typology of Metonymy. In K.-U. Panther & G. Radden (Eds.), *Metonymy in Language and Thought*. Amsterdam, Benjamins: 169-191.
- Blake, Renée. (1997). Defining the envelope of linguistic variation: The case of 'don't count' forms in the copula analysis of African-American Vernacular English. *Language Variation and Change* 9: 57-80.
- Blakemore, Diane. (1987). *Semantic constraints on relevance*. Oxford, Blackwell.
- Blyth, Carl, Recktenwald, Sigrid and Wang, Jenny (1988). *I'm like, "Say what ?!"*. A new quotative in American oral narrative. Paper presented at NWAVE 17, Montreal.
- Blyth, Carl, Recktenwald, Sigrid and Wang, Jenny. (1990). I'm like, "Say what ?!". A new quotative in American oral narrative. *American Speech* 65: 215-227.
- Boeder, Winfred. (2003). Speech and thought representation in the Kartvelian (South Caucasian) languages. In T. Güldemann & M. von Roncador (Eds.), *Reported Discourse. A meeting ground for different linguistic domains. Typological studies in Language*. Amsterdam, Philadelphia, Benjamins: 3-48.
- Bolden, Galina. (to appear). The Quote and beyond: Defining Boundaries in Conversational Russian. *Journal of Pragmatics*.
- Bolinger, Dwight. (1976). Meaning and Memory. *Forum Linguisticum* 1(1): 1-14.
- Bolinger, Dwight. (1977). *Meaning and Form*. London, Longman.
- Bolinger, Dwight. (1989). *Intonation and its Uses. Melody and Discourse*. London, Arnold.

- Bortoni-Ricardo, Stella Maris. (1985). *The urbanization of rural dialect speakers: a sociolinguistic study in Brazil*. Cambridge, Cambridge University Press.
- Bourdieu, Pierre and Boltanski, Luc. (1975). Le fétichisme de la langue. *Actes de la recherche en sciences sociales* 4: 2-32.
- Briggs, Charles. (1992). "Since I am a woman, I will chastise my relatives." Gender, reported speech, and the (re)production of social relations in Warao ritual wailing. *American Ethnologist* 19: 337-361.
- British Heart Foundation. (2000). *Couch kids: The growing epidemic - looking at physical activity in children in the UK*.
- Brugman, Claudia. (1988). *The story of over: Polysemy, semantics, and the structure of the lexicon*. New York, Garland Publishing, Inc.
- Bucholtz, Mary. (1996). Geek the girl: Language, femininity and female nerds. In N. Warner, J. Ahlers, L. Bilmes, M. Oliver, S. Wertheim & M. Chen (Eds.), *Gender and Belief Systems*. Berkeley, CA, Berkeley Women and Language Group, University of California: 119-131.
- Bucholtz, Mary. (1999). "Why Be Normal?": Language and Identity Practices in a Community of Nerd Girls. *Language in Society* 28(2): 203-223.
- Bucholtz, Mary. (2000). *Language and Youth Culture*. *American Speech* 75 (3): 280-283.
- Buchstaller, Isabelle. (1999). *Causality, Conditionality and Concessivity in Hawaiian Creole*. Unpublished Master's Thesis (Zulassungsarbeit zum Staatsexamen). University of Konstanz.
- Buchstaller, Isabelle. (2001a). An alternative view of LIKE: its grammaticalization in conversational American English and beyond. *Edinburgh Working Papers in Applied Linguistics* 11.
- Buchstaller, Isabelle. (2001b). *He goes and I'm like: The new quotatives re-visited*. Paper presented at NWAVE 30, University of North Carolina.
- Buchstaller, Isabelle. (2002a). *BE like USE? BE goes USE?* Paper presented at NWAVE 31, Stanford University.
- Buchstaller Isabelle. (2002b). The Co-occurrence of Quotatives with Mimetic Performances. *Edinburgh Working Papers in Applied Linguistics* 12.
- Buchstaller, Isabelle. (2003). *Putting perception to the reality test: The case of go and like*. Paper presented at NWAVE 32, University of Pennsylvania.
- Buchstaller, Isabelle (in preparation). *The go-scenario: A radial structure account for go's synchronic functions*.

- Buck, Carl. (1915a). Words of speaking and saying in the Indo-European Languages. *American Journal of Philology* 36(1):1-18, (2):125-154.
- Butters, Ronald. (1980). Narrative Go 'Say'. *American Speech* 55: 304-07.
- Butters, Ronald. (1982). Editor's note [on be like 'think']. *American Speech* 57: 149.
- Bredin, Hugh. (1998). Comparisons and Similes. *Lingua* 105 (1-2): 67-78.
- Brinton, L. J. (1996). *Pragmatic Markers in English*. Mouton de Gruyter, Berlin, New York.
- Britain, David. (1997). Dialect contact and phonological reallocation: "Canadian raising" in the English Fens. *Language in Society* 26: 15-46.
- Britain, David. (1999). Where did it all start?: dialect contact, the 'Founder Principle' and the so-called (-own) split in New Zealand English. *Transactions of the Philological Society*. 99: 1-27.
- Britain, David. (2002). Diffusion, Levelling, Simplification and Reallocation in Past Tense BE in the English Fens. *Journal of Sociolinguistics* 6 (1): 16-43.
- Britain, David and Trudgill, Peter. (1999). Migration, New-Dialect Formation and Sociolinguistic Refunctionalisation: Reallocation as an Outcome of Dialect Contact. *Transactions of the Philological Society* 97 (2): 245-256.
- Brown, Penelope and Levinson, Steven. (1987). *Politeness: Some Universals on Language Use*. Cambridge, Cambridge University Press.
- Bybee, Joan. (1985). *Morphology. A Study of the Relation between Meaning and Form*. Amsterdam, Philadelphia, Benjamins.
- Bybee, Joan. (1988). Semantic substance versus contrast in the development of grammatical meaning. *Berkeley Linguistic Society* 14: 247-64.
- Bybee, Joan. (1995). Regular Morphology and the Lexicon. *Language and Cognitive Processes* 10 (5): 425-455.
- Bybee, Joan. (1997). Semantic aspects of morphological typology. In J. Bybee, J. Haiman & S. Thompson (Eds.) *Essays on language function and language type*. Amsterdam, Benjamins.
- Bybee, Joan and Hopper, Paul (Eds.) (2001). *Introduction. Frequency and the emergence of linguistic structure*. John Benjamins.
- Bybee, Joan and Pagliuca, William. (1985). Cross-linguistic comparison and the development of grammatical meaning. In J. Fisiak (Ed.), *Historical Semantics, Historical Word Formation*. The Hague, Mouton de Gruyter: 59-83.



- Bybee, Joan and Pagliuca, William. (1987). The evolution of future meaning. In A. Giacalone Ramat, O. Carruba & G. Bernini (Eds.), *Papers from the VIIth International Conference on Historical Linguistics*. Amsterdam, Benjamins: 109-122.
- Bybee, Joan, Pagliuca, William & Perkins, Revere. (1991). Back to the Future. In E. Traugott, & B. Heine (Eds.) *Approaches to Grammaticalization*. Vol. 2. Amsterdam, Benjamins: 17-58.
- Bybee, Joan, Perkins, Revere & Pagliuca, William. (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages around the World*. Chicago, University of Chicago Press.
- Cameron, Deborah and Coates, Jennifer. (1988). Some problems in the sociolinguistic explanation of sex differences. In D. Cameron & J. Coates (Eds.), *Women in speech communities: New Perspectives on Language and Sex*. London and New York, Routledge.
- Cameron, Richard. (1998). A variable syntax of speech, gesture, and sound effect: Direct quotations in Spanish. *Language Variation and Change* 10: 43-83.
- Cameron, Richard. (2000). Language Change or Changing Selves? Direct Quotation Strategies in the Spanish of San Juan, Puerto Rico. *Diachronica* 17 (2): 249-292.
- Carbaugh, Donal. (1988). *Talking American: Cultural Discourses on DONAHUE*. Norwood, New Jersey, Ablex.
- Cedegren, Henrietta. (1973). *The interplay of social and linguistic factors in Panama*. Unpublished Cornell Dissertation.
- Cedergren, Henrietta. (1984). *Panama Revisited: Sound Change in Real Time*. Paper presented at NWAVE 10, University of Philadelphia.
- Cedergren, Henrietta and Sankoff, David. (1974). Variable rules: performance as a statistical reflection of competence. *Language* 50: 333-55.
- Chafe, Wallace. (1980). *The Pear stories: cognitive, cultural, and linguistic aspects of narrative production*. Norwood, N.J.: Ablex.
- Chafe, Wallace. (1982). Integration and Involvement in Speaking, Writing, and Oral Literature. In D. Tannen (Ed.), *Spoken and Written Language: Exploring Orality and Literacy*. Norwood, NJ, Ablex: 35-53.
- Chafe, Wallace. (1994). *Discourse, Consciousness and Time*. Chicago and London, University of Chicago Press.
- Chambers, Jack K. (1995). *Sociolinguistic Theory. Linguistic Variation and its Social Significance*. Oxford and Cambridge, Blackwell.

- Channell, Joanna. (1994). *Vague Language*. Oxford, Oxford University Press.
- Cheshire, Jenny. (1982). *Variation in an English Dialect: A sociolinguistic study*. Cambridge, Cambridge University Press.
- Cheshire, Jenny. (1987). Syntactic variation, the linguistic variable, and sociolinguistic theory. *Linguistics* 25: 257-282.
- Cheshire, Jenny. (2000). The teller or the tale? Narrative and gender in adolescent friendship networks. *Journal of Sociolinguistics* 4: 234-262.
- Clark, Herbert and Gerrig, Richard. (1990). Quotations as Demonstrations. *Language* 66: 764-805.
- Claudi, Ulrike and Heine, Bernd. (1986). On the metaphorical basis of grammar. *Studies in Language* 10 (2): 297-335.
- Clifford, James. (1986). Introduction: Partial truths. In J. Clifford & H.G. Markus (Eds.), *Writing culture: The poetics and politics of ethnography*. Berkeley, University of California Press: 1-26.
- Clift, Rebecca. (2001). Meaning in Interaction: The Case of *actually*. *Language* 77 (2): 245-291.
- Cohen, David, Simeone-Senelle, Marie-Claude & Vanhove, Martine. (2003). The grammaticalization of "say" and "do": An areal phenomenon in East Africa. In T. Güldemann & M. von Roncador (Eds.), *Reported Discourse. A meeting ground for different linguistic domains. Typological studies in Language*. Amsterdam, Philadelphia, Benjamins: 227-252.
- The Collins Cobuild's Student's Dictionary online. COLLINS Birmingham University International Language Database (<http://www.linguistics.ruhr-uni-bochum.de/ccsd/>)
- Comrie, Bernard. (1985). *Tense*. Cambridge, Cambridge University Press.
- Comrie, Bernard. (1986). Conditionals: A Typology. In E. Traugott et al. (Eds.), *On Conditionals*. Cambridge, Cambridge University Press.
- Coulmas, Florian. (1981). *Conversational routine: explorations in standardized communication situations and pre-patterned speech*. The Hague, Mouton de Gruyter.
- Coulmas, Florian. (1983). *Direct and indirect speech*. Berlin, Mouton de Gruyter.
- Coulmas, Florian. (1985). Direct and indirect speech: General problems and problems of Japanese. *Journal of Pragmatics* 9: 41-63.

- Coulmas, Florian. (1986). Nobody dies in Shangri-La: direct and indirect speech across languages. In D. Tannen & J.E. Alatis (Eds.), *Language and Linguistics: The Interdependence of Theory, Data, and Applications*. Washington, DC, Georgetown University Press: 140-163.
- Couper-Kuhlen, Elizabeth. (1999). Varieties of conditionals and their emergence in discourse. In A. Lahiri, A. Patschovsky & Ch. Schwarze (Eds.), *Issues in Interdisciplinary Research on the Lexikon. Arbeitspapier Nr. 99, Fachgruppe Sprachwissenschaft der Universität Konstanz*: 89-130.
- Craig, Collette. (1991). Ways to Go in Rama: A Case Study in Polygrammaticalization. In E. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*. Vol. 2. Amsterdam, Benjamins: 455-492.
- Cukor-Avila, Patricia. (2001). *Change over time in the use of verbs of quotation in African American Vernacular English*. Paper presented at the University of York.
- Cukor-Avila, Patricia. (2002). She Say, She Go, She Be Like: Verbs of Quotation over Time in African American Vernacular English. *American Speech* 77 (1): 3-31.
- Culler, Jonathan. (1982). *On deconstruction: Theory and criticism after structuralism*. Ithaca, NY: Cornell University Press.
- Cutler, Cecilia. (1996). *Yorkville Crossing: a Case Study of the Influence of Hip Hop Culture on the Speech of a White Middle Class Adolescent in New York City*. Paper presented at NWAVE 25, Las Vegas.
- Cutler, Cecilia. (2002). *Crossing over: White Youth, Hip-Hop and African American English*. PhD dissertation, New York University.
- Dahl, Östen. (1985). *Tense and Aspect Systems*. Oxford, Blackwell.
- Dahl, Östen. (2001). Inflationary effects in Language and elsewhere. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. Amsterdam, Benjamins.
- Dailey-O'Cain, Jennifer. (2000). The sociolinguistic distribution and attitudes towards focuser like and quotative like. *Journal of Sociolinguistics* 4 (1): 60-80.
- Danchev, Andrej and Kytö, Merja. (1994). The construction *be going to + infinitive* in Early Modern English. In D. Kstovsky (Ed.), *Studies in Early Modern English*. Berlin, Mouton de Gruyter: 59-77.
- Danchev, Andrej and Kytö, Merja. (2002). The Go-Futures in English and French Viewed as an Areal Feature. *NOWELE* 40: 29-60.

- Davies, Catherine Evans. (2002). *Assessing the Effects of Language in the Media on Linguistic Variation*. Paper presented at NWAVE 31, Stanford University.
- Declerk, Renaat and Tanaka, Kazuhiko. (1996). Constraints on tense choice in reported speech. *Studia Linguistica* 50: 283-301.
- De Roeck, Marijke. (1994). Typology of speech reports. In E. Engberg-Pedersen et al. (Eds.), *Function and expression in functional grammar*. Berlin, Mouton de Gruyter:331-351.
- Dines, Elizabeth R. (1980). Variation in Discourse-"and Stuff Like That". *Language in Society* 9 (1): 13-31.
- Dinsmore, John. (1991). *Partitioned Representation*. Dordrech, Kluwer Academic Publishers.
- Dirven, Rene, Goosens, Louis, Putseys, Yvan and Vorlat, Emma. (1982). *The Scene of linguistic action and its perspectivization by speak, talk, say and tell*, Amsterdam, Benjamins.
- Dougherty, Kevin A. and Strassel, Stephanie. (1998). *A new look at variation in and perception of American English quotatives*. Paper presented at NWAVE 27, Athens, Georgia.
- Downes, William. (1984). *Language and Society*. London, Fontana.
- Downing, Pamela and Noonan, Michael. (Eds), (1995). *Word Order in Discourse*. Amsterdam, Benjamins.
- Dubois, Betty Lou. (1989). Pseudoquotation in Current English Communication: "Hey, She Didn't Really Say It". *Language in Society* 18 (3): 343-359.
- Ducrot, Oswald. (1984). Esquisse d'une théorie de d'énonciation. *Le Dire et le Dit*. Paris, Minuit: 171-233.
- Du Feu, Veronica. (1996). *Rapanui*. London, Routledge.
- Duncan, Otis D. (1961). A socioeconomic index for all occupations. In A. Reiss (Ed.), *Occupations and social status*. New York, The Free Press.
- Dutkova-Cope, Lida. (2001). Texas Czech: The Language of Texans Who Say They Speak 'A Different Type of Czech'. *Southwest Journal of Linguistics* 20 (1): 29-69.
- Eckert, Penelope. (1988). Adolescent social structure and the spread of linguistic change. *Language in Society* 17: 183-208.
- Eckert, Penelope. (1989). The whole woman: sex and gender differences in variation. *Language Variation and Change* 1: 245-67.

- Eckert, Penelope. (1998). *Vowels and nail polish: The emergence of linguistic style in the preadolescent heterosexual marketplace*. Proceedings of the 1996 Berkeley Women and Language Conference. Berkeley: Berkeley Women and Language Group.
- Eckert, Penelope. (2000). *Linguistic Variation as Social Practice: The Linguistic Construction of Social Meaning in Belten High*. Oxford, Blackwell.
- Eckert, Penelope. (2003). Language in Gender and Adolescence. In J. Holmes & M. Meyerhoff (Eds.), *Handbook of Language and Gender*. Oxford, Blackwell: 381-400.
- Eckert, Penelope and McConnell-Ginet, Sally. (1999). New Generalizations and Explanations in Language and Gender Research. *Language in Society* 28 (2): 185-201.
- Eckert, Penelope and McConnell-Ginet, Sally. (2003). *Language and Gender*. Cambridge, Cambridge University Press.
- Ehrlich, Susan. (1990). *Point of View: A Linguistic Analysis of Literary Style*. New York, Routledge, Chapman & Hall.
- Einstein, Albert, (1938). *The evolution of physics: the growth of ideas from early concepts to relativity and quanta*. Cambridge, Cambridge University Press.
- Enfield, Nick. (2003). Demonstratives in Space and Interaction: Data from Lao Speakers and Implications for Semantic Analysis. *Language* 79 (1): 82-117.
- Erickson, Mats. (1995). A case of grammaticalization in Modern Swedish: The Use of *ba* in Adolescent Speech. *Language Sciences* 17: 19-48.
- Fader, Ayala. (forthcoming). Controlling knowledge and languages: Literacy, bilingualism and gender in a Hasidic community in Brooklyn. *Linguistics and Education*
- Faircloth, Norman. (1992). *Discourse and social change*. Cambridge, Polity Press.
- Fasold, Ralf. (1990). *The sociolinguistics of language*. Cambridge, MA, Blackwell.
- Fauconnier, Gilles. (1985). *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge MA, MIT Press.
- Fauconnier, Gilles and Sweetser, Eve. (1996). Cognitive Links and Domains: Basic Aspects of Mental Space Theory. In G. Fauconnier & E. Sweetser (Eds.), *Spaces, Worlds and Grammar*. Chicago, London, University of Chicago Press: 1-28.
- Feagin, Crawford. (1979). *Variation and Change in Alabama English*. Washington DC., Georgetown University Press.

- Feagin, Crawford. (1980). Woman's Place in Nonstandard Southern Whites: Not So Simple. In R. Shuy & A. Shnukal (Eds.), *Language Use and the Uses of Language*. Washington, Georgetown University Press: 88-97.
- Fenk-Oczlon, Gertraud. (2001). Familiarity, information flow and linguistic flow. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. Amsterdam, Benjamins.
- Ferrara, Kathleen. (1992). The Interactive Achievement of a Sentence: Joint Productions in Therapeutic Discourse. *Discourse Processes* 15: 207-228.
- Ferrara, Kathleen and Bell, Barbara. (1990). *Discourse Variation and verbs of Saying: say, go, be+like*. Paper presented at NWAVE 19, University of Pennsylvania.
- Ferrara, Kathleen and Bell, Barbara. (1995). Sociolinguistic Variation and Discourse Function of Constructed Dialogue Introducers: The Case of be + like. *American Speech* 70 (3): 265-290.
- Fillmore, Charles. (1990). Epistemic Stance and Grammatical Form in English Conditional Sentences. *Papers from the Regional Meetings*. Chicago Linguistic Society 1: 137-162.
- Fillmore, Charles. (1992). Toward a frame-based lexicon: the semantics of risk. In A. Lehrer & E. Kittay (Eds.), *Frame, fields and contrasts*. Hillsdale, NJ, Lawrence Erlbaum Associates: 75-102.
- Fillmore, Charles. (1997). *Lectures on Deixis*. Stanford: Center Study Language and Information.
- Flaubert, Gustave. (1966). *Bouvard et Pécuchet*. Paris, Garnier Flammarion.
- Flaubert, Gustave. (1966). *Le second volume de Bouvard et Pécuchet. Dictionnaire des Idées Reçues*. Paris, Edition Dénoël.
- Fleischman, Suzanne. (1982). *The future in thought and language: diachronic evidence from Romance*. Cambridge, Cambridge University Press.
- Fleischman, Suzanne and Yaguello, Marina. (to appear). *Discourse Markers across Languages. Evidence from English and French*.
- Fludernick, Monika. (1993). *The fictions of Language and the languages of fiction*. London, Routledge.
- Foley, William. (1997). *Anthropological linguistics: An Introduction*. Malden, MA, Blackwell.

- Fonagy, Ivan. (1986). Reported Speech in Hungarian. In F. Coulmas (Ed.), *Direct and Indirect Speech*. Berlin, Mouton de Gruyter: 333-60.
- Fought, Carmen. (1999). A Majority Sound Change in a Minority Community: /u/-fronting in Chicano English. *Journal of Sociolinguistics* 3(1): 5-23.
- Foulkes, Paul and Docherty, Gerard. (1999). Introduction. In P. Foulkes & G. Docherty (Eds.), *Urban voices: Accent studies in the British Isles*. London, Arnold: 1-24.
- Foulkes, Paul and Docherty, Gerard. (2000). Another chapter in the story of /r/: 'labiodental' variants in British English. *Journal of Sociolinguistics* 4: 30-59.
- Fox Barbara and Jespersen, Robert. (1995). A syntactic exploration of repair in English conversation. In P. Davis (Ed.), *Alternative Linguistics. Descriptive and theoretical modes*. Amsterdam, Philadelphia, Benjamins.
- Frajzyngier, Zygmunt. (1991). The *De Dicto* Domain in Language. In E. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization vol. 1*, Amsterdam, Benjamins: 219-251.
- Frajzyngier, Zygmunt. (1996). *Grammaticalization of the complex sentence: A case study in Chadic*. Studies in language companion Series 32. Amsterdam, Benjamins.
- Fuller, Janet. (2000). *Interviewers are, like, people too: Gender and convergence in the use of like in interviews*. Paper presented at NWAVE 29, Michigan State University, East Lansing.
- Gal, Susan. (1980). *Language Shift: Social Determinants of Linguistic Change in Bilingual Austria*. New York, Academic Press.
- Ganzeboom, Harry, Luijkx, Ruud and Tremain, Donald. (1989). Intergenerational class mobility in comparative perspective. *Research in Social Stratification and Mobility* 8: 3-84.
- Gaudio, Rudolf. (1994). Sounding Gay: Pitch Properties in the Speech of Gay and Straight Men. *American Speech* 69 (1): 30-57.
- Genette, Gerard. (1980). *Narrative discourse: An essay in method*. (trans. By I.E. Levin). Ithaca, NY, Cornell University Press.
- Giddens, Anthony. (1997). *Sociology*. Cambridge, Polity Press.
- Givon, Talmy. (1973). The time-axis phenomenon. *Language* 49 (4): 890-925.
- Givon, Talmy. (1975). Serial verbs and syntactic change: Niger-Kongo. In Ch. Li (Ed.), *Word order and word change*. Austin/London, University of Texas Press.

- Givon, Talmy. (1979). *On Understanding Grammar*. New York, San Francisco and London, Academic Press.
- Givon, Talmy. (1984/1990). *Syntax: A functional-typological introduction*. 2 Vols. Amsterdam, Philadelphia, Benjamins.
- Goffman, Erving. (1974). *Frame analysis*. Cambridge, MA, Harvard University Press.
- Goffman, Erving. (1978). Response Cries. *Language* 54 (4): 787-815.
- Goffman, Erving. (1979). Footing. *Semiotica*. 25 (1-2): 1-29.
- Goffman, Erving. (1981). *Forms of Talk*. Philadelphia, University of Pennsylvania Press.
- Golato, Andrea. (2000). An innovative German quotative for reporting on embodied actions. Und ich so/ und er so 'and I'm like / and he's like'. *Journal of Pragmatics* 32: 29-54.
- Golato, Andrea. (2003). Self-quotation in German: Reporting on past decisions. In T. Gueldemann & M. von Roncador (Eds.), *Reported Discourse: A meeting ground for different linguistic domains*. Amsterdam, Benjamins: 49-70.
- Grice, H. P. (1975). Logic and Conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics*, Vol. 3, New York, Academic Press, pp. 41-58.
- Gruyner, Helena and Tavares de Macedo, Alizia. (1981). La prononciation du *s* post-vocalique: deux processus de changement linguistique en portugais. In D. Sankoff & H. Cedergren (Eds.), *Variation Omnibus*. Alberta, Linguistic Research: 135-40.
- Güldemann, Thomas. (2001). *Quotative Constructions in African Languages: a Synchronic and Diachronic Survey*. Unpublished 'Habilitationsschrift', University of Leipzig.
- Güldemann, Thomas. (2002). When "say" is not say: the functional universality of the Bantu quotative marker *ti* with special reference to Shona. In T. Güldemann & M. von Roncador (Eds.), *Reported discourse: A meeting ground for different linguistic domains*. Amsterdam, Benjamins: 253-288.
- Günthner, Susanne. (1998). Polyphony and the Layering of Voices in Reported Dialogues. An Analysis of the Use of Prosodic Devices in Everyday Reported Speech. *InLiSt - Interaction and Linguistic Structures* 3.
- Günthner, Susanne. (2001). Constructing scenic moments: grammatical and rhetoric-stylistic devices for staging past events in everyday narratives. *InLiSt - Interaction and Linguistic Structures* 22.



- Gumperz, John. (1989). Contextualization Cues and Metapragmatics: The Retrieval of Cultural Knowledge. *Papers from the Regional Meetings*. Chicago Linguistic Society 25 (2): 77-88.
- Gumperz, John. (1992). Contextualization Revisited. In P. Auer & A. di Luzio. (Eds.), *The Contextualization of Language*. Amsterdam, Benjamins: 39-53.
- Gumperz, John. (1997). On the Interactional Bases of Speech Community Membership. In G. Guy, C. Feagin, D. Schiffrin & J. Baugh (Eds.), *Towards a social science of language: Papers in Honor of William Labov, Volume 2: Social Interaction and Discourse Structure*. Amsterdam, Benjamins: 183-203.
- Gundel, Jeanette, Hedberg, Nancy and Zacharshi, Ron. (1993). Cognitive status and the form of referring expressions in discourse. *Language* 69(2): 274-307.
- Guy, Gregory. (1988). Advanced varbrul analysis. In K. Ferrara, B. Brown, K. Walters & J. Baugh (Eds.), *Linguistic change and contact*. Department of Linguistics, University of Texas in Austin. Austin, Texas: 124-136.
- Guy, Gregory. (1990). The sociolinguistic types of language change. *Diachronica* 7: 47-67.
- Guy, Gregory. (1993). The quantitative analysis of linguistic variation. In D. Preston (Ed.), *American Dialect Research*. Amsterdam, Benjamins: 223-249.
- Guy, Gregory, Horvath, Barbara, Vonwiller, Julia, Daisley, Elaine and Rogers, Inge. (1986). An Introductory Change in Progress in Australian English. *Language in Society* 15: 23-52.
- Great Britain. Office of Population Censuses and Surveys. (1995). *1991 Census: General report*. London, HMSO.
- Haberland, Hartmut. (1986). Reported speech in Danish. In F. Coulmas (Ed.), *Direct and Indirect Speech*. Berlin, Mouton de Gruyter: 219-253.
- Haeri, Nilofaar. (1987). Male/female differences in speech: an alternative interpretation. In K. Denning et al. (Eds.), *Variation in Language: N.W.A.V.E.-XV at Stanford*. Stanford, Department of Linguistics: 173-82.
- Haeri, Niloofar. (1994). A Linguistic Innovation of Women in Cairo. *Language Variation and Change* 6 (1): 87-112.
- Haiman, John. (1989). Alienation in Grammar. *Studies in Language* 13: 129-70.
- Haiman, John. (1991). Motivation, repetition, and emancipation: the bureaucratization of language. In H.C. Wolfart. (Ed.), *Linguistic studies presented to John L. Finlay*. Winnipeg, Manitoba, Algonquian and Iroquoian Linguistics, Algonquian and Iroquoian Linguistics Memoir 8.

- Haiman, John. (1994). Ritualization and the development of language. In W. Pagliuca (Ed.), *Perspectives on Grammaticalization*. Amsterdam, Rutworth: 3-28.
- Hajicova, Eva. (1996). Topic-Focus Articulation – A matter of language or parole? The Case of Negation. In R. Sackmann (Ed.), *Theoretical Linguistics and Grammatical Description. Papers in Honour of Hans-Heinrich Lieb on the occasion of his 60<sup>th</sup> Birthday*. Amsterdam, Benjamins.
- Harris, Marvin. (1964). *The Nature of Cultural Things*. New York, Random House.
- Harris, Marvin. (1968). *The Rise of Anthropological Theory*. New York, Harper & Row.
- Haspelmath, Martin. (1999). Why Is Grammaticalization Irreversible? *Linguistics* 37, 6(364): 1043-1068.
- Heeschen, Volker. (1978). The metalinguistic vocabulary of a speech community in the highlands of Irian Jaya (West New Guinea). In A. Sinclair, R. Jervalla & W. Levelt (Eds.), *The child's conception of language*. Berlin, Heidelberg and New York, Springer: 155-187.
- Heine, Bernd. (1992). Grammaticalization chains. *Studies in Linguistics* 16 (2): 335-368.
- Heine, Bernd, Claudi, Ulrike and Hünemeyer, Friedericke. (1991). *Grammaticalization: A conceptual framework*. Chicago: The University of Chicago Press.
- Heine, Bernd and Reh, Mechthild. (1984). *Grammaticalization and Reanalysis in African Languages*. Hamburg, Buske Verlag.
- Heine, Bernd. (1994). Grammaticalization as an explanatory Parameter. In W. Pagliuca (Ed.), *Perspectives in Grammaticalization*. Current Issues in Linguistic Theory 109. Amsterdam, Benjamins.
- Heritage, John. (1984). A change-of-state token and aspects of its sequential placement. In J. Maxwell Aitkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis*. Cambridge, Cambridge University Press: 299-345.
- Hibiya, Junko. (1988). *Social stratification in Tokyo Japanese*. University of Pennsylvania Dissertation.
- Hockett, Charles. (1950). Age-grading and linguistic continuity. *Language* 26: 449-57.
- Holmes, Janet. (1984). Hedging your bets and sitting on a fence: some evidence for hedges as structures. *Te Reo* 27: 47-62.

- Holmes, Janet. (1986). Functions of *you know* in women's and men's speech. *Language in Society* 15 (1): 1-21.
- Holmes, Janet. (1990). Hedges and Boosters in Women's and Men's Speech. *Language and Communication* 10 (3): 185-205.
- Holmes, Janet. (1992). *An Introduction to sociolinguistics*. Essex, Longman.
- Holmes, Janet and Meyerhoff, Miriam. (2003). Introduction. In J. Holmes & M. Meyerhoff (Eds.), *Handbook of Language and Gender*. Oxford, Blackwell: 1-17.
- Holmquist, Jonathan. (1988). *Language Loyalty and Linguistic Variation: A Study in Spanish Catambria*. Dordrecht, Foris.
- Hopper, Paul. (1987). Emergent Grammar. In J. Aske et al. (Eds.), *General Session and Parasession on Grammar and Cognition*. Berkeley Linguistics Society.
- Hopper, Paul. (1991). On some principles of grammaticalization. In E. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*. Amsterdam and Philadelphia, Benjamins: 17-35.
- Hopper, Paul. (1998). Emergent Grammar. In M. Tomasello (Ed.), *The new Psychology of Language. Cognitive and functional approaches to Language structure*. New Jersey and London, Lawrence Erlbaum Associate Publishers.
- Hopper, Paul. (2000). *Prototype or family resemblance? Grammatical constructions and their discourse origins*. University of Essen, LAUD Paper 508.
- Hopper, Paul. (2001). Hendiadys and auxiliation in English. In J. Bybee & M. Noonan (Eds.), *Complex sentences in grammar and discourse*. Amsterdam, Benjamins: 145-174.
- Hopper, Paul and Bybee, Joan. (2001). Introduction. In P. Hopper & J. Bybee (Eds.), *Frequency and the emergence of linguistic structure*. Amsterdam, Benjamins.
- Hopper, Paul and Traugott, Elizabeth. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- Hopper, Paul and Traugott, Elizabeth. (2003). *Grammaticalization*. 2<sup>nd</sup> edition. Cambridge: Cambridge University Press.
- Horn, Laurence. (1984). Towards a New Taxonomy for Pragmatic Inference: Q-based and R-based Implicature. In D. Schiffrin (Ed.), *Meaning, Form and Use in Context*. Washington, Georgetown University Press: 11-42.
- Horvath, Barbara. (1985). *Variation in Australian English: The Sociolects of Sydney*. Cambridge, Cambridge University Press.

- House, Juliane and Kaspar, Gabriele. (1981). Politeness markers in English and German. In F. Coulmas (Ed.), *Conversational Routine. Explorations in Standardized Communications Situations and Prepatterned Speech*. The Hague, Mouton de Gruyter: 157-185.
- Hudson, Richard. (1985). The Limits of Subcategorization. *Linguistic Analysis* 15 (4): 233-256.
- Hudson, Richard. (1996). *Sociolinguistics*. Cambridge, Cambridge University Press.
- Hutchby, Ian and Wooffitt, Robin. (1998). *Conversation analysis: Principles, practices, and applications*. Cambridge, Polity Press.
- Hymes, Dell. (1975). Breakthrough into Performance. In D. Ben-Amos & K. Goldstein (Eds.), *Folklore: Performance and Communication*. The Hague, Mouton de Gruyter: 11-74.
- Hymes, Dell. (1987). A Pattern of Verbal Irony in Chinookan. *International Journal of the Sociology of Language* 65: 97-110.
- Igoe, Matthew, Lamb, Nell, Gilman, John and Kim, Ron. (1999). *The further grammaticalization of be like and some observations on be all*. Paper given at NWAVE 28. University of Toronto.
- Irvine, Judith. (2001). "Style" and distinctiveness: The culture and ideology of linguistic differentiation. In P. Eckert & J. Rickford (Eds.), *Stylistic Variation in Language*. Cambridge, Cambridge University Press: 21-43.
- Jackendoff, Ray. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA, MIT Press.
- Jackendoff, Ray. (1983). *Semantics and Cognition*. Cambridge, MA, MIT Press.
- Jakobson, Roman. (1971). Shifters, Verbal Categories, and the Russian Verb. *Roman Jakobson: Selected Writings II. Word and Language*. The Hague, Mouton de Gruyter: 130-47.
- Janda, Richard. (2001). Beyond "pathways" and "unidirectionality" on the discontinuity of language transmission and the counterability of grammaticalization. *Language Sciences* 23: 265-340.
- Jeffers, Robert and Zwicky, Arnold. (1980). The evolution of Clitics. In E. Traugott et al. (Eds.), *Papers from the fourth International Conference on Historical Linguistics*. Amsterdam, Benjamins.
- Johnson, Annie. (2001). And he was like, "I got to go and move the pig": *Development of BE LIKE as a constructed dialogue introducer*. Unpublished term paper. University of York.

- Johnson, Ellen. (1993). The relationship between lexical variation and lexical change. *Language Variation and Change* 5(3): 285-303.
- Johnson-Laird, Philip. (1975). Models of deduction. In R. Falmagne (Ed.), *Reasoning: Representation and Process*. Springdale, NJ, Erlbaum.
- Johnson-Laird, Philip. (1986). Conditionals and mental models. In C. Ferguson, J. Reilly, A. ter Meulen & E. Traugott (Eds.), *On Conditionals*. Cambridge, Cambridge University Press.
- Johnstone, Barbara. (1987). *He says... so I said*": verb tense alternation and narrative depictions of authority in American English. *Linguistics* 25: 33-52.
- Johnstone, Barabara. (1988). *Gender and Power in Midwestern Personal Storytelling*. Paper presented at Discourses of Power conference, Tempe, Arizona.
- Johnstone, Barbara. (1990). *Stories, Community and Place: Narratives from Middle America*. Bloomington, Indiana University Press.
- Johnstone, Barbara. (1991). Individual style in an American public opinion survey: Personal performance and the ideology of referentiality. *Language in Society* 20: 557-76.
- Johnstone, Barbara. (1993). Community and Conquest Midwestern Men and Women creating their Worlds in conversational Storytelling. In D. Tannen (Ed.), *Gender and Conversational Interaction*. Oxford, Oxford University Press: 62-80.
- Joseph, Brian. (2001). Is There Such a Thing as "Grammaticalization"? *Language Sciences* 23 (2-3): 163-186.
- Juillard, Michel. (2000). Did the Huns Ride a Horse into the Library? Or the Freedoms of the Free Indirect Style. *Cahiers Chronos* 5: 71-90.
- Kahr, Joan Casper. (1976). Adpositions and Locational: Typology and Diachronic Development. *Stanford University Working Papers on Language Universals* 19: 21-54.
- Kammerzell, Frank and Peust, Carsten. (2002). Reported speech in Egyptian: forms, types, and history. In T. Güldemann & M. von Roncador (Eds.), *Reported speech as a meeting ground for different linguistics domains*. Amsterdam, Benjamins: 289-322.
- Kang, Agnes. (1998). Triadic Participation in Organizational Meeting Interaction. *Issues in Applied Linguistics* 9 (2): 139-150.
- Karkkainen, Elise Orvokki. (1999). The Marking and Interactional Functions of Epitemistic Stance in American English Conversational Discourse. *Dissertation Abstracts International, A: The Humanities and Social Sciences* 59 (10), 3802-A.

- Katz, Elihu and Lazarsfeld, Paul. (1955). *Personal Influence*. Glencoe, Ill., Free Press.
- Kendon, Adam. (1980). Gesticulation and speech: two aspects of the process of utterance. In R. Key (Ed.), *The relationship of verbal and nonverbal communication. Contributions to the sociology of language*. The Hague, Mouton de Gruyter: 207-227.
- Kendon, Adam. (1992). Some recent work from Italy on 'Quotable gestures'. *Journal of Linguistic Anthropology* 2: 92-108.
- Kendon, Adam. (1994). Review Article: Faces, Gestures, Synchrony,... *Semiotica* 102 (3-4): 311-322.
- Kendon, Adam. (1996). Gesture. *Annual Review of Anthropology* 26: 109-128.
- Kerswill, Paul. (2003). *The levelling of English in Britain: some new theoretical perspectives*. Paper presented at the 4<sup>th</sup> UK Language Variation and Change Conference, Sheffield.
- Kerswill, Paul and Williams, Ann. (2002). Dialect Recognition and Speech Community Focusing in Old and New Towns in England: The Effects of Dialect Levelling, Demography and Social Networks. In D. Preston (Ed.), *Handbook of Perceptual Dialectology*. Vol 2. Amsterdam, Benjamins: 173-205.
- Klamer, Miriam. (2003). "Report" constructions in Kambera (Austronesian). In T. Güldemann & M. von Roncador (Eds.), *Reported speech as a meeting ground for different linguistics domains*. Amsterdam, Benjamins: 323-340.
- Klewitz, Gabriele and Couper-Kuhlen, Elizabeth. (1999). Quote-Unquote? The Role of Prosody in the Contextualization of Reported Speech Sequences. *Journal of Pragmatics* 4: 459-485.
- Kretzschmar, William. (2003). Mapping Southern English. *American Speech* 78 (2): 130-149.
- Kroch, Anthony. (1989). Function and Grammar in the history of English periphrastic "do". In R. Fasold & D. Schiffin (Eds.), *Language Variation and Change*. Orlando, Harcourt Brace Jovanovich: 132-177.
- Kroon, Caroline. (1995). *On causal and adversative Discourse markers in Latin, with a survey of the existing literature*. Amsterdam, Gieben.
- Krug, Manfred. (2001). Frequency, iconicity, categorization: Evidence from emerging modals. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. Amsterdam, Benjamins: 328-356.
- Kuno, Susumo. (1980). *The scope of Question and Negation in some Verb-final Languages. Papers from Regional Meetings*, Chicago Linguistics Society, Chicago, University of Chicago Linguistics Department 16:155-169.

- Kuryłowicz, Jerzy. (1964). On the methods of internal reconstruction. In H.G. Lundt (Ed.), *Proceedings of the Ninth International Congress in Linguistics*. The Hague, Mouton de Gruyter.
- Kuryłowicz, Jerzy. (1965[1975]). *The Evolution of grammatical categories*. *Diogenes* 51(1965), 55-71. Reprinted in *Esquisses linguistiques*, volume II. Fink Munich, 1975.
- Kvavik, Karen H. (1986). Characteristics of Direct and Reported Speech Prosody: Evidence from Spanish. In F. Coulmas (Ed.), *Direct and Indirect Speech*. Berlin, Mouton de Gruyter: 333-60.
- Labov, William. (1963). The social motivation of a sound change. *Word* 19: 273-309.
- Labov, William. (1966a). The linguistic variable as a structural unit. *Washington Linguistics Review* 3: 4-22.
- Labov, William. (1966b). *The Social Stratification of English in New York City*. Washington DC. Center for Applied Linguistics.
- Labov, William. (1969). Contraction, deletion, and inherent variability of the English copula. *Language* 45: 715-762.
- Labov, William. (1972a). *Sociolinguistic Patterns*. University of Pennsylvania Press.
- Labov, William. (1972b). Negative attraction and negative concord in English grammar. *Language* 48: 773-818.
- Labov, William. (1972c). *The Transformation of Experience in Narrative Syntax. Language in the inner City*. Philadelphia, University of Pennsylvania Press: 354-97.
- Labov, William. (1972d). *Language in the inner city: Studies in Black vernacular English*. Philadelphia, University of Philadelphia Press.
- Labov, William. (1978). Where does the linguistic variable stop? A response to Beatrice Lavandera. *University of Pennsylvania Sociolinguistic Working Paper 44*.
- Labov, William. (1984). At the intersection of sex and social factors in the course of language change. Paper presented at NWAVE 10, University of Philadelphia
- Labov, William. (1989). The Child as Linguistic Historian. *Language Variation and Change* 1 (1): 85-97.
- Labov, William. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2: 205-54.

- Labov, William. (1994). *Principles of Linguistic Change*. Volume I, Internal Factors. Malden, Cambridge, Blackwell.
- Labov, William. (1996). When intuitions fail. In L. McNair, K. Singer, L. Dorbin & M. Aucoin (Eds.), *CLS 32: Papers from the parasession on theory and data in linguistics*. Chicago, Chicago Linguistics Society: 77-105.
- Labov, William. (2001). *Principles of Linguistic Change*. Volume II, Social Factors. Malden, Cambridge, Blackwell.
- Labov, William. (2002). Driving Forces in Linguistic Change. Paper presented at the International Conference on Korean Linguistics, Seoul National University.
- Labov, William, Cohen, Paul, Robbins, Clarence and Lewis, John. (1968). *A study of the non-standard English of Negro and Puerto Rican speakers in New York City*. Cooperative research Grant 3288. New York, Columbia University.
- Labov, William and Waletzky, Joshua. (1967). Narrative analysis. In J. Helm (Ed.), *Essays on the Verbal and Visual Arts*. Seattle, University of Washington Press: 12-44.
- Ladegaard, Hans L. (1998). Boys, girls, language and identity: Language attitudes and linguistic behaviour in a rural community in Denmark. *International Journal of Applied Linguistics* 8: 3-25.
- Ladegaard, Hans J. (2000). Language attitudes and sociolinguistic behaviour: Exploring attitude-behaviour relations in language. *Journal of Sociolinguistics* 4 (2): 214-233.
- Lakoff, George (1972). *Hedges: A study in meaning criteria and the logic of fuzzy concepts*. Eighth Regional Meeting of the Chicago Linguistic Society. Chicago, University of Chicago Press: 183-228.
- Lakoff, George. (1987). *Women, Fire, and Dangerous Things*. Chicago, University of Chicago Press.
- Lakoff, George, and Johnson, Mark. (1980). The metaphorical Structure of the Human Conceptual System. *Cognitive Science* 4(2): 195-208.
- Lakoff, Robin. (1973). *Language and Woman's Place*. New York, Harper.
- Lambrecht, Knud. (1994). *Information structure and sentence for: Topic, focus, and the mental representations of discourse referents*. Cambridge, Cambridge University Press.
- Lambrecht, Knud and Polinsky, Maria. (1997). Typological Variation and Sentence Focus Constructions. Chicago, Chicago Linguistics Society 33. Papers from the panels on Linguistic Ideology: 189-206.



- Lance, Donald. (1999). Regional Variation in Subjective Dialect Divisions in the United States. In D. Preston (Ed.), *Handbook of Perceptual Dialectology*. Amsterdam, Benjamins.
- Langacker, Richard. (1985). *Foundations of Cognitive grammar, Vol. 1: Theoretical prerequisites*. Stanford, Stanford University Press.
- Lange, Deborah. (1986). *Attitudes of teenagers towards sex-marked language*. Unpublished manuscript, Georgetown University, Department of Linguistics.
- Lass, Roger. (2000). Remarks on (Uni) Directionality. In O. Fischer, A. Rosenbach & D. Stein (Eds.), *Pathways of Change*. Amsterdam, Benjamins: 207-227.
- Lavandera, Beatrice. (1978). Where does the sociolinguistic variable stop? *Language and Society* 7: 171-182.
- Leap, William. (1996). *Word's out: Gay men's English*. Minneapolis, MN., University of Minnesota Press.
- Leech, Geoffrey. (1974). *Semantics*. Harmondsworth, Penguin.
- Leech, Geoffrey and Short, Michael. (1981). *Style in fiction: A linguistic introduction to English fictional prose*. London, Longman.
- Lehrer, Adrienne. (1989). Remembering and representing Prose: Quoted speech as a data source. *Discourse Processes* 12: 105-125.
- Lehmann, Christian. ([1982]/1995). *Thoughts on Grammaticalization. A Programmatic Sketch. Vol. I*. Köln: Universität zu Köln, Institut für Sprachwissenschaft [Arbeiten des Kölner Universalien-Projekts 48].
- Lehmann, Christian. (1985). Grammaticalization: Synchronic Variation and Diachronic Change. *Lingua e Stile* 20 (3): 303-318.
- Lehmann, Christian. (1991). Grammaticalization in Modern German. In E. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*. Amsterdam, Benjamins.
- Lehmann, Christian. (1988). Towards a typology of clause linkage. In J. Haiman & S. Thompson (Eds.), *Clause Combining in Grammar and Discourse*. Amsterdam, Benjamins: 181-227.
- Lenneberg, Eric. (1967). *Biological foundations of language*. New York, Wiley and Sons.
- Le Page, Robert and Tabouret-Keller, André. (1985). *Acts of identity: Creole-based approaches to Language and Identity*. Cambridge, Cambridge University Press.
- Levey, Stephen. (2003). Reported dialogue and pragmatic particles in the narratives of preadolescents. *World Englishes* 22(3): 305-321.

- Levinson, Stephen. (1983). *Pragmatics*. Cambridge, Cambridge University Press.
- Li, Charles. (1986). Direct Speech and indirect speech: A functional study. In F. Coulmas (Ed.), *Direct and Indirect speech*. Berlin, Mouton de Gruyter: 29-45.
- Lichtenberk, Frantisek. (1991). Semantic Change and Heterosemy in Grammaticalization. *Language* 67 (3): 475-509.
- Lindblom, Bjorn. (1998). Systemic Constraints and Adaptive Change in the Formation of Sound Structure. In J. Hurford, M. Studdert-Kennedy & C. Knight (Eds.), *Approaches to the Evolution of Language: Social and cognitive bases*. Cambridge University Press: 242-264.
- Lindstrom, Therese. (2001). *Lexicalised Grammaticalisation?*. Paper Presented at The History of Linguistics in Texts and Concepts. Universität Potsdam.
- Lindstrom, Therese. (2002). *Unidirectionality – myth, truth, or tautology?* Paper presented at the New Reflections on Grammaticalization 2 Conference, University of Amsterdam.
- Lippi-Green, Rosina. (1989). Social Network Integration and Language Change in Progress in a Rural Alpine Village. *Language in Society* 18 (2): 213-234.
- Longacre, Robert. (1985). Sentences as combinations of clauses. In T. Shopen (Ed.), *Language Typology and syntactic description*, Vol.3. Cambridge, Cambridge University Press: 235-286.
- Longacre, Robert. (1994). The Dynamics of Reported Dialogue in Narrative. *Word* 45 (2): 125-143.
- Lord, Carol. (1976). *Evidence for syntactic reanalysis: From verb to complementizer in Kwa*. Chicago Linguistic Society. Papers from the Parasession on Diachronic Syntax: 179-191.
- Lord, Carol. (1993). *Historical Change and Serial Verb Constructions*. Amsterdam, Benjamins.
- Lucy, John. (1993). *Reflexive language: reported speech and metapragmatics*. Cambridge, Cambridge University Press.
- Macaulay, Ronald. (1976). Social class and Language in Glasgow. *Language in Society* 5: 173-88.
- Macaulay, Ronald. (1977). *Language, Social Class, and Education: A Glasgow study*. Edinburgh, Edinburgh University Press.
- Macaulay, Ronald. (2001). *You're like 'why not?'*. The quotative expression of Glasgow adolescents. *Journal of Sociolinguistics* 5(1): 3-21.

- Macaulay, Ronald. (2002). *ADS Annual Lecture. I'm Off to Philadelphia in the Morning: A Scotsman Looks at Dialect in America*. *American Speech* 77 (3): 227-241.
- Mackeigan, Terri. (in progress). *The impact of artefacts on colour concepts: A network study of new knowledge*. PhD dissertation. University of Edinburgh.
- MacLaury, Robert. (1998). *Linguistic relativity and plasticity of categorization: universalism in a new key*. Paper read to the LAUD Symposium on Linguistic Relativity. University of Duisburg.
- MacWhinney, Brian. (2001). Emergentist approaches to language. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. Amsterdam, Benjamins.
- Maling, Joan. (2002). *From Passive to Active: A Syntactic Change in Progress*. Germanic Linguistics Annual Conference (GLAC). Indiana University, Bloomington.
- Malnando, Concepción. (1991). *Discurso directo y discurso indirecto*. Madrid, Taurus Universitaria.
- Mann, William and Thompson, Sandra. (1986). Relational propositions in discourse. *Discourse Processes* 9(1): 57-90.
- Mandler, Jean. (1984). *Stories, scripts, and scenes: Aspects of schema theory*. Hillsdale, NJ, Lawrence Erlbaum.
- Markusen, Anne. (1996). Sticky places in slippery space: A typology of industrial districts. *Economic Geography* 72: 293-313.
- Marckwardt, Albert. (1967). Lexical redistribution of Modern English *say* and *tell*. In W. Austin (Ed.), *Papers in linguistics in Honour of Leon Dostert*. The Hague, Mouton de Gruyter: 118-122.
- Marslen-Wilson, William, Levy, Elena, and Tyler, Lorraine Komisarjewski. (1982). Proceeding interpretable discourse: The establishment and maintenance of reference. In R. Jarvella and W. Klein (Eds.), *Speech, place and action: Studies in deixis and related topics*. New York, Wiley: 339-378.
- Martinet, André. (1955). *Economie des changements phonétiques*. Bern, Francke.
- Massamba, Daniel. (1986). Reported Speech in Swahili. In F. Coulmas (Ed.), *Direct and Indirect Speech*. Berlin, Mouton de Gruyter: 99-120.
- Mathisen, Anne Grethe. (1999). Sandwell, West Midlands: Ambiguous perspectives on gender patterns and models of change. In P. Foulkes & G. Docherty (Eds.), *Urban Voices. Accent studies on the British Isles*. London, Arnold: 107-123.

- Mathis, Terri and Yule, George. (1994). Zero quotatives. *Discourse Processes* 18: 63-76.
- Mayes, Patricia. (1990). Quotation in spoken English. *Studies in Language* 14: 325-363.
- Maynard, Senko. (1996). Multivoicedness in speech and thought representation: The case of self-quotation in Japanese. *Journal of Pragmatics* 25: 207-226.
- McGregor, William. (1994). The Grammar of Reported Speech and Thought in Gooniyandi. *Australian Journal of Linguistics* 14 (1): 63-92.
- McQuail, Denis. (2000). *McQuail's Mass Communication Theory*. London, Sage.
- Meehan, Teresa. (1991). It's like, what's happening in the evolution of like?: A theory of grammaticalization. *Kansas Working Papers in Linguistics* 16: 37-51.
- McElhinny, Bonnie. (2003). Theorizing Gender in Sociolinguistics and Linguistic Anthropology. In J. Holmes & M. Meyerhoff (Eds.), *Handbook of Language and Gender*. Oxford, Blackwell: 21-42.
- McMahon, April. (1994). *Understanding Language Change*. Cambridge, Cambridge University Press.
- McNeill, David. (1985). So you think gestures are non-verbal? *Psychological Review* 92(3): 350-371.
- McNeill, David. (1992). *Hand and Mind: What Gestures Reveal About Thought*. Chicago, University of Chicago Press.
- Meillet, Antoine. (1921). *Linguistique Historique et Linguistique générale, tome I*. Paris, Librairie Ancienne Honoré Champion.
- Meillet, Antoine. (1929/1936). *Linguistique historique et linguistique générale, tome II*. Klincksieck, Paris.
- Mendoza-Denton, Norma. (forthcoming) *Homegirls: Symbolic Practices in the Making of Latina Youth Styles*. London, Blackwell.
- The Merriam-Webster online. Merriam-Webster, Inc. (<http://www.m-w.com/home.htm>)
- Meyerhoff, Miriam. (1992). „We've all got to go one day, eh?": Powerlessness and solidarity in the functions of a New Zealand tag. In K. Hall et al. (Eds.), *Locating power: Proceedings of the Second Berkeley Women and Language Conference*. Berkeley, Berkeley Women and Language Group: 409-419.

- Meyerhoff, Miriam. (1994). Sounds Pretty Ethnic, Eh?: A Pragmatic Particle in New Zealand English. *Language in Society* 23(3): 367-388.
- Meyerhoff, Miriam. (1999). *Towards a typology of linguistic variables*. Paper presented at NWAVE 28, University of Toronto.
- Meyerhoff, Miriam. (2000). *Constraints on Null Subjects in Bislama (Vanuatu): Social and linguistic factors*. Canberra, Pacific Linguistics.
- 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. Typological studies in Language*. Amsterdam, Philadelphia, Benjamins.
- Meyerhoff, Miriam and Niedzielski, Nancy. (1995). *On the discourse functions of olsem in Bislama*. Paper presented at NWAVE 25. University of Pennsylvania.
- Meyerhoff, Miriam and Niedzielski, Nancy. (1998). The syntax and semantics of olsem in Bislama. In M. Pearson (Ed.), *Recent Papers in Austronesian Linguistics*. Los Angeles, UCLA Occasional Papers in Linguistics.
- Meyerhoff, Miriam and Niedzielski, Nancy. (2002). *Standards, the Media, and Language Change*. Paper presented at NWAVE 31, Stanford University.
- Meyerhoff, Miriam and Niedzielski, Nancy. (2003). The globalization of vernacular variation. *Journal of Sociolinguistics*. 7(4): 534-555.
- Miller, Jim and Weinert, Regina. (1995). The Function of *like* in dialogue. *Journal of Pragmatics* 23: 365-393.
- Milroy, James. (1980). Lexical alternation and the history of English: evidence from the urban vernacular. In E. Traugott, R. Labrum & S. Shepherd (Eds.), *Papers from the 4<sup>th</sup> International Conference on Historical Linguistics*. Amsterdam, Benjamins.
- Milroy, James. (1981). *Regional Accents of British English*. Belfast, Blackstaff.
- Milroy, James and Milroy, Lesley. (1978). Belfast: Change and variation in an urban vernacular. In P. Trudgill (Ed.), *Sociolinguistic Patterns in British English*. London, Arnold: 19-36.
- Milroy, James and Milroy, Lesley. (1985). Linguistic Change, Social Network and Speaker Innovation. *Journal of Linguistics* 21 (2): 339-384.
- Milroy, James and Milroy, Lesley. (1997). Network Structure and Linguistic Change. In N. Coupland and A. Jaworski (Eds). *Sociolinguistics: A reader*. New York, St. Martin's Press: 199-211.

- Milroy, James and Milroy, Lesley. (1998). Mechanisms of Change in Urban Dialects: The Role of Class, Social Network and Gender. In P. Trudgill & J. Cheshire (Eds.), *The Sociolinguistic Reader. Vol.1: Multilingualism and Variation*. London, Arnold: 179-195.
- Milroy, James, Milroy, Lesley and Hartley, Sue. (1994). Local and supra-local change in British English: The case of glottalisation. *English World-wide* 15: 1-33.
- Milroy, Lesley. (1976). Phonological correlates to community structure in Belfast. *Belfast Working Papers on Language and Linguistics* 1: 1-44.
- Milroy, Lesley. (1980). *Language and Social Networks*. Oxford, Blackwell.
- Milroy, Lesley. (1987). *Observing and Analysing Natural Language: A Critical Account of Sociolinguistic Method*. Oxford, Blackwell.
- Milroy, Lesley. (1999). Women as innovators and norm-creators: The sociolinguistics of dialect leveling in a northern English city. *Proceedings of the Berkeley Conference on Women and Language*.
- Milroy, Lesley and Milroy, James. (1992). Social Network and Social Class: Toward an Integrated Sociolinguistic Model. *Language in Society* 21: 1-26.
- Milroy, Lesley, Milroy, James and Docherty, Gerard. (1997). *Phonological Variation and Change in Contemporary Spoken British English*. Final Report to the ESRC, R00 234892.
- Mougeon, Raymond, Beniak Edouard, and Valli, André. (1988). VAIS, VAS, M'A in Canadian French: a sociohistorical study. In K. Ferrara et al. (Eds.), *Linguistic Change and Contact: NWAV-XVI*. Austin, Department of Linguistics, University of Texas: 250-62.
- Morales, Humberto Lopez. (1986). Velarization of /-N/ in Puerto Rican Spanish. In D. Sankoff (Ed.), *Diversity and Diachrony*. Philadelphia, Benjamins: 105-13.
- Mosegaard-Hansen, Maj-Britt. (1995). Marqueurs métadiscursifs en français parlé: l'exemple de bon et de ben. *Le Français moderne* 63(1): 20-41.
- Mosegaard-Hansen, Maj-Britt. (1997). Alors and donc in spoken French: A reanalysis. *Journal of Pragmatics* 28:153-187.
- Mosegaard-Hansen, Maj-Britt. (1998). *The function of discourse particles: a study with special reference to spoken standard French*. Amsterdam, Benjamins.
- Munro, Pamela. (1982). On the transitivity or say-Verbs. In P. Hopper & S. Thompson (Eds.), *Studies in Transitivity. Syntax and Semantics*. New York, San Francisco, London, Academic Press: 301-318.

- Myers, Greg. (1999). Unspoken Speech: Hypothetical Reported Discourse and the Rhetoric of Everyday Talk. *Text* 19 (4): 571-590.
- Nagy, Naomi and Blondeau, Hélène. (1998). Subject Pronoun Variation in Montreal French as L2. Paper presented at NWAVE 27, Athens, GA.
- Nagy, Naomi and Blondeau, Hélène (2000). Double subject marking in L2 Montreal French. *University of Pennsylvania working papers in Linguistics* 6 (3).
- Naden, Tony. (1989). Gur. In J. Bendor-Samuel (Ed.), *The Niger-Congo Languages*. Lanham, University of South Africa: 136-168.
- Nakao, Keiko and Treas, Judith. (1990). *Computing 1989 Occupational Prestige Scores. GSS Methodological Report No. 60*. Chicago, National Opinion Research Council.
- Nichols, Patricia. (1976). Black women in the rural south: Conservative and innovative. In B. Dubois & I. Crouch (Eds.), *Proceedings of the Conference on the Sociology of the Languages of American Women*. San Antonio, Texas, Trinity University.
- Nicolle, Steve. (1997). A relevance-theoretic account of *be going to*. *Journal of Linguistics* 33: 355-377.
- Noble, Grant. (1975). *Children in Front of the Small Screen*. London, Constable.
- Nolde, Muriel. (2001). Deflexion as a counterdirectional factor in grammatical change. *Language Sciences* 23: 231-264.
- Nølke, Henning. (1983). *Les adverbes paradigmatiques: fonction et analyse*. Copenhagen, Akademisk Verlag.
- Nordberg, Bengt. (1984). Om ungdomars samstalsstil. Några preliminära iakttagelser. *Svenska Studier* 64: 5-27.
- Nordberg, Bengt and Sundgren, Eva. (1998). On observing language change: A Swedish case study. *FUMS Rapport nr. 190*. Institutionen för norsiska språkvid, Uppsala Universitet.
- Ochs, Elinor. (1997). Narrative. In T. Van Dijk (Ed.), *Discourse as structure and process: Discourse studies: A multidisciplinary introduction. Vol. 1*. London, Sage: 185-207.
- Östman, Jan-Ola. (1995). Pragmatic Particles Twenty Years After. *Anglicana Turkuensia* 14: 95-108.
- Overstreet, Maryann. (2000). *Whales, Candlelight, and Stuff like that: General Extenders in English Discourse*. Oxford, Oxford University Press.

- Ochs, Elinor. (1996). Linguistic resources for socializing humanity. In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity*. Cambridge, Cambridge University Press: 407-437.
- Ochs, Elinor and Capps, Lisa. (1996). Narrating the Self. *Annual Review of Anthropology* 25: 19-43.
- The Oxford English Dictionary Online. Oxford University Press 2003 (<http://dictionary.oed.com>)
- Page, Norman. (1973). *Speech in the English Novel*. London, Longman.
- Parameswaran, Radhika. (2002). Local culture in global media: Excavating colonial and material discourses in National Geographic. *Communication Theory* 12: 287-315.
- Partee, Barbara. (1973). The Syntax and Semantics of quotation. In S. Anderson & P. Kiparsky (Eds.), *A Festschrift for Morris Halle*. New York, Rinehart and Winston: 410-418.
- Pascal, Roy. (1977). *The dual voice: Free indirect speech and its functioning in the nineteenth century European novel*. Manchester, Manchester University Press.
- Payne, Doris. (Ed.), (1992). *Pragmatics of word order flexibility*. Amsterdam, Benjamins.
- Peireira Scherre, Maria Marta and Naro, Anthony. (1991). Marking in Discourse: "Birds of a feather". *Language Variation and Change* 3: 23-32.
- Pereira Scherre, Maria Marta and Naro, Anthony (1992). The serial effect on internal and external variables. *Language Variation and Change* 4: 1-13.
- Pérez, Aveline. (1990). Time in motion: Grammaticalization of the *be going to* construction in English. *La Trobe University Working Papers in Linguistics* 3: 49-64.
- Perrin, Mona. (1974). Direct and indirect Speech in Mambila. *Journal of Linguistics* 10: 27-37.
- Peters, Ann M. (1983). *The Units of language Acquisition*. Cambridge: Cambridge University Press.
- Pike, Kenneth L. (1967). *Language in Relation to a unified Theory of the Structure of Human Behavior*. The Hague, Mouton de Gruyter.
- Pintzuk, Susan. (1986). *VABRUL: Version 2.1*. Department of Linguistics, University of Philadelphia.
- Plato (1998). *The Republic* (translated with an introduction and notes by Robin Waterfield). Oxford, Oxford University Press.



- Polanyi, Livia. (1982). Linguistic and Social Constraints on Storytelling. *Journal of Pragmatics* 6 (5-6): 509-524.
- Poplack, Shana and Tagliamonte, Sali. (1999). The Grammaticization of *going to* in (African American) English. *Language Variation and Change* 11 (3): 315-342.
- Pomerantz, Anita. (1979). Agreeing and Disagreeing with Assessments: Some Features of Preferred/Dispreferred Turn Shapes. In M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversational analysis*. Cambridge, Cambridge University Press: 57-101.
- Precht, Kirsten. (2000). *Patterns of Stance in English*. Unpublished doctoral dissertation. Northern Arizona University, Flagstaff, Arizona.
- Precht, Kirsten. (2003). Stance Moods in Spoken English: Evidentiality and Affect in British and American Conversation. *Text* 23 (2): 239-257.
- Przedlacka, Joanna. (2001). Estuary English and RP: some recent findings. *Studia Anglica Posaniensia* 36: 35-50.
- Preston, Dennis R. (1991). Sorting out the variables in sociolinguistic theory. *American Speech* 66(1): 33-56.
- Pujolar, Joan. (2000). *Gender, Heteroglossia and Power: A Sociolinguistic Study of Youth Culture*. Berlin, New York, Mouton de Gruyter.
- Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey and Svartvik, Jan. (1985). A comprehensive grammar of the English language.
- Rabatel, Alain. (2001). Representations of Inner Speech: Interior Monologue, Free Direct and Indirect Discourse, Point of View. *Langue Francaise* 132: 72-95.
- Random House Webster's Unabridged Dictionary. (1999). NY. Random House, Inc.
- Ransom, Evelyn. (1988). The Grammaticalization of complementizers. *Berkeley Linguistics Society* 14: 364-374.
- Reid, Euan. (1978). Social and stylistic variation in the speech of children: some evidence from Edinburgh. In P. Trudgill (Ed.), *Sociolinguistic patterns in British English*. London, Arnold.
- Reilly, Judy Snitzer. (1983). What Are Conditionals For? *Papers and Reports on Child Language Development* 22: 124-132.
- Reyes, Graciela. (1993). *Los procedimientos de cita: Estilio directo y estilio indirecto*. Madrid, Arco/Libros.

- Rickford, John and McNair-Knox, Faye. (1994). Addressee and Topic-Influenced Style Shift: A Quantitative Sociolinguistic Study. In D. Biber, & E. Finegan (Eds.), *Sociolinguistic Perspectives on Register*. New York, Oxford, Oxford University Press: 235-276.
- Rimmer, Sharon. (1988). *Sociolinguistic Variability in Oral Narrative*. Unpublished Doctoral Dissertation, Aston University.
- Romaine, Suzanne. (1978). Post-vocalic /r/ in Scottish English: sound change in progress? In P. Trudgill (Ed.), *Sociolinguistic patterns in British English*. London, Arnold: 144-158.
- Romaine, Suzanne (1984a). *The language of children and adolescents: The Acquisition of Communicative Competence*. Oxford, Blackwell.
- Romaine, Suzanne. (1984b). On the problem of syntactic variation and pragmatic meaning in sociolinguistic theory. *Folia Linguistica* 18: 409-437.
- Romaine, Suzanne. (2003). Variation in language and gender. In J. Holmes & M. Meyerhoff (Eds.), *The Handbook of Language and Gender*. Oxford, Blackwell: 98-119.
- Romaine, Suzanne and Lange, Deborah. (1991). The use of *like* as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech* 66: 227-279.
- Rooth, Mats. (1996). Focus. In S. Lappin (Ed.), *The Handbook of contemporary semantic theory*. Oxford, Blackwell: 271-297.
- Rosch, Eleanor. (1975). Family Resemblances: studies in the internal resemblance of Categories. *Cognitive Psychology* 8: 382-439.
- Rosch, Eleanor. (1978). Principles of Categorization. In E. Rosch & L. Lloyd (Eds.), *Cognition and Categorization*. New York, Hillsdale: 27-48.
- Royster, Jane F. and Steadman, John M. (1923/1968). The going to future. In *The Manly anniversary studies in language and literature*. Freeport, NY, Books for Libraries Press: 394-403.
- Sacks, Harvey, Schegloff Emanuel, and Jefferson, Gail. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50 (4): 696-735.
- Sanchez, Tara and Charity, Anne. (1999). *The use of like and other verbs of quotation in a predominantly African-American community*. Paper presented at NWAVE 28. Toronto, University of Toronto.

- Sanders, José and Redecker, Gisela. (1996). Perspective and representation of Speech and Thought in Narrative Discourse. In G. Fauconnier & E. Sweetser (Eds.), *Spaces, Worlds, and Grammar*. Chicago, University of Chicago Press.
- Sankoff, Gillian, Thibault, Pierrette, Nagy, Naomi, Blondeau, Helene, Fonollosa, Marie-Odile and Gagnon, Lucie. (1997). Variation in the use of discourse markers in a language contact situation. *Language Variation and Change* 9: 191-217.
- Sankoff, Gillian. (1980a). Above and beyond phonology in variable rules. In G. Sankoff (Ed.), *The Social Life of language*. Philadelphia, University of Pennsylvania Press: 81-93.
- Sankoff, Gillian. (1980b). Variation, pidgins and creoles. In A. Valdman & A. Highfield (Eds.), *Theoretical Orientations in Creole Studies*. New York, Academic Press: 139-164.
- Sankoff, David and Sankoff, Gillian. (1973). Sample survey methods and computer assisted analysis in the study of grammatical variation. In R. Darnell (Ed.), *Canadian Language in its social context*. Edmonton, Linguistic Research Inc: 7-64.
- Sattel, Jack. (1983). Men, inexpressiveness, and power. In B. Thorne, C. Kramarae & N. Henley (Eds.), *Language, Gender and Society*. Rowley, MA, Newbury House: 118-124.
- Saxena, Anju. (1995). Unidirectional Grammaticalization: Diachronic and cross-linguistic evidence. *Sprachtypologie und Universalienforschung* 48 (4): 350-372.
- Schank, Roger C. and Abelson, Robert P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ, Lawrence Erlbaum.
- Schegloff, Emmanuel. (1990). On the Organization of Sequences as a Source of Coherence in Talk in Interaction. In B. Dorval (Ed.), *Conversational Organization and its Development*. Norwood, New York.
- Schegloff, Emmanuel. (1996). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. Schegloff & S. Thompson (Eds.), *Interaction and Grammar*. Cambridge, Cambridge University Press.
- Schiffrin, Deborah. (1981). Tense variation in narrative. *Language* 57: 45-62.
- Schiffrin, Deborah. (1987). *Discourse Markers*. Cambridge, Cambridge University Press.
- Schilling-Estes, Natalie. (2002). American English Social Dialect Variation and Gender. *Journal of English Linguistics* 30 (2): 122-137.

- Schneider, Stefan. (2002). *Pragmatic functions of Spanish parenthetical verbs*. Paper held at the 1st Symposium on intercultural, cognitive and social pragmatics (EPICS 1), Sevilla.
- Schourup, Lawrence. (1982a). Common Discourse Particles in English Conversation. *Working Papers in Linguistics* 28, Ohio State University, Department of Linguistics.
- Schourup, Lawrence. (1982 b). Quoting with go 'say'. *American Speech* 57: 149.
- Schourup, Lawrence (1985). *Common discourse particles in English conversation*. New York, Garland Press.
- Semino, Elena, Short, Mick and Culpeper, Jonathan. (1997). Using a corpus to test a model of speech and thought presentation. *Poetics* 25: 17-43.
- Short, Mick, Semino, Elena and Culpeper, Jonathan. (1996). Using a corpus for stylistics research: Speech and thought presentation. In J. Thomas & M. Short (Eds.), *Using Corpora in Language Research*. London, Longman: 110-131.
- Silva-Corvalán, Carmen. (1981). Extending the sociolinguistic variable to syntax: the case of pleonastic clitics in Spanish. In D. Sankoff and H. Cedergren (Eds.), *Variation Omnibus*. Alberta, Linguistics Research: 335-42.
- Simon-Vandenberg, Anne-Marie. (2000). The Functions of *I Think* in Political Discourse. *International Journal of Applied Linguistics* 10 (1): 41-63.
- Singler, John. (2001). Why you can't do a VABRUL study of quotatives and what such a study can show us. *University of Pennsylvania Working Papers in Linguistics* 7(3): 257-278.
- Singler, John and Woods, Laurie. (2002). *The Use of (Be) Like Quotatives in American and Non-American Newspapers*. Paper presented at NWAVE 32, Stanford University.
- Smith, Jennifer and Steele, Hazel. (2003). *Caregiver, community and child in the acquisition of variable forms in a Scottish dialect*. Paper presented at the Linguistic Circle, University of Edinburgh.
- Sperber, Dan and Wilson, Deidre. (1986). *Relevance: Communication and Cognition*. Cambridge Massachusetts, Harvard University Press.
- Stafford, Laura and Daly, John. (1984). Conversational memory: The effects of recall mode and memory expectancies on remembrances of natural conversation. *Human Communication Research* 10: 379-402.
- Sternberg, Meir. (1982). Proteus in quotation-land: Mimesis and form of reported discourse. *Poetics Today* 3: 107-56.

- Strang, Barbara. (1970). *A History of English*. London, Methuen.
- Stuart-Smith, Jane. (2001). *Should variationists worry about the influence of TV on speech?* Paper presented at the 3<sup>rd</sup> UK Conference on Variation and Change, University of York.
- Steever, Stanford B. (2003). Direct and indirect discourse in Tamil. In T. Güldemann & M. von Roncador (Eds.), *Reported Discourse. A meeting ground for different linguistic domains. Typological studies in Language*. Amsterdam, Philadelphia, Benjamins: 91-108.
- Strasburger, Victor. (1995). *Adolescents and the media: Medical and psychological impact*. London, Sage.
- Streek, Jürgen. (1988). The significance of gesture: How it is established. *International Pragmatics Association, Papers in Pragmatics* 2 (1-2): 60-83.
- Stein, Ellen. (1990). *I'm sittin' there*: Another new quotative? *American Speech* 65: 303.
- Stern, Gustaf. (1931). *Meaning and Change of Meaning*. Reprint by Bloomington, Indiana University Press 1965.
- Sturtevant, Edgar. (1947). *An Introduction to Linguistic Science*. New Haven, Yale University Press.
- Sweetser, Eve. (1987). Metaphorical Models of Thought and Speech: A comparison of historical directions and metaphorical mappings in the two domains. In J. Aske, N. Beery, L. Michaelis & H. Filip (Eds.), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, Berkeley: 446-459.
- Sweetser, Eve. (1988). Grammaticalization and semantic bleaching. *Berkeley Linguistics Society* 14: 389-405.
- Sweetser, Eve. (1991). *From Etymology to Pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge, Cambridge University Press.
- Tabor, Whitney and Traugott, Elisabeth. (1998). Structural Scope Expansion and Grammaticalization. In A. Giacalone Ramat & P. Hopper (Eds.), *The Limits of Grammaticalization*. Amsterdam, Benjamins.
- Tagliamonte, Sali. (1998). Was/Were Variation across the Generations: View from the City of York. *Language Variation and Change* 10 (2): 153-191.
- Tagliamonte, Sali and D'Arcy, Alex. (to appear). *He's like, she's like*: The quotative system in Canadian youth. *Journal of Sociolinguistics*.

- Tagliamonte, Sali and Hudson, Rachel. (1999). *Be like et al.* beyond America: The quotative system in British and Canadian Youth. *Journal of Sociolinguistics* 3: 147-172.
- Tagliamonte, Sali and Roberts, Chris. (2004). So Cool; So Weird; So Innovative. *The use of intensifiers in the television series 'Friends'*. Paper presented at the American Dialect Society Annual Meeting, Boston.
- Taeldeman, Johan. (1989). A typology of dialect transitions in Flanders. In M. Schouten and P. Van Reenen (Eds.), *New methods in dialectology*. Dordrecht, Foris: 155-163.
- Tannen, Deborah. (1986). Introducing constructed dialogue in Greek and American conversational and literary narrative. In F. Coulmas (Ed.), *Direct and Indirect Speech*. Amsterdam, Mouton de Gruyter: 311-332.
- Tannen, Deborah. (1987). *The private Language Problem*. Dissertation Abstracts International A: The Humanities and Social Sciences 47.
- Tannen, Deborah. (1988). *Connecting Observation and Understanding*. Lectures from the 1985 LSA/TESOL and NEH Institutes. Vol. XIX of Discourse Processes. Norwood NJ, Ablex Publishing.
- Tannen, Deborah. (1989). *Talking voices*. New York, Cambridge University Press.
- Tannen, Deborah. (Ed.) (1993). *Framing in discourse*. Oxford, Oxford University Press.
- Thomason, Sarah and Kaufman, Terrence. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley, University of California Press.
- Thompson, Geoff. (1996). Voices in the text: Discourse perspectives on language reports. *Applied Linguistics* 17: 501-530.
- Thompson, Geoff and Hopper Paul. (2001). Transitivity, clause structure, and argument structure: Evidence from conversation. In J. Bybee & P. Hopper (Eds.), *Frequency and the Emergence of Linguistic Structure*. Amsterdam, Benjamins.
- Thorne, Barrie. (1993). *Gender Play*. New Brunswick, NJ, Rutgers University Press.
- Thumm, Markus. (2000). The Contextualization of Paratactic Conditionals. *InList* 20, University of Konstanz.
- Thumm, Markus (in preparation). *Paratactic conditionals in English*. PhD thesis, University of Konstanz.
- Tillery, Jan, Wikle, Tom and Bailey, Guy. (2000). The Nationalization of a Southernism. *Journal of English Linguistics* 28: 280-94.

- Torres Cacoullous, Rena. (2001). From Lexical to Grammatical to Social Meaning. *Language in Society* 30 (3): 443-478.
- Traugott, Elizabeth. (1972). *A history of English syntax. A transformational approach to the history of English sentence structure*. New York, Holt, Rinehart and Wilson.
- Traugott, Elizabeth. (1978). On the expression of spatio-temporal relations in language. In J. Greenberg (Ed.), *Universals in Human Language. Volume 3: Word Structure*. Stanford, Stanford University Press.
- Traugott, Elisabeth. (1982). From Propositional to Textual to Expressive Meanings: Some Semantic-Pragmatic Aspects of grammaticalization. In W. Lehmann, & Y. Malkiel. (Eds.), *Perspectives on Historical Linguistics*. Amsterdam, Benjamins: 245-271.
- Traugott, Elizabeth. (1986). 'Conventional' and 'dead' metaphors revisited. In W. Paprotté & R. Dirven (Eds.), *The Ubiquity of Metaphor: Metaphor in Language and Thought*. Amsterdam, Benjamins: 17-56.
- Traugott, Elizabeth. (1989). On the Role of Epistemic Meaning in English: An Example of Subjectification in Semantic Change. *Language* 65: 31-55.
- Traugott, Elizabeth. (1991). English speech act verbs. A historical perspective. In L. Waugh & S. Rudy (Eds.), *New Vistas in Grammar: invariance and variation*. Current Issues in Linguistic Theory 49. Amsterdam, Benjamins: 387-406.
- Traugott, Elizabeth. (1995). Subjectification in Grammaticalization. In D. Stein & S. Wright (Eds.), *Subjectivity and Subjectification: Linguistic Perspectives*. Cambridge: Cambridge University Press.
- Traugott, Elizabeth. (2001). Legitimate counterexamples to unidirectionality. Paper presented at Freiburg University.
- Traugott, Elizabeth and Heine, Bernd (1991). Introduction. In E. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*. Amsterdam, Benjamins.
- Traugott, Elizabeth and Heine, Bernd. (Eds), (1991). *Approaches to Grammaticalization*. Amsterdam, Benjamins.
- Traugott, Elizabeth and König, Ekkehard. (1991). The Semantics-pragmatics of Grammaticalization revisited. In E. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization*. Amsterdam, Benjamins: 189-218.
- Trudgill, Peter. (1972). Sex, covert prestige, and linguistic change in the urban British English of Norwich. *Language in Society* 1: 179-96.
- Trudgill, Peter. (1974). *The social Differentiation of English in Norwich*. Cambridge, Cambridge University Press.

- Trudgill, Peter. (1980). *On Dialect: Social and Geographic Factors*. Oxford, Blackwell.
- Trudgill, Peter. (1983). *Sociolinguistics: An Introduction to Language and Society*. Harmondsworth, Penguin.
- Trudgill, Peter. (1985). New dialect formation and the analysis of colonial dialects: The case of Canadian Raising. In H. Warkentyne (Ed.), *Papers from the 5th International Conference on Methods in Dialectology*. Victoria, University of Victoria. 34-45.
- Trudgill, Peter. (1986). *Dialects in contact*. Oxford, Blackwell.
- Trudgill, Peter. (1988). Norwich Revisited. Recent Linguistic Changes in an English Urban Dialect. *English World-Wide* 9: 33-49.
- Trudgill, Peter (1994). *Dialects*. London, Routledge.
- Trudgill, Peter. (1999). A window on the past: 'Colonial Lag' and New Zealand Evidence for the phonology of 19<sup>th</sup>-century English. *American Speech* 74: 1-11.
- Ullian, Russel. (1972). The nature of future tenses. In J. Greenberg, C. Ferguson & E. Moravcsik (Eds.), *Universals of human language. Vol 3: Word structure*. Stanford, Stanford University Press.
- Underhill, Robert. (1988). Like is, like, focus. *American Speech* 63: 234-246.
- Vallduví, Enric and Engdahl, Elisabet. (1996). The Linguistic Realization of Information Packaging. *Linguistics* 34, 3(343): 459-519.
- Ver, Co. (1993). Conditiojns d'emploi et interprétation des temps futurs du français. *Verbum* 4: 71-84.
- Vincent, Diane and Dubois, Sylvie. (1996). A study of the use of reported speech in spoken language. In J. Arnold et al. (Eds.), *Sociolinguistic Variation: Data, Theory, and Analysis*. Stanford, CA, CSLI: 361-374.
- Visser, Fredericus T. (1970). *A historical syntax of the English language*. Leiden, Brill.
- Vološinov, Valentin N. (1973) *Marxism and the Philosophy of Language*. Cambridge, MA, Harvard University Press.
- von Hippel, Eric. (1994). 'Sticky Information' and the Locus of Problem Solving: Implications for Innovation. *Management Science* 40 (4): 429-439.



- von Roncador, Manfred. (1988). *Zwischen direkter und indirekter Rede: Nichtwörtliche direkte Rede, erlebte Rede, logophorische Konstruktionen und Verwandtes*. Tübingen, Max Niemeyer.
- Vuillaume, Marcel. (2000). Signaling the Free Indirect Style. *Cahiers Chronos* 5: 107-130.
- Wardhaugh, Ronald. (2002). *An Introduction to Sociolinguistics*. Malden, Oxford, Blackwell.
- Waksler, Rachelle. (2001). A new *All* in conversation. *American Speech* 76 (2): 128-138.
- Walters, Keith. (2002). *How and Why Media Language Has Altered the Nature of Variation in Arabic*. Paper presented at NWAVE 31, Stanford University.
- Websters New World Dictionary of American English. (1994). MacMillan Reference Books
- Weinreich, Uriel, Labov, William and Herzog, Marvin. (1968). Empirical foundations for a theory of Language change. In W. Lehmann & Y. Malkiel (Eds.), *Directions for historical linguistics: A Symposium*. Austin, University of Texas Press.
- Wells, John C. (1982). *Accents of English*. Cambridge, Cambridge University Press.
- Wekker, Herman C. (1976). *The expression of future time in contemporary British English*. Amsterdam, North-Holland.
- Werner, Heinz and Kaplan, Bernard. (1963). *Symbol-formation. An organismic-developmental approach to language and the extension of thought*. New York, London, Sidney, Wiley and Sons.
- Wilson, Deirdre and Sperber, Dan. (2002). Truthfulness and Relevance. *Mind* 111: 583-632.
- Wierzbicka, Anna. (1974). The Semantics of Direct and Indirect Discourse. *Papers in Linguistics* 7: 267-307.
- Wiesemann, Ursula. (1990). A Model for the Study of Reported Speech in African Languages. *The Journal of West African Languages* 20 (2): 75-80.
- Wittgenstein Ludwig. (1978). *Philosophical Investigations*. London, Blackwell.
- Winter, Joanne. (2002). Discourse Quotatives in Australian English: Adolescents Performing Voices. *Australian Journal of Linguistics* 22 (1): 5-21.

- Williams, Ann and Kerswill, Paul. (1999). Dialect levelling: Change and continuity in Milton Keynes, Reading and Hull. In P. Foulkes & G. Docherty (Eds.), *Urban Voices: Accent studies in the British Isles*. London, Arnold: 141-62.
- Wodak, Ruth and Benke, Gertraud. (1997). Gender as a sociolinguistic variable. In F. Coulmas (Ed.), *The Handbook of Sociolinguistics*. Oxford, Blackwell.
- Wolfram, Walter. (1969). *A Sociolinguistic Description of Detroit Negro Speech*. Washington DC, Center for Applied Linguistic.
- Wolfram, Walter and Fasold, Ralf. (1974). *The study of social dialects in American English*. Englewood Cliffs, NJ, Prentice Hall.
- Wolfram, Walt. (1989). Structural variability and phonological development: Final nasals in vernacular Black English. In R. Fasold & D. Schiffrin (Eds.), *Language Change and Variation*. Philadelphia, Benjamins.
- Wolfram, Walt and Schilling-Estes, Natalie. (1998). *American English*. Oxford, Blackwell.
- Wolfson, Nessa. (1981). The Conversational Historical Present Alternation. *Language* 55: 168-182.
- Wolfson, Nessa. (1982). *CHP The Conversational Historical Present in American English Narrative*. Dordrecht, Foris.
- Young, Richard and Bailey, Robert. (1996). VARBRUL analysis for second language acquisition research. In R. Bayley & D. Preston (Eds.), *Second Language Acquisition and Linguistic Variation*. Amsterdam, Benjamins: 253-306.
- Yule, George. (1993). Vera Hayden's Dilemma. Or the indirection in direct speech. In M. Eid & G. Iverson (Eds.), *Principles and prediction: The analysis of natural language*. Amsterdam, Philadelphia, Benjamins.
- Yule, George and Terrie, Mathis. (1992). The role of staging and constructed dialogue in establishing speaker's topic. *Linguistics* 30: 199-215.
- Zilles, Ana. (2003). *Real, or apparent, or both? Three types of evidence for a grammaticalization change in progress in Brazilian Portuguese*. Paper presented at NWAWE 32, University of Pennsylvania.