2002

A corpus-based study of direct speech quotatives in American English conversation

Federica Barbieri
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Bilingual, Multilingual, and Multicultural Education Commons, English Language and Literature Commons, and the First and Second Language Acquisition Commons

Recommended Citation
https://lib.dr.iastate.edu/rtd/7107

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
To my grandmother,
Carmen Bruscagin
A corpus-based study of direct speech quotatives in American English conversation

by

Federica Barbieri

A thesis submitted to the graduate faculty in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

Major: Teaching English as a Second Language/Applied Linguistics (English for Specific Purposes)

Program of Study Committee:
Dan Douglas (Major Professor)
Carol Chapelle
Susan Conrad
Frederick O. Lorenz

Iowa State University
Ames, Iowa
2002

Copyright © Federica Barbieri, 2002. All rights reserved.
This is to certify that the master's thesis of

Federica Barbieri

has met the thesis requirements of Iowa State University

Signatures have been redacted for privacy
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>VI</td>
</tr>
</tbody>
</table>

## CHAPTER I: INTRODUCTION

- Importance of this study ........................................ 3
- Research questions ............................................... 4
- Preview of the study ............................................. 5

## CHAPTER II: LITERATURE REVIEW

- Textbook grammar and real language use: the gap .................. 6
  - Reported speech in ESL/EFL textbooks and grammars .......... 8
  - From sentence grammar to discourse grammar .................. 10
  - The contribution of Corpus Linguistics ...................... 12
- Reported speech ................................................. 16
  - Traditional accounts of reported speech ..................... 16
  - Discourse-pragmatic accounts of reported speech .......... 18
- Direct speech in conversation .................................. 19
  - Reported speech as Constructed Dialogue ................... 19
  - Other approaches to direct reported speech ............... 21
- Types of quotatives .............................................. 27
  - Zero Quotatives ............................................... 27
  - Be like ......................................................... 29
  - Go ............................................................. 35
  - Be all ......................................................... 37
- Summary .................................................................... 38

## CHAPTER III: METHODOLOGY

- Corpora .................................................................... 39
  1. The Longman Spoken and Written English (LSWE) Corpus: Conversation ........ 40
  2. The TOEFL 2000 Spoken and Written English Academic Language (T2K-SWAL) Corpus: Service Encounters and Office Hours ........ 41
- Service Encounters ................................................. 41
- Office Hours ......................................................... 43
- Procedures ............................................................ 46
  - Data collection: simple and advanced searches ............. 46
Retrieval of “real” quotative forms ............................................. 48
Excluded forms ................................................................. 50
It’s like, it was like ............................................................ 50
Conditional, past progressive with gonna and going to, future forms ................................. 51
Data storage ........................................................................... 54
Analytical procedures .......................................................... 54
Quantitative Analysis .......................................................... 54
Procedures in the quantitative analysis ...................................... 55
Qualitative analysis ............................................................. 57
Framework for discourse-pragmatic analysis of the quotatives ................................. 60
Procedures in the qualitative analysis ...................................... 67
SUMMARY ............................................................................. 70

CHAPTER IV: RESULTS AND DISCUSSION ........................................ 71

QUANTITATIVE ANALYSIS ....................................................... 71
1. Analysis of frequency of the quotatives within each register .................................. 71
   Conversation ..................................................................... 71
   Service Encounters .......................................................... 73
   Office Hours ..................................................................... 75
2. Comparison of frequency of use of the quotatives across registers ......................... 77
   Say .................................................................................... 79
   Be like ............................................................................. 79
   Go .................................................................................... 80
   Be all .............................................................................. 81
3. Analysis of frequency of use of the quotatives with grammatical person ..................... 81
   Be like ............................................................................. 83
   Go .................................................................................... 84
   Be all .............................................................................. 84
   Summary of quantitative analysis ......................................................................... 85

QUALITATIVE ANALYSIS ........................................................... 86
Say ..................................................................................... 88
Be like .............................................................................. 89
Go ..................................................................................... 91
Be all .............................................................................. 92
   Summary of discourse-pragmatic analysis of the quotatives ...................................... 93
   First person singular subjects .............................................................................. 93
   Third person singular subjects ............................................................................ 94
   SUMMARY ........................................................................... 95

CHAPTER V: CONCLUSIONS ......................................................... 96

MAJOR FINDINGS ................................................................... 96
LIMITATIONS OF THE STUDY ................................................. 98
IMPLICATIONS FOR FURTHER RESEARCH ............................... 100
ABSTRACT

This study looks at a specific aspect of reported speech: the use of the new quotatives be like, go, be all to introduce direct speech reports in contemporary spoken American English. The study investigates the use of simple present and simple past tense forms of be like, go, be all and the traditional quotative say in three small corpora of spoken American English representing different registers of spoken interaction: conversation, campus-related service encounters, and academic office hours consultations. The study includes quantitative analyses of frequency of occurrence of the quotatives, of frequency in association with grammatical subjects, and comparison of frequency across corpora; and qualitative analysis of the discourse-pragmatic function of the quotatives.

The results show that while the traditional quotative say is still the most frequent direct speech quotative in American English (particularly its past tense form said), the new quotative be like is a stable, significantly frequent quotative which competes with traditional forms within the quotative system as a whole. Go also appears as a relatively stable element, while be all is still rather infrequent. The results reveal that there are clear patterns of association between frequency of occurrence of the quotatives and grammatical person, as well as patterns of association between grammatical person of the quotative and discourse function of the quotative. All new quotatives are used most often in association with first and third person singular grammatical subjects. Say and go generally introduce plausible quotation both when used with first and third person grammatical subjects; on the contrary, be like generally introduces plausible speech when used with third person singular subjects, and improbable speech when used with first person singular subjects.
Finally, this study shows that type and frequency of use of the quotatives vary across corpora representing different registers of spoken interaction.

These corpus findings have potential for grammatical description. In addition, they can be used to inform ELT materials, and thus to provide more accurate and authentic renditions of how real speakers make speech reports in everyday life.
CHAPTER I: INTRODUCTION

Reporting speech is a very important feature of the most quintessentially human characteristic: language. Quoted speech, or direct speech, is an important and extremely frequent feature of conversation. McCarthy (1998) has spelled out what we intuitively know: “Hardly any stretch of casual conversational data is without reports of prior speech” (150).

While speech reporting seems very natural in our first language, however, it may not be so natural in a second or foreign language. McCarthy indeed goes on to state that “it is hard to conceive of achieving any intermediate level of competence in a foreign language without needing to know how the speakers of that language make speech reports” (1998: 150). Therefore, as he points out, language teaching materials and grammars should contain lessons on speech reporting, offer opportunities to practice reports, and present accounts of reported speech that reflect the way speakers actually make speech reports in everyday life (150). Yet, when looking at English language teaching (ELT) materials and traditional grammars we find a simplified, impoverished and inadequate coverage of what actually happens in everyday speech, and, therefore in spoken data (McCarthy 1998, Carter & McCarthy 1995, Yule et al. 1992, Yule 1995). Reported speech is indeed generally depicted as a neat binary system in which, through a number of transformations known as “backshift”, direct speech is transformed into indirect speech.

For example, Focus on Grammar (Fuchs & Bonner 1995), a mainstream, recent reference and practice grammar textbook for high intermediate learners of English, introduces reported speech with two separate tables, one for direct speech and one for indirect speech. The clearly invented examples of direct speech (“He said, ‘The check is in...”)

...
the mail’/’The dress looks good on you’/’The traffic was bad’”) included in the first table are
then transformed into indirect speech in the second table, using the two traditional reporting
verbs *say* and *told* in the past tense (“He said/told (her/Jennifer/his client) (that) the check
was in the mail/the dress looked good on her/the traffic had been bad”). The unit then goes
on with a “grammar notes” page including definitions of direct speech or quoted speech, and
indirect or reported speech, followed by a series of instructions on verb tense, pronouns, and
possessives backshift (271-272). The textbook claims that “we usually use the past tense of
reporting verbs such as *say* or *tell* to report speech” and no other reporting verbs are indeed
mentioned. Moreover, the textbook focuses largely on indirect speech, thus leaving the
learner with the impression that indirect speech is the most important or more common way
to make speech reports in real language use.

Recent research in sociolinguistics, discourse analysis, and the grammar of spoken
English has drawn attention to a number of “reporting devices”, generally called *quotatives,
dialogue introducers, or direct speech introducers*, that real speakers of English use in the
everyday business of reporting speech, and that receive little or no coverage in ELT materials
and traditional, non corpus-based grammars. These include the introduction of quotation
with no overt introducer (“zero quotative”), and the new quotatives *be, be like, be all, go*.

In this study I will be concerned with a small number of these new quotatives in an
attempt to contribute to the description of the quotative system of present-day American
English, and shed more light on the use of these new quotatives. I will look at simple present
and simple past tense forms of the quotatives *be like, go, and be all* in three small corpora of
contemporary spoken American English representing three registers, namely casual
conversation, campus-related service encounters, and academic office hours.
Importance of this study

The importance of this study lies in the need for more comprehensive, up-to-date, and corpus-based accounts of the use of these quotatives in casual conversation and other registers of spoken interaction in order to shed more light on their use within the quotative system as a whole. While the use of these quotatives in everyday spoken American English has indeed been taken up in studies looking at various aspects of speech reporting (Yule 1992, 1993) and in sociolinguistic studies looking at the quotative *be like* in American English (Blyth et al. 1990, Romaine & Lange 1991, Ferrara & Bell 1995) and other varieties of English, e.g. Scottish English (Macaulay 2001), and Canadian and British English (Tagliamonte & Hudson 1999), there are no studies of the use of all these quotatives in principled, diverse corpora of contemporary spoken American English. In addition, virtually all research on direct speech quotation, particularly research on the quotatives (e.g. *be like*), has been based exclusively on conversational storytelling. In other words, no research has been done on the use of these new quotatives in other registers of spoken interaction, such as, for example, academic registers. Because corpus-based research consistently shows that grammatical patterns vary systematically across registers, there is a need for studies looking at the use of new quotatives not only in casual conversation but also in other registers of spoken interaction.

Previous research on the quotative *be like* has looked at the use of the quotative with grammatical person and at the type of quotation introduced by the quotative in order to investigate the process of spread and the discourse-pragmatic function of the quotative. Therefore in this present study I will also look at association of the quotatives with grammatical person and at the content of the quotation introduced by the quotatives in order
to investigate further the discourse function of the quotatives. In particular, I will be concerned with shedding more light on the degree to which these new quotatives are used to introduce quotation that may be taken as plausibly uttered speech or they are used to introduce quotation that is improbable as it simply resembles a representation of some inner thought or speech of the speaker in the reported situation.

**Research questions**

The purpose of this study is to investigate the use of the new quotatives *be like, go, be all* in present-day American English by looking at their overall frequency, association with grammatical person, and discourse-function in three corpora of spoken American English representing different registers of spoken interaction. Therefore, this study addresses the following research questions:

1) What is the frequency of use of the new quotatives *be like, go, be all* and in present-day American English?

1a) How frequent are the quotatives *be like, go, be all* and the traditional quotative *say* in the three corpora used in this study, Conversation, Service Encounters and Office Hours?

1b) Are there differences in frequency of occurrence between casual conversation and other registers of spoken interaction as represented in Service Encounters and Office Hours?

1c) What is the frequency of use of the quotatives in association with grammatical person?
2) What is the discourse-pragmatic function of the new quotatives *be like, go, be all*? Specifically, what type of quotation do they introduce? Is it quotation that may be taken as plausibly uttered speech, or does it resemble a representation of some inner speech or thought of the quoted speaker?

**Preview of the study**

In Chapter Two I provide a comprehensive review of various areas of research informing this study: studies of the grammar of spoken English, corpus linguistics methodology, literature on reported speech and direct quotation in conversation, and studies of quotatives. In Chapter Three I describe the data, the corpora, and the methodology used in the study. In Chapter Four I present and discuss the results of the study. Chapter Five summarizes the major findings of the study and includes a discussion of possible limitations as well as implications for further research that might be drawn from this study.
CHAPTER II: LITERATURE REVIEW

A corpus-based study on new quotatives in contemporary spoken American English calls for a critical review of different areas of research, ranging from reported speech, direct quotation in conversation, studies of quotative forms, but also empirical studies of the grammar of spoken English and corpus linguistics methodology. This chapter includes four main sections. In the first section I look at literature pointing to the limitations of treatments in traditional grammars of reported speech, in the context of a broader lack of fit between grammar and textbook descriptions and real language use; I introduce recent calls for a discourse, corpus-based approach to grammar as a response to the gap; and I describe the strengths of a corpus-based approach to the study of language. In the second section I review traditional and discourse-pragmatic accounts of reported speech. In the third section I focus on direct speech and discuss different approaches to direct quotation in conversation. Finally, in the last section I provide a comprehensive review of previous studies of quotatives.

Textbook grammar and real language use: the gap

In recent years, a growing number of studies (Carter & McCarthy 1995, McCarthy & Carter 1995, Di Vito 1991, Glisan & Drescher 1993, Walz 1986, Lawson 2001) have compared the textbook and grammar descriptions of a target language with the language used in real life by real language users. Not surprisingly, English has been the object of the majority of these studies, but it was interesting to find out that other Western European
languages, i.e. French (e.g. see Di Vito 1991, Walz 1986, Lawson 2001) and Spanish (e.g. see Glisan & Drescher 1993), have also been involved into this type of research.

Underpinning this bulk of research is the need to assess the extent to which target language textbook descriptions provide an authentic picture of the language used in everyday life by real language users, in order to identify possible gaps and eventually generate implications for materials writing and language teaching practice. Underpinning this research is also the belief that, in a communicative approach to language teaching emphasizing that language learners should acquire the ability to speak in a natural way, the language presented in language materials should accurately reflect real language user norms, and, further, that frequency of occurrence (of vocabulary and grammatical structures) in the target language should inform decisions about what to prioritise in the presentation of materials.

Surprisingly, these studies unanimously show that there is a great divide, a lack of fit, between grammar and textbook descriptions of the target language and real language use. All of the studies mentioned above indeed demonstrate that, despite over two decades of language teaching aiming at fostering speaking skill and natural spoken interaction, textbooks neglect important and frequent features of the language spoken by real language users, present a patchy, confusing, and often inadequate treatment of common features of the grammar of the spoken language, and, in sum, do not reflect actual use.

The lack of fit between textbook descriptions and real language use may be attributed to several factors: 1) textbook descriptions often rely on intuitions, rather than on empirical, corpus-based data; 2) textbooks are not informed by empirical, corpus-based evidence about the occurrence and frequency of occurrence of language features; 3) textbooks usually
present grammatical and lexical patterns as equally generalizable and equally important communicatively, thus neglecting information about register-specific or discourse-context-specific use; 4) textbooks are usually based on written norms only, thus ignoring the spoken language.

**Reported speech in ESL/EFL textbooks and grammars**

One important area of spoken English that receives inadequate coverage in traditional reference grammars and English language teaching (hereinafter ELT) materials is reported speech: traditional reference grammars and ELT materials present reported speech as a binary system consisting of *direct* and *indirect* speech. They focus on the syntactic feature of backshift and mechanical exercises for converting direct speech into indirect speech, while providing virtually no coverage of the various *choices* open to native speakers in the everyday business of reporting speech in casual conversation. Yule et al. (1992; See also Yule 1993, 1995) point out that at the formal level there is a continuum of choices, and not a binary choice between direct and indirect speech: when reporting discourse, speakers may choose between indirect speech, direct speech introduced by a range of quotative verbs in addition to the traditional *say (be, be like, go)*, and direct speech presented without an overt introducer, using a null form or zero introducer ("zero quotative").

Yule et al. (1992) and Yule (1995) argue that there are some drawbacks to the presentation of reported speech as a binary, mechanical system: first, the procedure of presenting one mode (direct speech) as primary, and the other (indirect speech) as derived is just "a dangerous illusion" (1992: 246), not only because in writing one mode of reported speech is generally selected as a composing decision without necessarily having actual
recorded speech as a source, but also because there are instances of reported speech that do not have one single, easily constructed direct speech equivalent. Secondly, this presentation does not account for the fact that direct speech and indirect speech have different discourse functions, and, consequently, "one form simply does not take the place of the other" (1995: 186). Thirdly, at the formal level, there appears to be a continuum of forms available for reported speech, and not a binary choice. Yule et al. (1992; See also Yule 1993, Yule 1995) discuss a number of options used to present reported speech in real written and spoken English, such as indirect speech, free indirect discourse, direct speech as constructed dialogue, direct speech introduced by the new quotatives be, be like, go and by no overt introducer.

Carter & McCarthy (1995; see also McCarthy & Carter 1995; McCarthy 1998) for example, use corpus data from the CANCODE Spoken Corpus to explore several grammatical features of spoken English that are neglected in ELT materials and major reference grammars, such as situational ellipsis, left dislocation and topical information, tails (or "right dislocation")\(^1\), and past continuous as a reporting verb in indirect speech, as in *Tony was saying they should have the heating on by about Wednesday*. Carter & McCarthy point out that the past continuous as a reporting verb in indirect speech is neglected not only by ELT textbooks and grammars, but also by secondary literature on indirect speech, including studies of real spoken and written data, such as Yule et al. (1992) and the corpus-based COBUILD Reporting Guide.

\(^1\) The phenomenon by which the speaker inserts grammatical patterns which reinforce or amplify what (s)he is saying or has said, as in: She's a really good actress, *Clare*; Singapore's far too hot for me it is. (McCarthy 1998: 180)
The importance of the studies reviewed above for this present study is twofold: first, they point to a gap in the literature on reported speech and, thus, to the need for more research; secondly, they provide a starting point, a ‘springboard’ for the identification of forms to be taken into consideration in this study. Yule et al. (1992) and Yule (1995), for example, note the use, in contemporary spoken American English, of various quotative forms, including *be like, go, be, be all*, and what they name ‘zero quotatives’, i.e.: the presentation of reported speech with no introducing verb and attributed speaker, as in:

She’s like, *'So what time did you get in?' We got in at two thirty. 'Well I got home around a little after one' cause they sleep like the dead—they don’t hear us come in anyway and eh so 'Did you all have a nice time?' —'yeah'.* (example taken from Yule et al. 1992: 249)

From sentence grammar to discourse grammar

One of the main implications of much of the research reviewed above, particularly research by Carter, McCarthy and Hughes, is that sentence-based accounts are inadequate if we wish to provide an authentic description of the language spoken everyday by native speakers. There are two main reasons why sentence grammar is inadequate.

First, sentence grammar is insensitive to context: it provides a static, decontextualised picture of the target language. Di Vito (1991) points out that “in order to use any target language structures in native-like ways, learners must develop native-like expectations and native-like intuitions about the use of those structures. In other words, these language learners must not only be aware of what structures are possible in the target language, but also what structures are both *probable and appropriate* in specific discourse contexts” (383).

By providing no information about the *sociolinguistic appropriateness* of linguistic structures, language materials induce learners to assume that the various grammatical
structures presented are equally generalizable, equally important communicatively, and equally productive in the target language. Di Vito (1991) argues that such inaccurate representations of the target language mislead students as to the nature of the target language, and encourage them to form false intuitions about real spoken language.

Secondly, sentence grammar is mostly based on the written language, as Carter & McCarthy point out (1995; see also McCarthy & Carter 1995, Carter 1998, Hughes & McCarthy 1998, Carter et al. 1998). Carter et al. (1998) observe that “correct grammar” has come to mean ‘correct grammar as represented by the written language’, so that “many perfectly normal and regularly occurring utterances made by standard English speakers [...] have by omission come to be classified as ‘ungrammatical’” (67). Thus, Carter and McCarthy (1995) argue that “just as it would be questionable to base a writing skills course on grammatical statements based only on informal spoken data, [...] it is equally the case that spoken language instruction based solely or mainly on written language description is an unsound methodological foundation upon which to build” (141-142).

For all these reasons, researchers like Carter, Hughes and McCarthy have been strongly arguing for the development of grammars of the spoken language and of ‘discourse grammars’, as opposed to sentence based grammars, reflecting written norms. In a ‘discourse grammar’ perspective, grammar is “an aspect of discourse, rather than something that operates only within the boundaries of the clause or sentence” (Hughes & McCarthy 1998: 264); in a discourse-grammar approach, “grammar exists only as a trace of the discourse process, and it is best viewed as the regular patterns left behind by millions of conversational and written texts in which the exigencies of communication are paramount [...]” (279). A discourse grammar, indeed, aims at acknowledging the fact that “the kinds of
grammatical choices that speakers and writers make often depend on contextual features” (265). Further, a discourse grammar is corpus-based as it avails itself of real data, and is probabilistic in that it accounts for the “probabilistic correlations” of forms and contexts that can be observed using corpus evidence. Therefore, a discourse grammar may require “the redefinition of traditional paradigms in line with real choices in discourse” (278), and the re-examination of prescriptive and deterministic rules in the light of real, corpus-based data.

A discourse grammar is also necessary to account for differences between spoken and written grammar. Carter, McCarthy and Hughes (Carter & McCarthy 1995, Carter 1998, Hughes & McCarthy 1998, Carter et al. 1998) illustrate a discourse-based treatment of various structures (e.g. various forms of ellipsis, right dislocation, and different forms of reported speech) that are typical of the spoken language and receive no coverage in traditional grammars, which, as mentioned above, are based on the written norms.

The contribution of Corpus Linguistics

Another major implication of much of the work reviewed at the beginning of this chapter is the potential of Corpus Linguistics for the development of more authentic descriptions of the target language. The studies by Lawson (2001), Di Vito (1998), Glison & Drescher (1993), Carter & McCarthy (1995) are based on small corpora of spoken language. The use of corpora has allowed these researchers to assess to what extent textbook descriptions authentically reflect real language use, proving that intuition, which underlies much of the choices determining presentation of topics in ELT materials, is often misleading and fallacious. On the other hand, the work of Carter and McCarthy demonstrates the validity of small corpora as a testbed for intuitions about language.
An exhaustive account of corpus linguistics, its methods and its advantages is beyond the scope of this study. Here, I will only highlight a few main points.

In a nutshell, corpus linguistics may be defined as “the empirical study of language relying on computer-assisted techniques to analyse large, principled databases of naturally-occurring language” (Conrad 2000: 548). Recent work by Biber and Conrad highlights three major strengths of the corpus-based approach: 1) its methodology; 2) its ability to demonstrate unreliability of intuitions about use; 3) its ability to show the centrality of register for language use. While the first point refers to the intrinsic characterisation of corpus-based research, the second and the third points are generalisations that Biber and Conrad define as “crucial for ESL/EFL teaching” (Biber & Conrad 2001: 332). I will now briefly illustrate these three strengths.

1. Methodology

Corpus-based research has four essential characteristics (Biber et al. 1998, Conrad 1999):

1) It is empirical, in that it analyses the actual patterns of language use in naturally-occurring texts. This means that corpus-based research is concerned with the actual language used in naturally occurring texts, rather than in what is theoretically possible, as in traditional studies of structure.

2) It uses a large and principled collection of naturally-occurring texts in the corpus as the basis for analysis. It is important that the corpus be large, as if it is too small it will not include representative samples of the item being studied. How large the corpus should be, however, will depend on the type of linguistic feature under investigation. For example, lexicographic studies require the use of corpora of several million words,
whereas studies of discourse require much smaller corpora. Broadly speaking, the overwhelming amount of data yielded by large multimillion word corpora make qualitative, discourse analytical studies impractical. It is important that the corpus be a principled database, in the sense that it should be representative of the varieties of language (dialect, register, etc.) that it is supposed to represent. For example, a corpus of written texts is not suitable to obtain information about the spoken language, just as a corpus of newspaper articles would not be representative of scientific language, nor of all written language. Representativeness is a key notion in corpus design.

3) It uses computers for analysis. The use of the computer to process data from the corpus and conduct analyses is necessary as analyses of large quantities of data would otherwise not be feasible, in addition to being inaccurate. Corpus-based research avails itself of both automatic and interactive techniques. Automatic techniques are suitable for the analysis of non ambiguous linguistic features, while interactive techniques (i.e.: techniques requiring human interaction) are necessary for the analysis of ambiguous linguistic features.

4) It relies on both quantitative and qualitative, functional interpretations of language use. Biber and Conrad strongly argue against the view that quantitative analyses consist merely of elaborate bean counting. While quantitative analyses are necessary to determine patterns of frequency and association with other linguistic features, qualitative, functional analyses are necessary to go beyond simple counts and interpret and describe the communicative functions that correspond to the quantitative patterns.

2. The Unreliability of intuitions about use
Corpus-based research shows that the intuitions on which linguists and teachers may often rely are often incorrect and fallacious when tested empirically against the actual patterns of use in large text corpora.

It appears that humans have a tendency to notice unusual occurrences more than typical occurrences, and therefore intuition can be unreliable. The literature on corpus-based research contains a number of examples of corpus-based analyses which have yielded results proving intuitions or commonly accepted ‘facts’ about language wrong (See for example, the two case studies on common lexical verbs and aspect across registers in Biber & Conrad 2001).

3. The centrality of register for studies of language use

Corpus-based research consistently shows that grammatical patterns differ systematically across registers (i.e., varieties of language determined by their purpose and situation of use) at all linguistic levels. Strong patterns in one register often represent only weak patterns in other registers. This means that few descriptions are adequate for a language as a whole, because languages are not homogeneous in their linguistic characteristics across registers.

In a recent talk on the Longman Grammar, Biber (2002) declared that “one major finding in doing the grammar was the belief that register is so important that there is no such thing as a general use”. Conrad (2000) predicts that one of the changes that will revolutionise the teaching of grammar in the 21st century is that “monolithic descriptions of English grammar will be replaced by register-specific descriptions” (549)
Reported Speech

In this section I present a more comprehensive overview of the literature strictly on reported speech relevant to this study. Reported speech has attracted the interest of researchers in several fields, including poetics, logic, philosophy, and literature. A comprehensive review of such a vastly treated linguistic phenomenon would be beyond the scope of this study. Thus, while providing an overview of major approaches to the study of reported speech, this review will devote more space to studies that are more relevant to direct reported speech, rather than indirect reported speech.

The literature on reported speech may be critically divided into two broad categories: traditional or syntactic accounts, and discourse-pragmatic accounts of reported speech. In this section I will first briefly overview traditional or syntactic approaches, which emphasize the syntactic dimension of speech reporting, and major criticism to this approach, which emphasizes the interaction of syntactic, pragmatic and stylistic factors in discourse. These approaches are what Baynham calls discourse-pragmatic approaches (Baynham 1991, 1996). I will then proceed to a more in-depth review of discourse-pragmatic accounts of direct speech reporting. I will review Tannen’s work on ‘constructed dialogue’ and various other studies of direct speech quotation in conversation and other contexts. Finally, I will look at the literature on quotatives.

Traditional accounts of reported speech

Traditional accounts of reported speech distinguish between Direct Reported Speech (sometimes also called Direct Quotation) and Indirect Reported Speech. The distinguishing criterion has usually been considered the relationship between the report and the reported or
original utterance: direct reported speech is said to report the words of the original speaker
verbatim, whereas indirect reported speech is said to express the content or the sense of what
was originally said, but not necessarily the form, i.e., the words. Take, for example, the
account of reported speech given by one of the most authoritative reference grammars of
English: *The Comprehensive Grammar of the English Language* (hereinafter: CGEL) by
Quirk et al. (1985):

"Direct speech purports to give the exact words that someone (who may be the
reporter) utters or has uttered in speech or in writing. Indirect speech, on the other
hand, conveys in the words of what has been said or written by the original speaker or
writer (who again may be the same person as the reporter). Contrast the direct speech
in [1] with the indirect speech in [1a]:
David said to me after the meeting, ‘In my opinion, the arguments in favour of
radical changes in the curriculum are not convincing.’ [1]
David said to me after the meeting that in his opinion the arguments in favour of
radical changes in the curriculum were not convincing. [1a]" (Quirk et al. 1985:
1021)

In traditional accounts, reported speech is seen as a neat two-term system in which a
series of syntactic transformations carried out on direct speech allows to transform it into
indirect speech. These syntactic transformations involve pronoun shift, tense shift, mood
shift, and embedding of the reported clause in a matrix sentence dominated by a verb of
saying or communication. A paradigmatic example of this theoretical approach to reported
speech, with its syntactic and deictic shifts, is presented by Lyons (1968):

"Take for instance a passage such as the following: The prime minister said that he
deeply regretted the incident. He would do everything he could to ensure that it did
not happen again. On the following day he would confer with his colleagues. He was
confident that… Once again, passages of this kind are best accounted for in two
stages: first of all, by describing a set of independent sentences in their ‘direct’ form
(I deeply regret…; I will do everything I can…; Tomorrow I will confer…; I am
confident that…) and then, by specifying the secondary grammatical rules which will
transpose each of these sentences into the corresponding ‘indirect’ form when they
occur in sequence after a ‘verb of saying’." (Lyons 1968: 174)
Looking at this quote from Lyon’s *Introduction to theoretical linguistics* it is clear that underlying treatments such as that in the CGEL is a theoretical, purely syntactical approach to reported speech. The CGEL in turn informs many non corpus-based ESL/EFL grammars, which, as we have seen in the section “Reported Speech in ESL/EFL textbooks and grammars” present a treatment of reported speech that has been criticised in recent discourse-based work such as Yule et al. (1992), Yule (1993, 1995), Carter & McCarthy (1995) and McCarthy & Carter (1995).

**Discourse-pragmatic accounts of reported speech**

The limitations of traditional or syntactic approaches to reported speech have been noted by a number of researchers who have pointed out the inadequacy of traditional approaches to account for the complexity of speech reporting in discourse. Vološinov’s account of speech reporting is often cited in the literature on direct speech as the first attempt to account for the properties of reported speech in discourse. Criticising Peskovskij’s grammatical transformation approach, Vološinov ([1929] 1978) argues:

“Peskovskij makes a typical grammarian’s error. His mechanical, purely grammatical model of translating reported speech from one pattern to another, without the appropriate stylistic reshaping, is nothing but a bogus and highly objectionable way of manufacturing classroom exercises in grammar. This sort of implementation of the patterns of speech reporting has nothing remotely to do with their real existence in a language. The patterns express some tendency in one person’s active reception of another’s speech.” (Vološinov, 1978: 160)

The discourse-pragmatic approach has been taken up from a theoretical perspective also in Bansfield (1982) and Coulmas (1985). Both Coulmas and Banfield criticise sentence approaches to reported speech on the basis of various categories of utterances that are problematic for mechanical treatments of speech reporting, even though from different
theoretical contexts: Banfield in an attempt to deal with the syntax of speech reporting within a variant of the Extended Standard Theory, Coulmas from the perspective of pragmatics.

Discussing the difference between direct and indirect speech, Coulmas observes that “where direct speech is expressive, indirect speech is descriptive”. He shows that the more expressive elements an utterance contains, the more changes are necessary for reporting it in indirect speech. Further, Coulmas identifies “restrictions on the kinds of expressions that cannot, or should not, be expected to occur in indirect speech” (45), i.e., a number of elements of direct speech that do not occur in indirect speech, such as expressive elements (interjections, directly addressed invectives, curses, etc.); terms of address; intonation; sentence moods (imperative, hortative, interrogative); elliptical sentences; discourse organizing signals (starters, pause fillers, hesitation signals, etc.); and other elements such as false starts, self-correction, repetition, etc. (48).

Direct Speech in Conversation

Reported speech as Constructed Dialogue

A number of studies observe that the reporting of what was said in direct speech forms in conversation is more likely to be a construction of the reporter than a verbatim report of the speech of any reportee. The most vigorous advocate of this reinterpretation of the nature of direct speech (or direct quotation) in conversation is Deborah Tannen, who argues that “the term “reported” speech is a misnomer” in that “what is commonly referred to as reported speech or direct quotation in conversation is constructed dialogue, just as surely as is the dialogue created by fiction writers and playwrights” (1986: 311). Tannen reiterates
this point in a number of articles and book chapters (Tannen 1982, 1986, 1988, 1989) in which she compares oral and literal strategies in literature (fiction) and spontaneous conversation arguing that “ordinary conversation and literary discourse have more in common than has been commonly thought” (1988: 89).

Tannen argues that “constructed dialogue in conversation and in fiction is a means by which experience surpasses story to become drama” (1986: 312; See also Tannen 1988). She further argues that through drama the speaker creates interpersonal involvement: “the creation of drama from personal experience and hearsay is made possible by and simultaneously creates interpersonal involvement among speaker or writer and audience” (1986: 313; See also Tannen 1988). In other words, for Tannen, direct speech in conversation resembles direct speech in fiction in two ways: first, it is a creation of the speaker, a construct, and as such “a means by which experience surpasses story” (1986: 312); second, it creates emotional “involvement” between the speaker and his audience. It creates involvement in the sense that it provides “particulars” by which speakers and listeners collaborate in imagining and participating in similar worlds: by recognizing particulars and reconstructing them from remembered associations with other particulars, listeners can imagine the scenes in which those particulars could occur, and in this way they can understand and appreciate a story. Thus, Tannen identifies in the notion of “particularity” the key to the creation of emotional involvement between writer/speaker and reader/listener: “the casting of thoughts and speech in dialogue creates particular scenes and characters, and it is the particular which moves readers by establishing and building on a sense of identification between speaker or writer and audience” (1988: 92).
In a later contribution, Tannen (1989) discusses her reinterpretation of "reported speech" as "constructed dialogue" more thoroughly and, by analysing examples more closely, distinguishes several types of constructed dialogue: dialogue representing what wasn't said, dialogue as instantiation (dialogue offered as an instantiation of a general phenomenon), choral dialogue (dialogue offered as an instantiation of what many people said), dialogue as inner speech (dialogue reporting thoughts), the inner speech of others (dialogue presenting the thoughts of a character other than oneself), dialogue constructed by a listener (when a listener constructs dialogue appropriate to someone else's narrative), fadeout, fadein (dialogue fading from indirect to direct), dialogue including vague referents, and dialogue cast in the persona of a nonhuman speaker.

For the purposes of this study, the interest of Tannen's work lies not only in her redefinition of 'direct speech' or 'direct quotation' as 'constructed dialogue' — a redefinition which, as we will see, has remained seminal — but also in her attention to how the dialogue is introduced and thus to dialogue introducers or 'quotatives', as we will see more in detail in the section "Be like", below in this chapter.

Other approaches to direct reported speech

While the work by Tannen on constructed dialogue in conversational storytelling has been very influential, it is not the only work that has been done on direct reported speech. In this section I provide an overview of other studies of direct speech. Most of these studies look at the functions of direct speech in conversational narrative, while a few others investigate the functions of direct speech in non-narrative discourse.
Macaulay (1987) analyses quoted direct speech in the oral narratives of the inhabitants of Ayr, in Southwest Scotland. Observing that although the use of quoted direct speech is not necessary to convey the prepositional content of the utterance, quoted direct speech is much more common than indirect speech in the narratives analysed in his study, Macaulay sets out to explore what purpose the use of direct speech serves. He identifies a number of functions of quoted direct speech in the narratives: indirection, embedded evaluation, mimicry, taboo expressions, authenticity, translation, self quotation. Here I only review the most important of these functions.

Macaulay shows that one important function of quoted direct speech is expressing indirection, in the sense that quoted direct speech is able to convey information implicitly that it might be more awkward to express explicitly. Another important function of quoted direct speech, according to Macaulay, is the expression of what Labov (1972) has called ‘embedded’ or ‘internal’ evaluation. While in external evaluation the narrator gives the justification about why the story is reportable or worth telling explicitly, in internal evaluation the point of the story is conveyed through the story itself, including the remarks of another person. Macaulay observes that the function of embedded evaluation is one of the most frequent uses of quoted direct speech presumably because responsibility for the remark is clearly attributed to someone other than the speaker. Authenticity is another major function of quoted direct speech because “the assumption underlying the use of quoted direct speech in narratives is that the speaker is reporting what was actually said” (14).

Macaulay criticizes Tannen’s claim that all reported speech or direct quotation in conversation is constructed dialogue: he argues that because there is evidence that people have a substantial memory for surface structure and not just meaning in experiments testing
recall of dialogue, it is not unreasonable that speakers should be able to recall specific remarks that made an impression on them, in non-experimental situations where there is emotional involvement. He argues that “it is equally hard to show that quoted direct speech is always constructed as it is to show that it has been recalled accurately” (15).

Mayes (1990) addresses the question of whether direct quotations are really quotations, in the sense of how authentic they are, of whether they represent actual previous utterances or whether instead they are inventions of the speaker. Mayes shows that many direct quotes in informal spoken discourse are invented and that speakers seem to prefer direct quotes over indirect quotes. She found that at least 50% of the direct quotes in her data do not represent actual previous utterances, and argues that this supports the suggestion that speakers are using direct quotes for specific discourse functions. Mayes identifies two main discourse functions of the use of direct quotation in conversation: 1) evaluative: direct speech is used to effectively dramatise key elements in narrative; 2) evidential: direct quotation is used as evidence which allows the speaker to present a more believable story.

Thus, Mayes, like Macaulay (1987), identifies in Labov’s concept of internal evaluation one of the main functions of the use of direct quotation in conversation. As mentioned above, for Labov, internal evaluation is implicit in that it is contained inside the story itself: the narrator shows rather than tells the audience what the point is, and very often avails himself of direct quotation. Direct quotation is one type of internal evaluation. It is used to show what the point of a story is and to dramatise it.

The second most important function of direct quotation for Mayes is that of providing evidence. The basis of this function lies in “the popular belief that direct quotes are exact and, therefore, more factual or reliable than indirect quotes” (353). As an example of
quotation used as evidence, Mayes cites a talk show story about a wedding disaster. In telling the story, the speaker quotes an expert to increase her credibility, since the story was somewhat unbelievable. The story is indeed about a bride (the storyteller) whose lung was punctured by a corsage pin, the night before her wedding, while she was being hugged by a guest. Afterwards, when she became dizzy and short of breath, nobody would believe she was seriously injured. The speaker quotes an expert (a radiologist) to provide evidence that, as incredible as it sounds, the corsage did puncture her lung.

For the purposes of this study, one of the most interesting aspects of Mayes’ study is her notion of a continuum from quotes which could be authentic renditions to quotes which are unquestionably invented. Mayes indeed agrees with Tannen’s claim that it is not possible that direct quotes could be exact repetitions of previous utterances and therefore direct quotation is ‘constructed dialogue’, but she adds the notion of a continuum from authenticity to invention. Mayes divides the continuum into four categories: 1) plausible quotes; 2) improbable quotes; 3) highly improbable quotes; 4) impossible quotes. She uses two types of evidence to place a quote in a category: internal evidence (“lexical or syntactic evidence inside the quote itself or inside the utterance containing the quote that indicates that the quote is (or may be) invented.” (331)) and situational evidence (“evidence concerning the entire discourse situation that casts doubt on the quote” (331)). Mayes’ study is interesting for the purposes of this present study as I will draw on Mayes’ framework for the classification of direct quotes and the analysis of the discourse function of the quotatives.

The evidential discourse function of direct quotation is taken up by Holt (1996), who shows that direct reported speech is “an effective and economical way of not only reporting a previous interaction but also giving evidence regarding what was said” (221). Holt contends
that the use of direct reported speech in conversation has an evidential function as it is used
to provide evidence of a previous comment or interaction.

A most influential account of direct reported speech in conversation is Clark &
Gerrig’s (1990), who argue that quotation is a kind of demonstration, a non-serious depiction
of the words of another. Clark & Gerrig argue that demonstrations differ from descriptions
and indications in two main ways: 1) they are non serious actions; 2) they depict, rather than
describe, their referents only selectively. As a consequence, quotations too are non-serious
actions and selective depictions. Moreover, speakers aren’t necessarily committed to trying
to reproduce a source utterance verbatim. Indeed, Clark & Gerrig, like Tannen, argue against
the verbatim assumption of traditional characterizations of reported speech:

“[…] almost every argument we have adduced for the demonstration theory is also an
argument against the verbatim assumption. By our account, what speakers commit
to themselves to in a quotation is the depiction of selected aspects of the referent.
Verbatim reproduction per se has nothing to do with it.” (795)

While most of the work on direct reported speech has traditionally considered
reported speech in narratives, some recent work on direct reported speech looks at the
occurrence of direct speech in non-narrative discourse.

Baynham (1996), for example, analyses the use of direct speech in non-narrative
contexts in classroom discourse, more specifically in a corpus of recordings of adult
mathematical discourse. Baynham argues for a situated account of speech reporting, which
is sensitive and takes into account of its functions in particular discourse contexts. Baynham
tackles the issue of whether it is adequate to treat direct speech reporting as the verbatim
reproduction of an original utterance, and argues that the analysis of speech reporting in non-
narrative discourse contexts further calls into question the verbatim assumption, in that the
functions of reported speech in the data analysed are much more diverse than the simple reproduction of an original utterance.

Baynham also criticizes Tannen’s idea of constructed dialogue: he argues that the focus on constructed dialogue leads to a loss of the grammatical relationship between direct and indirect speech and other polyvocal strategies for referring to the speech of others in discourse. Baynham calls for the necessity to look at “the alternation of reported speech with other means of referring to acts of speaking, including indirect speech reporting” (68).

Myers (1999) looks at the use of direct reported speech in focus groups (i.e.: a moderated discussion between selected participants led through a topic by a moderator). Myers argues that reported speech always suggests a shift in frame (in the sense of Goffman’s frame shift from the primary frame that we take to be immediate reality, to another frame shared for the purposes of the interaction) and that the shift can focus attention on different aspects of the reported utterance, such as setting, factuality, positioning and wording. Myers uses these shifts to categorise functions of reported speech in the focus groups. He draws on Clark & Gerrig’s (1990) distinction between “direct experience” and “detachment” to identify different interpretations within the four functions mentioned above (setting, factuality, positioning, and wording). He explains that this taxonomy of reported speech serves to link the conception of reported speech as depicting and detaching speech to the various uses observed in focus groups, and can account for the formal features characteristic of focus groups, but not of literary texts and grammar books.

Thus, Myers, like Baynham, points to the need for accounts of direct speech that are more sensitive to context, and stresses that the functions of reported speech are more complex than what is suggested by studies of written, and particularly of literary language.
He concludes stating that his work supports Baynham's, McCarthy's and others' contention that "language learners need much more than direct/indirect speech exercises of some books. Different as these studies are, they all suggest, as does mine, that learners should start, not with a grammatical question (what tense? what pronoun?) but with a discourse analytical question: why do people quote?" (397).

Types of quotatives

Zero Quotatives

In an article presenting a classification of types of constructed dialogue similar — though much less comprehensive — to the one discussed by Tannen (1986, 1988, 1989), Yule (1995) also notes constructed dialogue attributed to a non-human 'persona', 'zero quotatives', constructed dialogue as future possible thought rather than utterance, and dialogue giving voice to thoughts, feelings, and attitudes.

As mentioned above, indeed, Tannen's redefinition of reported speech as constructed dialogue has been very influential. One remarkable study in this line of research is Mathis & Yule (1994)'s Zero Quotatives, an in-depth study of direct speech forms presented with no introducing verb or attributed speaker. Analysing occurrences of zero quotatives in the casual conversational speech of American women, Mathis & Yule found that zero quotatives are used for a range of dramatic purposes, in a number of situations where a full quotative form (such as say, go, and be like) would be possible but is not used. In a large number of cases, zero quotatives are used when, in representing an interaction where two distinct participants have been identified, the speaker presents what each participant says in
sequence, without introductory quotative frames, to achieve a dramatic effect (for example, the urgency of the interaction being reported) which cannot be achieved with the presence of full quotatives. Zero quotatives are also found when a speaker is giving voice to her attitude, or the attitude of another character, in direct speech form, in reports where no actual interaction is being reported. Another use of direct speech forms with zero quotatives is when a second speaker adds his voice to the construction of a reported interaction, thus echoing in paraphrase the attitude of the first speaker, and supplying a line of dialogue. In these cases, the zero quotative allows the speakers to show their strongly convergent behaviour and stress their similarity by constructing utterances which may be spoken by either speaker. Finally, zero quotatives occur in sites when utterances attributed to another character are clearly marked as being constructed rather than reported. Here, again, speakers merge their voices to become that of a character and “underscore their sameness” (75).

Mathis & Yule (1994)’s account of zero quotatives is the most comprehensive and in-depth in the literature; however, the phenomenon of direct speech quotation presented with no introducing verb or attributed speaker had been previously analysed by Tannen (1986, 1988) in a study comparing the forms and functions of conversational and literary narrative in American English and modern Greek. Tannen found that ‘zero quotatives’, which she calls ‘unintroduced dialogue’, accounted for 26% in the American stories and for a comparable 22% in the Greek stories. While the use of no lexicalised introducer was found to account for a significant percentage of all four discourse types under examination (American conversational and literary narrative, and Greek conversational and literary narrative), what is surprising is the fact that the percentage of unintroduced quotes was larger (rather than smaller) in the conversational narratives than in fiction. Tannen explains that
this is possible thanks to the great versatility of the human voice, which allows speakers to change their ‘footing’ from that of narrator to that of character, and to present the dialogue of characters without introducing them, by talking on their voices by shifts in pitch, amplitude, voice quality, prosody and pacing.

The phenomenon of the presentation of direct speech quotation with the omission of the reporting verb is also briefly taken up by Romaine and Lange (1991), who call it “bold, unframed or unbracketed reporting” (235).

**Be like**

Stern (1995) reports that *like* occurs “a staggering 9,010 per million” words in the Longman Spoken American Corpus, a five million words database gathered from 12 regions across the United States, used for the compilation of Longman dictionaries. Since Americans say *like* once in every 110 words, *like* is “a word that warrants much more attention”, notes Stern. As she points out, indeed, *like* is not always just “the meaningless filler of space that it might seem to be at first glance” (5): *like* is often used with the auxiliary verb *be* to introduce direct speech, as a substitute of *say* and *go*, as in *...another song comes and she’s like, you wanna go dance? And I’m just like...alright* (7).

The use of *be + like* as direct speech quotative was first attested in 1982 in an Editor’s Note in the journal *American Speech* by Butters, who points out the use of “to be (usually followed by *like*) where what is being quoted is an unuttered thought” (1982: 149). Schourup (1985) accounts for *be + like* as direct speech quotative among the non-standard uses of *like*, where “non-standard” refers to the fact that they are not accounted for/mentioned in the Webster’s Third International Dictionary”. Schourup says that “*like* in
its quotative use introduces [...] a direct discourse rendering of what someone was thinking” (44) and reports that informants using the construction be + like as a quotative claim that “it prefaces not direct retrospective reports of speech, but internal speaker reactions – what the speaker had in mind to say but did not, or how the speaker felt at the time” (44). Schourup presents an interesting comparison between like and particles in five different languages which, as the English like, function both as particles meaning like and as direct speech introducers. According to Schourup, this functional correspondence supports the analysis of the origin of the quotative use of like as a way of conveying “the approximate nature of direct discourse reports” (47).

The use of be like as quotative is noted also by Tannen (1986, 1989) in her work on constructed dialogue. For Tannen, the quotative be + like “functions as a formulaic introducer [...]. If the literal meaning functions at all, it is to suggest that the dialogue is not being quoted but simply represents the kind of thing that character was saying or thinking.” (1986:321) Thus, much in line with Schourup, Tannen underlines the use of like to introduce approximate reports of speech and thought.

In her comparison of the forms and functions of conversational and literary narrative in American English and modern Greek, Tannen (1986, 1988) found that the most frequent introducers in all four types of discourse under consideration – American and Greek conversation and fiction – were forms of the verb say. But, while in the Greek spoken stories forms of say constituted 71% of the introducers, in American spoken stories, forms of say accounted only for 43% of the introducers, a difference partly accounted for by the use, in American English, of be like, tell, and go, none of which has a counterpart in Greek. In Tannen’s study go and like accounted for 21% of the English introducers (13% and 8%
respectively), while unintroduced dialogue (Yule’s “zero quotatives”) accounted for 26% in the American stories and for a comparable 22% in the Greek stories.

Blyth et al.’s (1990) small scale study based on interviews to 30 informants in New York is the first in-depth analysis of the discourse function of *be like*. Pointing out that, while most quotatives introduce either direct speech or inner monologue, the new quotative *be like* can introduce both kinds of reported speech, Blyth et al. argue that the quote which follows *be like* may represent a thought, a state of mind, or inner monologue, and therefore may be interpreted as never having been uttered” (222). They contend that the discourse function of *be like* as an introducer of thoughts, states of being, or inner monologue is related to its aspectual difference with other quotatives. Quotatives usually introduce an instantiation of speech which is typically conceived as a punctual or perfective event. While with the quotatives *say* and *go*, the aspectual difference is grammaticalised so that the progressive form marks continuous aspect, with *be like* aspect is not grammaticalised (*be* is a stative verb) and is highly context-sensitive. Blyth et al. claim that *be like* may be either perfective (introducing a completed and punctual speech act) or imperfective according to its discourse function: “when *be like* is interpreted as imperfective it introduces a thought, inner monologue, or a gestalt which summarizes the speaker’s frame of mind; when perfective it introduces direct speech” (222). Blyth et al. argue that *be like* reconciles “the apparent conflict between a perfective speech act and an imperfective state of mind”, typical of narrative discourse, where imperfective actions or states form the descriptive background for the foregrounded perfective actions (223). They make an interesting comparison between *be like* and the quotative *think*: as *think*, *be like* allows the speaker to present thought as direct
speech, thus rendering it more salient; however, *be like* differs from *think* in that it may introduce either direct speech or inner monologue.

Finally, Blyth et al. found that *say* and *go* behave similarly in that they are both predominantly used with the third person singular, while *be like* is rarely used in the third person singular. This finding is important, as we will see, not only in the light of later quantitative studies of the use of *be like* as direct speech quotative, but also in the light of findings of this present study.

Romaine & Lange’s (1991) is the first and up to now only exhaustive study of the historical origins of *be like* as an introducer of speech and thought. Their data consists in a total of almost 80 instances of *like* used in the new function of quotative of speech and thought, from recordings of teenagers and adults, and from various media sources. Although the bulk of the study is a reconstruction of the process and mechanism of grammaticalization of this new use of *like* as a marker of speech and thought, the study also contributes significantly to the discussion of the discursive functions and the sociolinguistic distribution of *be like* as a quotative of dialogue and thought.

Romaine & Lange (1991) present a thorough reconstruction of the ongoing grammaticalization of *be like* as an introducer of speech and thought. Showing that languages differ greatly in the extent to which they distinguish grammatically between direct and indirect speech, they suggest that “discourses introduced by *like* blur the boundaries between direct and indirect representations of both speech and thought report” (235). Romaine and Lange identify two important functions of the use of *like*: (1) its use to demarcate roles in the speech event through the alternation between *like* and other verbs of saying, such as *say* and *go*; (2) its use to convey aspects of the speaker’s subjectivity, such as
feelings and thoughts. They argue that these functions can be best understood by applying Goffman’s (1981) notion of Footing, i.e. the social and dramatic roles of speaker and hearer in conversation. In this perspective, the alternation among go, like and say has the function of signalling changes in what Goffman calls Production Format, that is the different combinations of speaker and hearer roles. Romaine and Lange, however, point out that “the choice of a verb of saying has implications which go beyond the presentation of self and extends to consideration such as what one undertakes to say/repeat not only to others, but also of others” (243). More specifically, be like can be used to report the discourse of others when the speaker wants to convey little commitment to what was actually said or thought. Thus, the construction of be like not only blurs the distinction between the direct and the indirect mode of reporting, but also the distinction between speech and thought: “since there is a subtle and often fuzzy boundary between reported speech and thought, like is a good choice since it creates only an example of something that could have been said or thought without implying the kind of commitment that say does.” (263) 

Finally, Romaine and Lange provide a convincing explanation of the sociolinguistic distribution of be like and of the fact that it is most frequently found in the informal conversational style of young people, especially women: they argue that the reason why like as a quotative of direct speech is found mainly in the speech of women and young people lies in the fact that direct speech is a feature of what Tannen calls “involved conversational style”: dialogue is perceived as more vivid because the emotive and affective aspects of speech are difficult to incorporate in the indirect mode, dialogue creates involvement because it allows listeners to imagine the recounted action or speech rather than hearing about it. Romaine and Lange point out that the primacy of direct speech on indirect speech can be
seen also in children’s language development. Moreover, it has been proved that women’s group talk is characterised by the use of features creating affect and emotional involvement and the use of strategies reinforcing group communication rather than provoking conflict and competition. Thus, “to the extent that like can be used to convey thought at the same time as it allows the speaker some leeway in commitment to what was said by another, it is consistent with many of the goals and values associated with the female communicative style.” (269)

Ferrara and Bell (1995) present an updated and comprehensive account of the development of the sociolinguistic diffusion of the use of be like as dialogue introducer. The study is based on 284 instances of be like drawn from three corpora of tape-recorded narratives, collected in 1990, 1992, and 1994, from 405 informants in Texas, aged 8-86. Ferrara & Bell provide evidence that while, in 1990 (as documented by Romaine and Lange 1991), be like was used predominantly by college-aged females, by 1992 the quotative was used approximately equally by males and females under the age of forty. Beside demonstrating that the form has gained in frequency and has diffused well beyond the areas in which it has been attested by previous studies (Ohio, New York, North Carolina, Washington DC) and is spreading equally among various ethnicities, Ferrara and Bell provide quantitative evidence that the form is being grammaticalised for third-person as well as first-person quotation. Indeed, their study shows that while in 1992 third person usage of be like accounted for 33% of total usage, in 1994 third person usage accounted for 47% of total usage.

The most interesting part of the study, however, lies probably in the contribution to the discussion of the discourse function of the new quotative. Ferrara and Bell claim that the
fact that the use of *be like* has increasingly spread into third person usage shows that the discourse function of the quotative is also expanding to include *introduction of speech*. They argue that the ability of *be like* to introduce constructed dialogue, thought and quotable gestures is determined by variation and expansion from a "core, paradigmatic case" (282), and to better explain the paradigmatic case of *be like* they apply Goffman’s concept of “Response Cries”, i.e. “openly theatrical, conventionalized utterances meant to clearly document or index the presumed inner state of the transmitter”. (282)

Finally, there is evidence that the use of *be like* as quotative of direct speech is being expanded from the United States not only to other varieties of English, such as British and Canadian English (Tagliamonte & Hudson 1999), Scottish English (Macaulay 2001) and Irish English (O’Keeffe 2002), but also to other European languages, such as German. Golato (2000), for example, presents a detailed account of the German quotative *und ich so* (‘and I’m like’).

**Go**

The use of *go* as direct speech quotative (as in “So George comes at Louis with the knife and Louis goes, ‘Don’t cut me, don’t cut me, I’ll do anything, just out up that blade.’” Butters 1980: 305) was first documented in 1980 in *American Speech* by Butters (1980), who calls the phenomenon ‘narrative go ‘say’’. Butters points out that the fact that the use of *go* as direct speech quotative is of new formation is strongly suggested by the fact that it is not accounted for in the dictionaries he consulted; in addition, he says, the form is commonly used by young speakers.
Butters notes that *go* as direct speech quotative usually occurs in the present tense, even when past time is indicated, although the use in the past tense is also permitted. He further notes that the form has not spread to indirect reported speech and interrogatives, as sentences like *He went that you are crazy* and *what did he go?* are ungrammatical. Butters points out that *go* may also be used “to introduce the mimicking of bodily actions, gestures, and postures, and thus has a somewhat broader range than mere *say*” (305). Finally, Butters suggests a possible explanation for the etymological change that has allowed *go* to become a quotative of direct speech.

The use of *go* as direct speech quotative is taken up by Shourup (1982) who claims that “there is only one generally applicable verb in English that unambiguously cues listeners to the onset of a direct quotation, and that is the recent semantic extension of *go*” (148). Shourup means that when a quote is introduced with *say*, as in *John said, “I was responsible for Lauren’s failure.”*, since speakers are without the benefit of quotation marks, ambiguity may result: it is unclear whether in the preceding sentence it is John or the speaker who is claimed to be responsible for Lauren’s failure. Thus, according to Shourup, *go* solves a small problem in spoken English for those who use it. Shourup also claims that the fact that narrative *go* ‘*say*’ has not spread to indirect speech introduction is not surprising, as of it had it would lose its function as quote marker.

There is evidence that *go* as direct speech quotative is widely used in Canadian, British, Scottish, and Irish English. The quotative use of *go* is indeed accounted for in the *Longman Grammar of Spoken and Written English* (hereinafter, Longman Grammar, 1999: 1119), which includes examples of the quotative use of *go* both from the American and the British component of the corpus on which it is based (*Longman Spoken and Written English*...
(LSWE) Corpus); in O’Keeffe’s study of the use of reporting forms in a sub-corpus of young Irish women’s casual conversations from the Limerick Corpus of Irish English (O’Keeffe 2002); in Tagliamonte & Hudson’s (1999) study of the quotative system of contemporary British and Canadian youth; and, finally, in Macaulay’s (2001) study of the quotative expressions of Glasgow adolescents. Macaulay found that go is used more frequently in the past tense.

**Be all**

While the literature on be like is extensive, and other quotative forms such as go and zero quotatives have been accounted for in various papers, systematic accounts of the direct speech quotative be all are practically non-existent. The use of be all as direct speech quotative is however accounted for in the Longman Grammar, which includes a few examples taken from the American English component of the LSWE Corpus, such as *He was all ‘Well, I wanted to stay out of it.’* (AmE) (1120). Finally, Macaulay (2001: 6) cites a paper presented at New Ways of Analysing Variation (NWAVE)-28 by Igoe et al., who reported on 311 quotatives collected by an introductory sociolinguistics class at the University of Pennsylvania, 18 of which were of be all. Macaulay specifies that the examples of be all were all provided by Californians, and adds that there is no evidence so far that this form has been adopted further east.

---

Summary

In this chapter I have reviewed literature pointing to the limitations of sentence-based approaches to reported speech; I have discussed the literature on reported speech in conversation, and reviewed studies of quotatives.

This review of literature in reported speech and quotatives has shown that research on direct quotation in spoken English has been mostly based on conversation, and more specifically on narrative, or conversational storytelling. Only few studies (Baynham 1996, Myers 19990, indeed, have looked at direct speech in non-narrative discourse. Similarly, all studies of quotatives have been based on small samples of quotatives drawn from tape-recorded narratives from informants from one single dialect area (New York in Blyth et al. 1990), Washington DC in Romaine & Lange 1991, Texas in Ferrara & Bell 1995) and have only looked at the new quotative be like. In other words, there are no studies providing quantitative, empirical evidence of the use of the new quotatives go and be all, while quantitative studies of be like may not be considered representative of American English as they are not based on principled, representative corpora. Furthermore, none of the studies on direct quotation reviewed here has adopted a corpus-based approach.

Thus, this chapter points to two main gaps in the literature on direct speech and the new quotatives: 1) the need for studies looking at quotation and quotatives in other registers of spoken interaction, particularly registers of spoken Academic English; 2) the need for empirical, corpus-based studies of the new quotatives, based on principled, diverse corpora.
CHAPTER III: METHODOLOGY

In this chapter I introduce the methodology used for the analysis of the use of the new quotatives be like, go, be all in contemporary spoken American English. I first describe the corpora used in the study, and procedures of data retrieval, data selection and data storage. I then illustrate the types and procedures of quantitative analysis carried out in the study. Finally, I describe the framework I designed for the analysis of the discourse-pragmatic function of the quotatives, and specify the steps in the analysis of discourse function.

As mentioned in Chapter I, I look at present tense and simple past tense forms of the new quotatives be like, go, be all. In order to obtain a broader picture of their frequency of use and of their discourse-pragmatic use, however, I compare them with the more ‘traditional’ quotative say. In other words, I use the standard quotative say as an ‘anchor’ for comparison with new quotatives.

Also, I only look at simple present and simple past tense forms of be like, go, be all, say when used with grammatical subjects referring to persons; that is, I do not consider forms of be like, be all, go and say when used in association with existential it.

Corpora

This corpus-based study of direct speech quotatives is based on data drawn from three small corpora of contemporary spoken American English: Conversation, Service Encounters and Office Hours. All corpora include text files and are untagged. The corpora are described below. In order to provide a more detailed picture of the corpora, I also briefly overview the
corpora from which they are drawn, the Longman Corpus of Spoken and Written English, and the TOEFL-2000 Spoken and Written Academic English Corpus, focusing on the spoken component of these corpora.

1. The Longman Spoken and Written English (LSWE) Corpus: Conversation

The LSWE Corpus contains over 40 million words (40,025,700) of text and 37,244 texts mostly produced after 1980, representing four registers: News, Conversation, Fiction, and Academic Prose. The LSWE Corpus was originally designed to include 10 million words in each register, evenly distributed between subcorpora of British and American English. However, the final size of each subcorpus is slightly different from the initial target, and ranges from 2.5 million to 5.4 million words. The LSWE Corpus also includes two supplementary registers: non conversational speech and general prose.

The subcorpus for conversation is defined as “probably the most representative sampling compiled to date” (Biber et al. 1999: 28), as it is much larger than previous corpora of conversation and has been collected in “genuinely natural settings” (28). The British component of the conversational subcorpus is a part of the British National Corpus (BNC) and comprises 3,436 texts for a total of 3,929,500 words. The American component has been collected by Professor Jacques Du Bois and his team at the University of California at Santa Barbara and includes 329 texts for a total of 2,480,800 words. The sampling for both components was carried out following demographical criteria, i.e. a set of informants representing the range of English speakers in the country across age, sex, social class, regional spread, was identified. The informants were then asked to tape-record all their conversational interactions over a period of a week. The demographic sampling of the
conversation is what ensures the representativeness of the corpus, i.e. it ensures that the corpus be not skewed toward one particular regional dialect, conversational register, etc.

The American English component of the conversation subcorpus includes conversations from 491 participants, 292 females and 199 speakers; most speakers are aged less than 50 (112 are less than 21, 114 aged 21-30, 94 aged 31-40, 82 aged 41-50), while speakers over 50 are about a fifth of the sample (51 aged 51-60, 38 aged over 61) (30).

**Conversation**

The Conversation corpus used in the present study is a subcomponent of the American English component of the LSWE Corpus subcorpus for conversation. It includes 65 text files, for a total of approximately 425,000 words. As a subcomponent of the LSWE Conversation corpus, it may be considered representative of contemporary spoken American English.

2. **The TOEFL 2000 Spoken and Written English Academic Language (T2K-SWAL) Corpus: Service Encounters and Office Hours**

The T2K-SWAL corpus was compiled in 1998, under commission of the Educational Testing Services (ETS), as part of a major revision of the Test of English as a Foreign Language (TOEFL). Overall, the corpus consists of over 2.7 million words (2,737,168) and includes 423 texts. It was compiled with the primary goal of providing "a representative basis for test construction and validation" (Biber et al. 2001: 49), and was designed to fill the gap represented by the lack of representative text collections of academic registers, particularly spoken academic registers.
The spoken component of the corpus consists of about 1.65 million words (1,665,516) and the registers represented include lectures and interactive classes, classroom management activities, study groups, office hours consultations, and campus-related service encounters (Biber et al. 2001; Douglas & Nissan 2001). All register categories (lectures, classes, labs, etc.) include an even distribution of graduate- and undergraduate-level texts, and cover subdisciplines in six major disciplines: natural science, humanities, business, engineering, social science and education. The texts were collected at four different academic sites (Northern Arizona University, Iowa State University, California State University at Sacramento, and Georgia State University), which may be considered broadly representative of the dialects of four major regions in the United States: the West Coast, the Rocky Mountain region, the Midwest, and the Deep South. The goal behind this design choice was representing the various types of spoken academic language that students can expect to encounter at virtually any American university campus, and thus avoiding marked skewing that might result from sampling texts from a single university setting (Biber et al. 2001).

The participants in the SWAL Corpus were primarily students, who were recruited to record their academic conversations with faculty, but also, faculty, who were recruited to record office hours, and university staff, who were recruited from offices and areas that regularly interact with students (e.g., the registrar’s office, the university bookstore, the library, etc.) to record service encounters (Biber et al. 2001, 2002). More specifically, students carried tape recorders over a period of two weeks, to capture academic speech as it occurred in the class sessions and study groups that they were involved in; faculty, with
students’ permission, left the tape recorders running during their office hours (Biber et al. 2002).

**Service Encounters**

This corpus includes 22 files, for a total of approximately 97,600 words and is a subcorpus of the spoken component of the T2K-SWAL Corpus. It consists of university service encounters recorded at different settings on a university campus, including the university bookstore, the copy shop, the coffee shop, the front desk in the dormitory, academic department offices, the library information desk, the media center, and student business services. (Douglas & Nissan 2001; Biber et al. 2002). Service Encounters were recorded “wherever students regularly interacted with staff to conduct the business of the university” (Biber et al. 2002:20).

Both Office Hours and Service Encounters are classified as “consultations”, one of the three text types identified for the new TOEFL Listening Test. Douglas & Nissan (2001) define a consultation as

"an extended piece of discourse, taking place between a student and an employee or employees of a school, such as an instructor or administrative and service personnel. The content of the interaction is relevant to a specific course or coursework in general, [...]. A consultation is intentional, and is usually initiated by a student to address a specific need or to solve a particular problem. The relationship between the interlocutors is asymmetrical in terms of status and knowledge, with the school employee having greater status as an expert, the person who holds the knowledge of skill sought by the students, and the student being non-expert” (Douglas & Nissan 2001: 3-4).

Douglas & Nissan (2001) also expect the discourse of service encounters to be

“highly structured, especially in how they are opened and closed” (4). They also stress that
“as with office visits, the relationship between the interlocutors is asymmetrical, with the school employee being considered an expert and the student a non-expert” (4).

This description of the discourse of Service Encounters represents what the corpus designer were expecting to obtain with the recordings. However, even an happenstance analysis of the files contained in this corpus would show that a substantial part of the interaction recorded consists in casual conversation between co-workers, who in many cases are part-time student workers. The following excerpts, retrieved while searching for occurrences of I'm like, consist of extended chunks of casual conversation not intermingled by any “service encounter” between customer and clerk. In fact, the speakers are talking about issues relating to their personal life:

(1): Oh OK.
  2: They're gonna be here the tenth through the fourteenth. They get in on Thursday at ten.
  2: OK.
  2: And they'll be here the fourteenth and I know I'm not gonna pack and everything now. My best friend here moved from San Diego. [1 sentence] My uncle's coming from California so he's coming from San Francisco, so.
  2: Oh how fun!
  2: So I kind of wanted to have like a week after that to sort of just like pack my stuff up my mom's like, well it's so close to Christmas [[I'm like]], mom like I've gotta go through my stuff.
  2: Mhm.
  2: Cos I have to ship it all home. [373] [copier done]
  2: I'm just worried nobody's gonna be up here. It's just gonna be me.
  2: Katie's here.
  2: How long is she here till?
  2: I don't know. But she's here.
  2: I'm afraid I'm just gonna be like so, cos Josh I didn't know that he was gonna home and work.
  2: Is he?
  2: So it's just gonna be like, solo mission.
  2: When is he leaving?
  2: He doesn't know yet. So it'll be after finals [5 ...
(TOEFL 2000 T2K-SWAL Corpus, servenc_n130.txt)³

(2) 26. ... exhilarating [laughs]
  4: Bye

³ Permission to use the TOEFL 2000 SWAL Corpus was granted by Educational Testing Service, the copyright owner.
2: You know a few minutes you know on the phone when I need you know I need more than just hi how's it going oh I had a really good day oh you had a really good day OK what are you what's your schedule for tomorrow OK Great that's your schedule oh this is my schedule oh OK great well I'll talk to you later 4: Don't you hate conversations like that?
2: Yeah, it's like
4: (that's what I try to tell my boyfriend look talk to me like I'm human and that [[I'm like]] your friend I bet something happened interesting today or something your upset about or whatever talk to me
2: Right
1: Right
2: And he's all excited about coming up here he's like he's like oh I wanna go to Sedona we'll do so a picnic on Monday and we'll go to the Grand Canyon on another and this and this and this and I'm like I don't know you know I got you know
4: What does that mean I don't know
2: I don't know
4: It means I don't know if I like you anymore I don't know ...

(TOEFL 2000 T2K-SWAL Corpus, servendo_n002.txt)

The likelihood that Service Encounters may contain extended chunks of casual conversation, rather than the type of interaction described in Douglas & Nissan (2001) may clearly represent a potential limitation for this study: if it cannot be ensured that a large part of the interaction contained in the corpus consists of "real" university service encounters, it may be questionable whether the Service Encounters corpus can really be considered representative of a spoken register distinct from conversation.

In a study of the spoken and written registers of the TOEFL 2000 SWAL Corpus using a corpus-based technique called multidimensional (MD) analysis, the Service Encounters register was found to be the most highly involved in the TOEFL 2000 SWAL Corpus, as it is characterised by the co-occurrence of features with positive loadings on Dimension I (Involved vs. Informational Production) (Biber et al. 2002). Service Encounters was also found to be the most situation-dependant register, since it has highly positive scores

---

4 Permission to use the TOEFL 2000 SWAL Corpus was granted by Educational Testing Service, the copyright owner.
along Dimension III (Situation Dependant vs. Elaborated Reference), and the most highly marked register for nonimpersonal style, due to its dense use of features with positive loadings on Dimension V (Nonimpersonal vs. Impersonal Style) (Biber et al. 2002). These findings are consistent with and provide possible evidence for the likelihood that the Service Encounters corpus mainly consists of casual conversation among co-workers or college students and it thus may not be possible to consider it as a register distinct from general conversation.

**Office Hours**

This is the smallest of the three corpora used in the study: it includes 12 files, totalling approximately 50,400 words. Like Service Encounters, Office Hours is a subcorpus of the spoken component of the T2K-SWAL corpus. It includes one-on-one consultations between students and university instructors.

**Procedures**

**Data collection: simple and advanced searches**

The present and simple past tense forms of the quotatives under investigation were retrieved using MonoConc Pro (version 2.0), a commercially available, sophisticated concordance package, that is a computer program that generates concordances automatically.

---

5 "A concordance is an index to a word in a text." (Sinclair 1991: 170). Computer-generated concordances are also called 'Key Words In Context' (KWIC).
Both simple and advanced concordance searches were used for the retrieval of quotative forms. For example, in order to search for the simple present tense forms of the quotative *be like* in association with a first person singular subject, a simple search for the form *I'm like* is sufficient, as there is no other grammatical subject but *I* that goes with "'m like". In the search field of the concordancer it will thus be sufficient to enter the search "I'm like". On the other hand, in order to search for the simple present tense of *be like* used with third person singular subjects referring to persons only, a simple search would not be adequate, as there is not one single subject that goes with the form "'s like" ("s like" can be preceded by *it, he, she, everybody*, and any first name). Using wildcards, i.e.: searching for "*'s like" (where the asterix is the wildcard character which matches zero or more characters), is a possible, yet not completely satisfactory option: a search for "*'s like" will indeed also yield such undesired forms as *that's like, there's like* and *it's like*. The best solution is thus using an advanced search. Staying on the example of third person singular present tense forms of *be like*, a possible advanced search would be "[^t]’s like".

Advanced searches are also useful in cases in which the same verb form may be used in association with several different grammatical subjects. For example, the singular simple past tense form of *be like*, namely *was like*, may be used in association with first and third person singular grammatical subjects (*he, she, it, everyone, everybody, first names*, etc.); similarly, the plural form *were like* can be used in association with second person singular (*you*), first person plural (*we, James and I, all of us*, etc.), second person plural (*you*), and third person plural (*they, Sarah and Jennifer*, etc.).

---

6 It was found that the non contracted form *I am like* is very rare in the corpora used in this study. For example, *I am like* occurs only 3 times in the Conversation corpus, against the 106 occurrences of *I'm like*.

7 The NOT symbol '^' ensures that all words ending with the character following it will be excluded.
Simple searches for *was like* and/or *were like* would yield concordances of these forms in association with all these different grammatical subjects, making it extremely difficult and complicated to keep track of the use of quotatives in association with different grammatical subjects. Therefore, in order to keep track of quotative use in association with different grammatical subjects, I used mostly advanced searches. For example, in order to generate concordances of the simple past of *be like* in association with third person singular subjects only, I used the following string:

```
[^it]\Wwas\b\Wlike\b (or simply: [^it]\Wwas\Wlike)
```

This string ensures that forms of *was like* preceded by *I* and *it* will be rejected.

Clearly, one weakness of this type of search is that it will also exclude forms of *was like* preceded by any first name ending in *i* or *t*. However, I see this drawback as a necessary compromise for the separation of quotatives used with different grammatical subjects, as it did not affect the analysis in any significant way.

**Retrieval of “real” quotative forms**

While the use of simple and complex searches allows the retrieval of all forms in which *be like, go, be all, say* are used in association with the specified grammatical subjects, neither simple nor complex searches will ensure that the yielded forms are forms in which *be like, go, be all* and *say* actually introduce direct speech and are thus used as quotatives. For example, a search for *I'm like* will also yield examples such as:

3) *I know, this was funny though but I said so do you ever see him, do you ever get to meet him Ted Casablanca she said I forgot, I guess not really but you know who she did is like Chris*

---

8 The meta-characters `\b` and `\W` are word-delimiters. `\b` is a simple word delimiter, while `\W` stands for any symbol that is a word delimiter, that is it covers all the characters listed as “word delimiters” in the MonoConc Pro’s “search options”, along with space, return, etc. (Barlow 2000: 60)
Connelly h[[e's like]] the editor of Premiere and um she says that he's really nice (Conversation)

4) it was just like he's just like the nicest guy, he's so patient, he never, I've like never seen him mad, when he's happy h[[e’s like]] quietly happy it's really funny (Conversation)

5) I kind of wanted to get there a little early to walk around Sausalito and we'd just be rushing there and then like you know and then so, we're doing our big dinner Thursday night so tomorrow [[we're like]] exchanging gifts and all of that. (Conversation)

6) Well they are kind of but [[they're like]] total innocent couple. (Conversation)

7) <1720>And you know [[they're like]] about that thick. (Conversation)

8) They, they don't even show until [[they're like]] maybe um five, six months pregnant. (Conversation)

9) Yeah Brian my old boss showed me like ... cause h[[e's like]] all into blues and stuff like that. (Conversation)

10) I mean this isn't supposed to be what, he was getting all whiny about that, one of the younger guys and then h[[e's like]] saying and you know I would really like it if we knew in advance what was going to be happening, what we were going to be talking about if someone would like make out a schedule and and inform everybody what we were going to be doing in the meeting (Conversation)

11) <2072> Oh. You slept with him? <2080> Oh god we used to sleep together for about a month and a half, h[[e's like]] just crazy. <2072> Girl you're a tramp. (Conversation)

In these examples be like is used as a discourse marker and does not introduce direct speech. In some cases (for ex. 7, 8), like means ‘approximately’, but in other cases like is semantically empty; it could be removed and the overall semantics of the utterance would remain unchanged. Underhill calls this use of like ‘focus like’ as like functions as marker of new information and focus. (Underhill 1988; See also Meehan 1991).

Because it is not possible to avoid such undesired forms automatically, i.e.: using a concordancer, human intervention is necessary to refine the output and eliminate undesired forms. The combination of automatic searches and human intervention for disambiguation is what makes this type of technique semi-automatic.
Excluded forms

*It’s like, it was like*

Forms of the quotative *be like* in association with existential *it*, i.e.: forms of *it’s like* and *it was like* were excluded from the analysis and will not be dealt with in this study.

These forms were excluded from the analysis presented in this study for two reasons. First, at a first glance at the examples of *it’s like*/*it was like* introducing (supposed) direct speech quotation, it is apparent that *it’s*/*it was like* is used in a different way, if nothing else because the quotation is unattributed to a speaker:

12) I was going to buy him a souvenir 'cause I thought I would see him 'cause [[It’s like]] they’ve broken up but they still kind of remain like in touch and I, you know, it’s like he’ll still come up and you know when I'm down there he’ll still like come with her to pick me up at the airport. (Conversation)

13) He was at the door, with his hand in the pizza bag, waiting for somebody to answer and [[It’s like]] yeah, I wouldn’t go answering a door too if it’s like that. (Conversation)

14) And Kim and I were talking to him and we walked away Kim’s like that is one ugly S O B. I never knew that. I’m like, how didn’t you know that? [[It’s like]], god, I never realized how ugly Steve is. Okay this isn’t really being mean. It’s like how, he’s ugly. (Conversation)

15) I was like quit being so goddam anal, I mean <2076> <nv_laugh> <2078> [[It’s like]] jeez, I mean he's getting so mini about everything, I was just like uh this is just a group of friends getting together to invest money, you know, this is not like a business meeting so fuck off. (Conversation)

16) he was just worried about the timing so I said no, it’s April and I told him to it and he was so shocked, [[It’s like]] oh I’m really disappointed, he said I was really optimistic we'd be able to get it through. (Conversation)

17) We got home like last night like at one thirty and the hamburger that we bought was still sitting on the couch. <1720> Oh. <1722> [[It was like]] um put this in the freezer, see how it turns out. (Conversation)

18) You wanna go <unclear> with your parents <unclear> <2123> <nv_laugh> <2125> And go, and they go oh <unclear> [[It was like]] don’t pay any attention. (Conversation)
The use of *it's like* has been pointed to as different from the use of *be like* with animate subjects by Schourup (1985: 59 ff.) and Ferrara & Bell (1995: 278). Schourup suggests that the *it* of *it's like* refers to what the speaker has in mind to express. (58) Ferrara & Bell claim that *it's like* is used "either chiefly by males to report the collective thoughts of a group, [...] or by both males and females to report a habitual style of thought or speech for one individual" (278). The use of *it was like* has not been pointed out in any of the studies on *be like* mentioned in Chapter II.

The use of *it's like*/*it was like* seems to be pragmatically different from the use of *be like* in association with animate grammatical subjects, and seems to deserve more attention than it has received so far. However, the treatment of such forms would have been beyond the scope of this study.

Secondly, *It's like* has no counterpart in the other new quotatives considered in this study: forms of *it goes, it's all*, etc. do not seem to introduce direct speech quotation.

*Conditional, past progressive with gonna and going to, future forms*

The search for simple present tense forms of *say* yielded a number of occurrences of forms such as *I was gonna say, I would (just) say, I was going to say, I wanted to say, I'll say, she'd say, he'd say, you might say, she'll say, I was about to say, you might wanna say*, etc.

In these forms the quotative *say* is preceded by modal and semi-modal auxiliaries, or other 'epistemically modalised' expressions (e.g., *about*), forming conditional (*would say*), future (*will say*) and progressive forms (*going to/gonna say*). These different quotative formats introduce different types of quotations, ranging from hypothetical quotation, to
quotation projected in the future, to quotation expressing habitual and past events. For example, in (19) and (20), the quotative introduces a quotation expressing habit:

19) *And she'd [[say]] well it's been two weeks I haven't even seen Jay just you know.* (Conversation)

20) *I mean she would go off and say something that had nothing, had something to do with what we were talking about but, like I have to ask her are you a veteran. And she would [[say]] well you know I really believe that if I just try hard enough that I can do something with my life and it's about time that I take a new step and a new face and then she would stare at me for like forty five seconds straight.* (Conversation)

In (21) and (22), instead, the quotative introduces a hypothetical quote:

21) *Cause all students should be able to call their professors <laughing> time to time. */laughing*/ Get book information. <1533> Yeah. <1534> But then you have that faculty that'll [[say]] you know I'm not working during the summer, don't bother me. <1534> Yeah.* (Conversation)

22) *They get up at, oh at seven they have to stand at attention to get letters and they do twenty-five pushups for every letter they receive. For every letter. I would [[say]] dear <unclear>, give me a tooth, don't write.* (Conversation)

This use of direct speech quotation has been taken up by Myers, who calls it “unspoken discourse” or “hypothetical reported discourse” (Myers 1999). Myers distinguishes three types of hypothetical reported discourse: imaginary, possible or conditional, and impossible or contrary to fact. Broadly speaking, these forms introduce hypothetical, and, thus, counterfactual, discourse.

In a number of cases, *say* occurs as part of progressive forms, as in the examples below:

23) *I know, I was going to [[say]] yeah he does look really skinny.* (Conversation)

24) *I was going to [[say]], they keep going down and sucking <unclear> I still think you've got to get some new sucker fish.* (Conversation)

25) *Oh yeah so it was like the lady I was waiting for she's like Amanda I'm so flat, and she's like I got these and she was like they have so much padding and I'm like I was about to [[say]]*
yeah my friend Alyssa told me those new ones out there had so, that's not that, well (Conversation)

26) I was thinking, well see 'cause it's like I know I could ask Lisa but she's working and I know she's going to [[say]] don't get it anyway. (Conversation)

The quotation format "[present/past tense] gonna/going to + say" seems very interesting, especially in light of the fact it has not been taken up in the literature on reported speech so far. Carter & McCarthy (1995; See also McCarthy 1998, McCarthy & Carter 1995) were the first to point out the use of the past continuous (or past progressive) forms *was/were saying/telling* to introduce indirect speech reports. These examples, however, are different from the type of progressive forms accounted by Carter & McCarthy in two ways: (1) the quotative format includes past or present forms of *going to*; (2) the quotative introduces direct speech, not indirect speech.

A more in-depth analysis of the function of this quotative format was beyond the scope of this study. It seems however that these quotative formats (at least in the examples above) introduce what Myers calls possible or conditional hypothetical or unspoken discourse (Myers 1999: 576-577). These forms were eliminated and not included in the study because their inclusion would have called for searches of hypothetical or unspoken speech introduced by comparable forms of *be like, be all, go*; i.e. for the search of quotative forms such as *I would be like, I would be all, I would go*, etc., thus adding a new dimension to the study. A quick search for such forms revealed that they are quite active in association with *be like* but not with *go* and *be all.*
Data storage

Once all undesired forms had been eliminated, the refined output of simple and/or advanced searches was saved as text file. The quotatives were saved with a surrounding context of 500 characters, in order to ensure enough context for the analysis of their discourse-pragmatic function. All text files were then imported into a computer program (FileMaker Pro) for database management. The use of a computer database management program proved extremely useful to keep track of the qualitative analysis of the quotatives discourse function in an efficient manner. The use of this program also allowed to keep track efficiently of further refinements of the output: although the concordances of the quotatives had been carefully scanned manually in order to eliminate the occurrences of be like, go, be all, say in which the verbs are not used to introduce direct speech, while carrying out the discourse-pragmatic analysis of the forms saved on text files and imported into FileMaker Pro, a few forms had to be eliminated as, on a closer analysis, it was found that they were used to introduce indirect speech rather than direct speech, or, the quality of the transcription was not sufficiently good to make any inference about the discourse-pragmatic function of the quotative.

Analytical procedures

Quantitative Analysis

The quantitative analysis is the bulk of this study, as it is through the analysis of the frequency of occurrence of the new quotatives be like, go, be all, that we can best determine the status of their use in the quotative system of spoken American English. The quantitative analysis includes three steps:
1. Analysis of the frequency of the quotatives within each register. The total frequency of use of the quotatives within each register was obtained adding all simple present and simple past tense forms of the quotatives, used in association with all grammatical subjects referring to 'animate' subjects (as mentioned above, forms of it's like, it was like were rejected). The analysis of the frequency of use of the quotatives in the simple present and in the past tense was also carried out.

2. Comparison of the frequency of use of the quotatives across registers. Because the corpora used in the study are of unequal size, the figures representing the frequency of the quotatives in the different registers were normed. The quantitative findings were normed on a sample of 100,000 words.9

3. Analysis of the frequency of use of the quotatives in association with different grammatical subjects. The analysis of frequency of use in association with the grammatical subjects (1st person singular, 2nd person singular, 3rd person singular, etc.) is important in the light of previous studies of the grammaticalization of the new quotatives arguing that the spread into third person singular usage suggests expansion of discourse function (Ferrara & Bell 1995)

**Procedures in the quantitative analysis**

After searching for the simple past and simple present tense forms of the new quotatives be like, be all, go and of the new quotative say, hand editing the output of the searches, saving the final output resulting from automatic searches and manual disambiguation in separate files, and then importing these files into a computerised database

---

9 That is, the number of occurrences were divided for the number of the words of the corpus (for example, 425,000 words in the case of occurrences in the Conversation corpus), and then multiplied by 100,000.
program, I proceeded to the quantitative analysis of the quotatives. More specifically, the quantitative analysis of the quotatives involved the following procedures.

First, I added up all simple present and simple past forms (in all grammatical persons) for each quotative, in order to calculate the total frequency of that specific quotative within the register.

For each quotative, I then calculated the frequency of use in the simple present tense and in the simple past by adding all simple present forms in all grammatical persons, and all simple past forms in all persons separately, thus obtaining the raw frequency of the quotative in the simple present and the simple past. I then calculated the relative frequency (i.e., proportion of use) of the total number of occurrences of each quotative (i.e., the sum of all simple past and simple present tense forms) and the relative frequency of use in the simple present and the simple past, against the total quotative use (i.e., the sum of all quotative forms considered in the study). I repeated this procedure for all quotatives, and for all corpora. This procedure was carried out in order to determine the frequency of occurrence of the quotatives within the quotative system as a whole, or at least within the set of forms considered in the present study.

In order to compare the frequency of use of the quotatives in the simple present, simple past, and the sum of simple present and simple past across the three corpora, I normed all the figures on a sample of 100,000 words. For example, in the Conversation corpus there are 193 occurrences of simple present tense forms of *be like*. By dividing 193 for the total number of words of the Conversation corpus (i.e., 425,000 words), and multiplying the result by 100,000, we will obtain the normed frequency of simple present tense forms of *be like* on a 100,000 words sample: 45.4 (See Table 4.4). I then calculated proportions of use against
total quotative use: staying on this example, this is done by the multiplying the normed frequency of simple present tense forms of be like (45,4) by 100,000, and then dividing the result for the normed total quotative use (300,1). The result (15,1) represents the proportion of simple present tense forms of be like against the total quotative use considered in this study.

Finally, I calculated the frequency of use of each quotative in association with grammatical person by adding all simple past and simple present tense forms in each person, e.g., all forms of I was like and I'm like were added together; forms of you were like were added to forms of you are like etc. The total frequency of use of each quotative in association with grammatical person was used to calculate the proportion of use of the quotative with grammatical person against the total use of that specific quotative, e.g., I calculated what proportion of the total use of be like, the sum of I am like and I was like accounts for.

I now proceed to the description of the second type of analysis carried out in this study: the analysis of the discourse-pragmatic function of the quotatives.

**Qualitative analysis**

The purpose of the qualitative analysis of the quotatives was determining which of two types of discourse-pragmatic function (i.e., introducing plausible speech vs. introducing improbable speech) the quotatives fulfil. The function of the quotative is related to the nature of the quotation, and more specifically to the level of plausibility/improbability of the quotation. In other words, I argue that introduction or quotation of plausible speech is a function, and this function is different from introduction or quotation of improbable speech.
In order to classify the quotatives according to their discourse function, I designed a framework that aims at capturing the main distinction between plausible and improbable quotation, while leaving other, more subtle distinctions unaccounted for. There are two main reasons for adopting a simplified framework for the classification of the quotatives. First and foremost, the main goal of this study is providing a description of the use of new or non-traditional quotatives in contemporary spoken American English; it is not that of providing an account of the nature of direct quotation itself. The fact that the focus of the study is not on quotation itself, but on the quotatives, justifies the choice of a simpler framework. Secondly, because this is a corpus-based study, the amount of data under examination is much larger than the amount of data used in previous, non-corpus based studies of direct speech quotation in conversational storytelling. Previous studies of direct quotation, as well as previous quantitative studies of the new quotatives (See, for example, Blyth et al. 1990, Ferrara & Bell 1995) have been based on transcriptions of recordings of conversations of small numbers of participants, and have thus dealt with much fewer occurrences of quotatives (e.g. 284 occurrences in Ferrara & Bell's (1995) study). A thorough, in-depth analysis of each occurrence of direct quotation introduced by simple present or simple past forms of the quotatives under examination would require the examination of extended chunks of discourse, and would be beyond the scope of this study.

The discourse-pragmatic analysis of the quotatives was limited to the quotatives retrieved in the Conversation corpus. There are two main reasons for limiting the discourse-pragmatic analysis to one of the three corpora used in this study and specifically to Conversation: (1) while it is not the main focus of the study, the analysis of the discourse-pragmatic use of the quotative is an extremely time consuming task; (2) the Conversation
corpus is by far the largest corpus used in the study, the only one that has yielded occurrences of all the forms searched for this study. The other two corpora, Service Encounters and Office Hours, for example, have not yielded any occurrence of the quotative be all. The Office Hours corpus did not yield any occurrence of go either. Thus, limiting the analysis of the discourse-pragmatic function to the Conversation corpus only was not considered as a limitation to the study.

In designing this framework for the analysis of the discourse-pragmatic function of the quotatives, I drew mainly from research by Tannen (1986, 1989), Macaulay (1987) and Mayes (1990) on direct quotation in conversation. Therefore, before proceeding to the definition of the framework for the analysis of the discourse-pragmatic function of the quotatives used in this study, I briefly review points in Tannen’s, Macaulay’s and Mayes’ work relevant to the definition of the framework.

As mentioned in Chapter Two, Deborah Tannen argues that all quoted speech or direct quotation in conversation is in fact “constructed dialogue”, “just as surely as is the dialogue created by fiction writers and playwrights” (1986: 311). She indeed compares the use of direct quotation in conversational narrative and literary discourse (fiction and drama), and identifies a number of different “types” of constructed dialogue (see Chapter Two for a full list) which demonstrate her argument that direct quotation in conversation cannot be a verbatim report of what someone said, and is instead just a construction of the reporter.

Tannen’s argument partly relies on studies reporting on experiments on human memory showing that, even in the most favourable circumstances, verbatim recall is faulty, as “humans cannot keep in their minds the precise words they have heard, even for a moment” (Tannen 1986: 313).
The idea that direct quotation in conversation (or spoken language in general, for that matter) can hardly be a verbatim report and therefore is reconstructed by the speaker is accepted here. However, while it is very unlikely – as Tannen points out – that speakers can make a verbatim report of what they have heard, it has to be recognised that there may be different levels of plausibility of the quote, and that, as Macaulay (1987) has noted, there is a continuum going from plausibility to impossibility, or, as Mayes puts it, “from quotes which could be authentic renditions to those which are, unquestionably, invented” (Mayes 1990: 331).

Mayes divides this continuum into four categories (plausible, improbable, highly improbable, and impossible quotes), and uses two types of evidence to place a quote in a category: internal evidence and situational evidence. She defines internal evidence as “lexical or syntactic evidence inside the quote itself or inside the utterance containing the quote that indicates that the quote is (or may be) invented”, and situational evidence as “evidence from the entire discourse situation casting doubt on the quote” (331).

**Framework for discourse-pragmatic analysis of the quotatives**

The framework used in this study combines and simplifies Mayes’ (1990) classification of quotes and Tannen’s (1989) analysis of constructed dialogue, with the goal of accounting for the major, broad distinction between introduction of plausible and introduction of improbable quotation. I now proceed to the definition of the framework.

The distinction between plausible and improbable quotation varies according to the number of the grammatical subject to which the quotation is attributed to. For example, as Macaulay (1987: 23) and Tannen (1989:113) have noted, a quotation attributed to a plural
subject, such as they, is highly improbable, as it is highly unlikely that the people referred to by the subject of the quote spoke in unison like a Greek chorus. This view is accepted here, and the direct consequence of it is that quotations introduced by a quotative referring to a plural subject may only be improbable and not plausible. Quotations introduced by quotatives referring to a singular subject, on the other hand, may be both plausible and improbable. This main, broad distinction between quotatives referring to a singular grammatical subject, and quotatives referring to a plural grammatical subject determines a classification of the functions of the quotatives represented in the table below:

Table 3.1: Framework for the discourse-pragmatic analysis of the quotatives

<table>
<thead>
<tr>
<th>Grammatical subject associated with the quotative</th>
<th>Type of quotation introduced by the quotative</th>
<th>Functions of the quotatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>First person singular subject (I)</td>
<td>Plausible</td>
<td>1. Quoting one's own speech</td>
</tr>
<tr>
<td></td>
<td>Improbable</td>
<td>2. Quoting one's own thought</td>
</tr>
<tr>
<td></td>
<td>(Inner speech/thought; gist of something)</td>
<td></td>
</tr>
<tr>
<td>Second person singular subject (You)</td>
<td>Plausible</td>
<td>3. Quoting the speech of a single person you are talking to.</td>
</tr>
<tr>
<td></td>
<td>Improbable</td>
<td>5. Quoting the gist of some speech or thought of a single person you are talking to or a third party.</td>
</tr>
<tr>
<td></td>
<td>(Inner speech/thought; gist of something)</td>
<td></td>
</tr>
<tr>
<td>Third person animate singular subjects (he, she, mom, etc.)</td>
<td>Plausible</td>
<td>4. Quoting the speech of a singular third party</td>
</tr>
<tr>
<td></td>
<td>Improbable</td>
<td>5. Quoting the gist of some speech or thought of a single person you are talking to or a third party.</td>
</tr>
<tr>
<td></td>
<td>(Inner speech/thought; gist of something)</td>
<td></td>
</tr>
<tr>
<td>First person plural subject (we)</td>
<td>Improbable (Greek chorus)</td>
<td>6. Telling the gist of some speech or thought of a group of people.</td>
</tr>
<tr>
<td>Second person plural subject (you)</td>
<td>Improbable (Greek chorus)</td>
<td></td>
</tr>
<tr>
<td>Third person plural subject (they, James and Paul, etc.)</td>
<td>Improbable (Greek chorus)</td>
<td></td>
</tr>
</tbody>
</table>
The broad classification presented in the second column of Table 3.1 may also be taken to a finer level: as shown in the third column, the 9 functions specified above can be merged into the following six functions of the quotatives:

1. quoting one's own speech (plausible quotation with first person singular subject);
2. "quoting" one's own thought (improbable quotation with first person singular subject);
3. quoting the speech of a single person you are talking to (plausible quotation with second person singular subject);
4. quoting the speech of a singular third party (plausible quotation with third person singular subject);
5. quoting the gist of some speech or thought of a single person you are talking to or a third party (improbable quotation with second and/or third person singular);
6. telling the gist of some speech or thought of a group of people (improbable quotation with all plural subjects).

For the purpose of the classification of quotation as plausible or improbable I use Mayes' internal and situational evidence. I now provide examples of quotatives introducing plausible quotation, and quotatives introducing improbable quotation, as well as examples of the six functions of the quotatives specified above.

1. **Quoting one's own speech (plausible quotation with first person singular subject):**

   (27) You want to call Eric? Actually he called me at like six thirty he's like like hello. Hi. [[[I'm like]]] what's wrong with you he's like I don't feel good. Why don't you go back to bed. I've been sleeping. Go back. No I can't sleep he said I have a headache, take something, what's wrong, they're still making my throat hurt. I don't want you to stay home and relax if you feel like you're going to puke then you're, I don't want you to come over for dinner and stuff and he goes no and I'm all just
come over and see me later on you know after bathing and stuff and he was like oh okay and I'm like ...

(28) I said I thought you knew, I tried to explain to her that I was waiting for somebody that's why I called you Nancy or whatever and she's like, oh, anyway, and like someone beeped and it was my landlord and [[I'm like]] <unclear> how are you doing and he like pulled up and like parked. <?>
<unclear> <?> What? <?> Anyway so like the Jehovah's witnesses said okay bye and they got in the car and left and then like my landlord comes up and says like what are they trying to sell you? And I said Jesus. (Conversation)

In both the examples above, the quotation introduced by I'm like is plausible in that it contributes to the advancement of the action narrated. Thus, there is situational evidence showing that the quotation was plausibly uttered in a form similar or close to the one reported.

2. Quoting one's own thought (improbable quotation with first person singular subject):

(29) I was looking like at the lemon and today 'cause you know how all the colors on sale are like yellow and purple= and like bright purple and orange and red, all the summer colors and [[I'm like]] I am so envious of people who can wear that stuff and get away with it 'cause I cannot get away with wearing anything like that. (Conversation)

(30) I want to but then it's like, it's um me and Chris's anniversary tomorrow and we didn't go out for dinner today. No, he called me at like eleven and it's like I'm really understanding 'cause it's kind of like I was kind of feeling upset to my stomach today in the morning anyway and I wasn't up to like getting ready for a big night on the town and he's just like well and I just kind of, I was thinking about it and [[I'm like]] it just makes sense because he isn't going to get off work until after six and it's like I kind of wanted to get there a little early to walk around Sausalito and we'd just be rushing there and then like you know and then so, we're doing our big dinner Thursday night so tomorrow we're like exchanging gifts and all of that. (Conversation)

(31) Well last year my sister said that I didn't have to buy for all the little kids if I didn't want to and [[I'm like]] yeah right, like you really mean it and then I always feel so awful because my one sister then she makes me feel good, or makes me feel bad because she'll say oh well look there's your uncle Michael, he, he doesn't have kids or anything and it's not like he can't afford, what does he do with his money, so she'll make comments like that and I just feel like shit< (Conversation)

(32) It was so sad. I went to the playground I can't use the stuff there, I'm too big. <1618> Which? <1615> At the zoo yesterday I was gonna go on the bar, monkey bars but [[I'm like]] either that or I'm too fat and I'm too heavy but I couldn't lift my leg, bend my legs and like swing myself from bar to bar]. <1618> [<laughing>I think it's called out of shape] Mandy. (Conversation)
In all of these last examples ((29), (30), (31), (32)), there is situational evidence suggesting that the quotation introduced by the quote I'm like was not uttered at the time of the reported situation/event and is thus improbable. The quotation is improbable because it represents what the speaker thought at the time of the recounted situation, rather than what he/she said. Clearly, there is no empirical evidence to demonstrate that, but our "knowledge of the world" suggests us that what the quote expresses is not the kind of thing that would have been considered appropriate to say in the reported situation.

In (30) the evidence is not only situational but also internal, in that I'm like is preceded by I was thinking and is probably a repair of the latter.

3. Quoting the speech of a single person you are talking to (plausible quotation with second person singular subject):

(33) ... who have bathing suits but you can't look at it 'cause I have to fix my boobs. <1615> Okay. Well it's kind of like it's really funny because it's like but then it's like so gross when I see people who should not be wearing bikinis and I know that they would say that about me. <1613> Okay look if I push it it looks like I have cleavage. <1615> Let me see. <1613> A little bit. <1615> Oh <nv_laugh> <1613> See <nv_laugh> 'cause I'm normally flat. <1615> It's like you tried it on to show me [[you're like]] see. <1613> <nv_laugh> I could actually <1615> <nv_laugh> I almost got a miracle bra. <1613> Did you? <1615> Okay, are you taking it off? <1613> Okay, I'll put my shirt on, I'll put my shirt on. <1615> Um <nv_clears_throat> because they were on sale but they didn't do anything for me and like <1613> Miracle bras don't do anything for me. <1615> I really want the kind of bra that you can like wear like really tight shirts, not tight, I don't wear like super tight shirts but I'm like cer (Conversation)

In (33) there is internal evidence that the quotative is introducing plausible quotation: the speaker introducing the quote (speaker's code <1615>) quotes part of the exact same wording ("See") of the preceding turn, which was uttered by her interlocutor (speakers' code <1613>).
4. Quoting the speech of a singular third party (*plausible* quotation with third person singular subject)

(34) 1720> He said something about getting your dog. I can't remember. <1722> I've thought of something. <1722> Oh. <1720> Oh it was that chocolate he had left over from making the cake. He was like he can have this I said no it'll make him sick. He goes why? I said I don't know why I just know that you're not supposed to give chocolate to dogs. <1723> It kills them. <1720> Yeah and I don't know <1722> Why. <1720> I don't know the clinical reason why but I know my vet told me not to give the dog chocolate, ice cream or pastries. (Conversation)

In (34) the evidence of the plausibility of the quotation seems situational: the quotation contributes to the advancement of the action narrated.

5. Quoting the gist of some speech or thought of a single person you are talking to or a third party (*improbable* quotation with second and third person singular subject):

(35) 1615> But I've heard people. <1621> Damn. <1615> Sunset sauna shop. Um, is you know how Chris <name> pretends to be that, that like wife who reads from the dining thing and then like Adam <name> who's her husband, his husband, have you seen that one? <1621> It's yeah. <1615> [[She's like]] well, da da da da da. <1621> Everything he does is funny. <1615> And like it's so funny 'cause I know he sounds like so many other people that I know. <1615> Are you tired? <1621> I just got tired. If I had not, uh, you have to come over here honey when I'm, no all the way. <1615> I can't. <1621> I can't meet you halfway. (Conversation)

(36) ... I'll have to tell you what that is. <?> <unclear> <?> Mac Daddy but if you're macking that means you're trying to pick up a girl but, if you are Mac Daddy that means you're a gigolo you're ladies you're a ladies man [[you go]] <mimicking> oh Mac Daddy <mimicking> and I'm macking to you. <unclear> <nv_laugh> <?> Mac Daddy <unclear> <?> There's a song called Mac Daddy. <?> There is? <?> I like that song <unclear> <?> Who does that? <?> Um some reggae um band <?> What's the name [of this?] <?> [I got] the tape ... uh ...(Conversation)

(37) ... w what I'm saying? <2072> Yeah, I think so. <2076> It's just, it's a mental picture where you go in and I would like go in and try on everything extra large or I'd, you know, I'd go in to buy a pair of shorts that I wanted to try on and I would take in my old size with me. Just automatically pick up the old size and then go you know like oh my god, they couldn't even, they would just fall down, they wouldn't even stay up on your hips and then you'd go in and I'd try the next size and [[you're like]] oh my god, these things are really mismarked because I've got to go down another size. <2072> Hmm. <2076> And you're trying on three and four and I'm like, I can't, no, I can't fit in thirty four's. This, you know <2072> Yeah I suppose it's hard to adjust to that mentally, you'd have to really <2076> Well I remember because I would, I'd walk into Emm's this is before, right before I started working at Emm's I think but you know I'd walk into the bar and I obviously wouldn't go on dri ... (Conversation)
In (35) there is internal evidence of the improbability of the quotation: the speaker uses the vague discourse markers *well* and *da da da da* to sum up the gist or the attitude of the speech of a third party. In (36) and (37) the speaker is mimicking some hypothetical speech of a hypothetical speaker. It seems that in both cases the second person singular pronoun *you* is used in its impersonal use.

6. **Quoting the gist of some speech or thought of a group of people (improbable quotation with all plural subjects)**

(38) ... out what they think about it before I commit. *I don't, see the thing of it is, soon as you sell, say Amway, people automatically go you. So when they see you coming and they're talking in a group they[y go] here comes that Amway idiot, quick everybody run. You know, I don't want that shit to happen.* <1003> *It's irrelevant if you want the money* <nv_laugh> <1001> *Yeah but I like what people think of me too.* <1003> *Yeah I really do* <nv_laugh> <1001> *Well I ...* (Conversation)

(39) ... think it's, it's important that each of us represent to the fullest, 'cause you're gonna represent whether you like it or not. *Once you get over twenty-one, young kids in your neighborhood look up and they see you and the[y go], that's my man up there, I'm a lookin', I'm a seein', I'm a watchin', see what he's doing to be successful. And if you selling drugs, they gonna emulate that, if you're a rapper, making money, they gonna emulate that, ...* (Conversation)

(40) <1722> *What's Marie Claire?* <1720> *The new magazine. It's that one that Billy had.* <1722> *Oh.* <1720> *You can get them since it's new for like twelve bucks for the year.* <1722> *That's cute.* <1720> *Actual size* <nv_laugh> <1722> *I know, I was thinking that. Your boom box is no bigger than our logo.* <1720> *[[They're like]] you know, don't open your shirt like that unless you've got some chest hair to show.* <1722> *Would you have guessed all this time he was paid for writing this?* <1720> *Hm. Are you sure?* <1723> *Look how many kids he has.* <1722> *She's his second wife.* (Conversation)

In (38) and (39) the quotation represents some hypothetical speech by some hypothetical group of people identified with a generic *they*. In all examples, the quotation is taken to be improbable as it is attributed to a plural subject, and it is improbable that people talk in unison as in a Greek chorus.
Procedures in the qualitative analysis

As mentioned above, I restricted the analysis of discourse-pragmatic function of the quotatives to the quotatives found in the Conversation corpus. Since the framework used for the discourse-pragmatic analysis of the quotatives used in this study determines that quotatives used in association with singular grammatical person may introduce plausible or improbable quotation, while quotatives used in association with plural subjects may introduce only improbable quotation, only quotatives used with singular person were actually analysed for discourse-pragmatic function.

The classification of the quotatives according to whether they introduce plausible or improbable quotation was carried out using the computer program for database management FileMaker Pro, which allows to create fields for different types of specifications, and to search for certain parameters by entering the code or labeling for that specific parameter. Thus, each FileMaker Pro file included at least two fields, one for the occurrences of the quotatives (field “example”), one for the classification of the quotatives according to discourse function (field “function”).

I analysed the discourse function of each occurrence of the quotatives by looking at the context surrounding the quotation, the content of the quotation, and applying the criteria specified above in the description of the framework for discourse-pragmatic function. Clearly, the classification of the discourse-pragmatic function of the quotatives according to whether the quotative introduces plausible or improbable speech was not always straightforward. An absolute, clearcut differentiation between quotation that represents plausible speech and quotation that represents inner speech or thought, or the gist of some speech, is indeed unfortunately not possible in all contexts. In cases of uncertainty, I looked
at whether the quotation is contained in a sequence of reported dialogue; I also looked more closely at the content of the quote, trying to establish to what extent the quote has some propositional content that somehow contributes to the advancement of the story-line. If these two conditions were satisfied, the quote was considered plausible. If the quote was part of a sequence of reported dialogue, but did not have propositional content contributing to the advancement of the story-line, it was considered improbable. An example of uncertain or ambiguous case is provided in (41) below:

41) <?> What do you mean? I'm no friends of any law enforcement organization.
<?> No friend of any law enforcement organization.
<?> <unclear>
<?> Like <unclear> these people that like they were just customers where I worked when I was a waitress and they invited me to go out with them. So I went out with them for sushi and the guy all of a sudden [[goes like]] well I'm like uh, I work for the D E A. <nv_laugh>
<?> That was like weird.
<?> He was like telling us about how like um,
<?> You don't have to fill out all that if you don't want to.
<?> Oh.
<?> <unclear>
<?> That's not necessary.
<?> Okay.
<?> They said just the first green one is fine.
<?> Okay cool.
<?> That's all I need.
(Conversation)

In (41), the quotation introduced by the quotative goes is part of a brief sequence of dialogue. However, it does not seem to contribute to the advancement of the story-line, as it only includes the discourse marker well, which has no particular content. In fact, the relationship between the quotation introduced by goes, and the quotation following it (I'm like uh, I work for the D E A) is all but clear. The quotation introduced by goes seems to be a representation of the attitude of the speaker, or of the gist of some speech, rather than a representation of plausibly uttered speech, and was thus classified as improbable.
Once a quotative was analysed and a decision about its discourse-function was made, the quotative was classified by entering the label “plausible” or “improbable” in the field “function”, situated below the field containing the example. When a quotative was found not to introduce direct speech, it was classified accordingly, and eliminated from the final counts.  

Once all the quotatives in the Conversation corpus had been carefully analysed, I calculated the total frequency of occurrence of the quotatives introducing plausible and improbable quotation for each grammatical person (i.e., first, second, and third grammatical person). This was easily done by setting the FileMaker Pro file in the “search mode”, and entering searches for “plausible” and “improbable” in the field “function”.

Finally, I calculated the proportion of use of the quotatives to introduce plausible and improbable quotation in association with grammatical person against the total frequency of the quotative in that specific grammatical person. For example, in Conversation, there are 163 occurrences of *be like* used in association with first person singular subjects. Of these, 45, or 27.7%, were found to introduce plausible quotation, and 118, or 72.4%, were found to introduce improbable quotation.

---

10 Of course, in this case, the counts of the qualitative analysis were also modified accordingly.
Summary

In this chapter I have described the corpora used in this study, procedures of data retrieval, selection, and storage. I have illustrated the two types of analyses carried out in the study, outlined the framework used for the analysis of the discourse-pragmatic function of the quotatives, and specified the procedures in the quantitative and qualitative analysis.

I now proceed to Chapter Four, where I will present and discuss the results of this study. The discussion of the results will follow the order of the analytical procedures described in this chapter: first, I will present the results of the quantitative analysis of frequency within each corpus, of the comparison of frequency across corpora, and of the frequency of use in association with grammatical person; then I will discuss the results of the analysis of the discourse-pragmatic function of the quotatives found in the corpus Conversation.
CHAPTER IV: RESULTS AND DISCUSSION

In this chapter I present and discuss the results of the analysis of the new quotatives carried out following the procedures described in Chapter Three. I first present the results of the quantitative analysis, and then move on to the discussion of the analysis of the discourse-pragmatic function of the quotatives.

Quantitative analysis

As specified in Chapter Three, the quantitative analysis of the quotatives be like, go, be all (and the "anchor" say) comprises three main steps: 1) analysis of frequency of the quotatives within each register; 2) comparison of the frequency of use of the quotatives across registers; 3) analysis of the frequency of use of the quotatives in association with different grammatical subjects. The discussion of results will follow this same order.

1. Analysis of frequency of the quotatives within each register

Conversation

The frequency and distribution of the quotatives in Conversation is displayed in Table 4.1, which shows the frequency of simple present tense of the quotatives, the frequency of simple past tense forms, and the total of simple present and simple past tense forms. The table specifies the raw and relative frequency of the quotatives, i.e. raw numbers resulting from the final output of the searches, and the proportion that these raw numbers represent against the total quotative use in the corpus (e.g., the sum of all the quotative forms considered in this present study retrieved in this particular corpus).
Table 4.1: Frequency of use of quotatives in Conversation corpus

<table>
<thead>
<tr>
<th>SAY</th>
<th>BE LIKE</th>
<th>GO</th>
<th>BE ALL</th>
<th>Total quotative use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Present</td>
<td>Raw %</td>
<td>Raw %</td>
<td>Raw %</td>
<td>Raw %</td>
</tr>
<tr>
<td>207</td>
<td>16.2</td>
<td>193</td>
<td>15.1</td>
<td>188</td>
</tr>
<tr>
<td>Simple Past</td>
<td>554</td>
<td>43.4</td>
<td>109</td>
<td>8.5</td>
</tr>
<tr>
<td>Total</td>
<td>761</td>
<td>59.6</td>
<td>302</td>
<td>23.6</td>
</tr>
</tbody>
</table>

The frequency of the quotatives in the Conversation corpus is also represented in Figure 4.1, which highlights the distribution across simple present and simple past tense, and across the sum of simple present and simple past:

The results displayed in Table 4.1 show that in Conversation the traditional quotative say accounts for almost 60% of all quotative use considered in this study. The new quotatives be like and go, however, also account for a significant portion of the quotative use in the corpus: be like accounts for almost one quarter (23.6%) of all quotative use; go accounts for about 15% (14.9) of all quotative use.
An important finding of the analysis of the frequency of use of the quotatives in Conversation is that the frequency of use of the quotatives *say, be like* and *go* in the Present Tense is very similar (see Figure 4.1), in that they each account for approximately 15% of all quotative use in the corpus. This means that while, when looking at the sum of simple present and simple past use of the quotatives, the traditional quotative *say* is by far the most frequent quotative in the corpus, when looking at the use of the quotatives in the simple present, there is almost no difference in frequency between the traditional quotative *say* and the new quotatives *be like* and *go*. Thus, when speakers of American English report direct speech in casual conversation introducing the quoted speech with a lexicalised quotative (i.e., not using "zero quotatives"), they choose among present tense forms of *say, be like* and *go* with about the same frequency.

Another interesting finding is that while the traditional quotative *say* is most often used in the simple past, the new quotatives seem to be used more often in the simple present: the use of simple present *be like* accounts for 15% of the total quotative use, against the 8.5% of simple past *be like*; *go* is almost never used in the simple past.

Service Encounters

The frequency and distribution of the quotatives in the Service Encounters corpus are displayed in the Table 4.2, which follows the same conventions of 4.1.

<table>
<thead>
<tr>
<th>QUOTATIVE</th>
<th>SAY</th>
<th>BE LIKE</th>
<th>GO</th>
<th>BE ALL</th>
<th>TOTAL QUOTATIVE USE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw %</td>
<td>Raw %</td>
<td>Raw %</td>
<td>Raw %</td>
<td>Raw %</td>
<td>Raw %</td>
</tr>
<tr>
<td>Simple Present</td>
<td>6 2.9</td>
<td>81 39.5</td>
<td>19 9.2</td>
<td>9 4.3</td>
<td>115 57.9</td>
</tr>
<tr>
<td>Simple Past</td>
<td>49 23.9</td>
<td>40 19.5</td>
<td>1 0.5</td>
<td>0 0</td>
<td>90 42</td>
</tr>
<tr>
<td>Total</td>
<td>55 26.8</td>
<td>121 59</td>
<td>20 9.7</td>
<td>9 4.3</td>
<td>205 100</td>
</tr>
</tbody>
</table>
Figure 4.2 also represents the frequency of the quotatives, highlighting their distribution across tense:

![Figure 4.2: Frequency of use of quotatives in Service Encounters corpus](image)

The results of the analysis of the frequency and distribution of the use of the quotatives in the Service Encounters corpus are especially interesting in that they show that, in this corpus, the new quotative *be like* is by far the most frequently used quotative, accounting for almost 60% of all quotative use. *Be like* is almost twice as frequent as the traditional quotative *say*, which accounts for only about 27% of all quotative use in the corpus.

As in Conversation, in Service Encounters, while the traditional quotative *say* is by far used most often in the simple past, the new quotatives *be like*, *go* and *be all* are used much more frequently in the present tense.

Finally, it should be mentioned that the relatively low frequency of forms of *say* in the simple present has to be understood in connection with the high number of conditional,
past progressive with gonna and going to and future forms, which, as explained in Chapter Three, were not included in the counts.

The finding that in Service Encounters be like, with its 60% share of total quotative use, is by far the most frequent of the set of quotatives analysed in this study and is much more frequent than the traditional say is most surprising, and has most probably to be interpreted in the light of the type of spoken discourse included in Service Encounters, which, as mentioned in Chapter Three, seems to mainly consist in casual conversation among co-workers (who often are college students). Also, as mentioned in Chapter Three, in a comprehensive study of the spoken and written registers of TOEFL 2000 SWAL using MD analysis, Service Encounters was found to be highly involved, highly situation-dependant and highly marked for nonimpersonal style. These findings are all consistent with the fact that Service Encounters contains much casual conversation between college students and has a high frequency of the new quotative be like. The discussion on the surprisingly high frequency of be like in Service Encounters will be taken up below in this chapter, in the section on the comparison of frequency of use of the quotatives across registers.

Office Hours

The frequency and distribution of the quotatives in the Office Hours corpus are displayed in Table 4.3, which follows the same conventions of Table 4.1. and Table 4.2.

<table>
<thead>
<tr>
<th>SAY</th>
<th>BE LIKE</th>
<th>GO</th>
<th>BE ALL</th>
<th>Total quotative use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>%</td>
<td>Raw</td>
<td>%</td>
<td>Raw</td>
</tr>
<tr>
<td>Simple Present</td>
<td>22</td>
<td>40</td>
<td>6</td>
<td>10.9</td>
</tr>
<tr>
<td>Simple Past</td>
<td>23</td>
<td>41.8</td>
<td>4</td>
<td>7.2</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>81.8</td>
<td>10</td>
<td>18.1</td>
</tr>
</tbody>
</table>
The first important finding of the analysis of frequency of use in Office Hours is that only the quotatives *say* and *be like* occur in this corpus: there are no occurrences of *go* and *be all*. The traditional quotative *say* represents almost 82% of all quotative use in the corpus, while the remaining 18% is accounted for by *be like*.

Figure 4.3 below highlights that in the Office Hours corpus, *say* is consistently the most frequent quotative across simple present and simple past, and, consequently, in the total representing the sum of simple present and simple past.

The low frequency of use of the quotatives in Office Hours (only 55 quotatives), combined with the low frequency of new quotatives and the predominance of the traditional quotative *say* is an aspect that deserves more attention; it would be interesting to investigate what contextual factors determine the low use of quotatives, and thus the low use of direct reported speech, in this particular register of spoken interaction.
2. Comparison of frequency of use of the quotatives across registers.

In this section I compare the frequency of use of the quotatives across the three registers of spoken interaction represented by the three corpora used in the study, Conversation, Service Encounters and Office Hours. Because the three corpora are of unequal size (425,000; 97,600; 50,400 words respectively), the figures representing the raw frequency within each register (displayed in Table 4.1, Table 4.2, Table 4.3) were normed to occurrences per 100,000 words. The results of the normalization of the frequency of occurrence of the quotatives across the three corpora are summarized in Table 4.4 below.

Table 4.4: Normed frequency of the quotatives across the three corpora.

<table>
<thead>
<tr>
<th></th>
<th>Conversation (Frequency per 100,000 words)</th>
<th>Service Encounters (Frequency per 100,000 words)</th>
<th>Office Hours (Frequency per 100,000 words)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw</td>
<td>%</td>
<td>Raw</td>
</tr>
<tr>
<td><strong>SAY</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple Present</td>
<td>48.7</td>
<td>16.2</td>
<td>6.1</td>
</tr>
<tr>
<td>Simple Past</td>
<td>130.3</td>
<td>43.4</td>
<td>50.2</td>
</tr>
<tr>
<td>Total</td>
<td>1879</td>
<td>59.6</td>
<td>56.3</td>
</tr>
<tr>
<td><strong>BE LIKE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple Present</td>
<td>45.4</td>
<td>15.1</td>
<td>82.9</td>
</tr>
<tr>
<td>Simple Past</td>
<td>26.6</td>
<td>8.5</td>
<td>40.9</td>
</tr>
<tr>
<td>Total</td>
<td>71</td>
<td>23.6</td>
<td>123.9</td>
</tr>
<tr>
<td><strong>GO</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple Present</td>
<td>44.2</td>
<td>14.7</td>
<td>19.4</td>
</tr>
<tr>
<td>Simple Past</td>
<td>0.5</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Total</td>
<td>44.7</td>
<td>14.8</td>
<td>20.4</td>
</tr>
<tr>
<td><strong>BE ALL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple Present</td>
<td>3</td>
<td>1</td>
<td>9.2</td>
</tr>
<tr>
<td>Simple Past</td>
<td>2.4</td>
<td>0.8</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>5.4</td>
<td>1.8</td>
<td>9.2</td>
</tr>
<tr>
<td><strong>Total quotative use</strong></td>
<td>300.1</td>
<td>100</td>
<td>210</td>
</tr>
</tbody>
</table>
The comparison of the frequency of the quotatives in the three corpora revealed a few interesting aspects. First of all, it appears that quotatives are used the most frequently in Conversation (300 occurrences), and the least frequently in Office Hours (109 occurrences), with Service Encounters in a somewhat median position (210 occurrences). This means that the use of direct quotation is more frequent in certain registers of spoken interaction than others, an aspect that, as mentioned above, would deserve further investigation. This suggests, indeed, that the type of spoken interaction affects the use of direct speech quotation and the type of quotative formats used to introduce direct quotation. In this perspective, it is perhaps not surprising that Conversation shows the highest frequency of quotatives among the three corpora: casual conversation has traditionally been considered a type of spoken interaction naturally generating storytelling, which in turn generally includes the use of direct quotation to create interpersonal involvement and for evaluative and evidential purposes (See Tannen 1986, 1988, 1989; Mayes 1990; Holt 1996; Macaulay 1987). Based on this assumption, it is intuitively not so surprising that Office Hours, a corpus consisting in one-on-one consultations between students and instructors, and thus presumably not including much storytelling, should yield a low use of direct quotation and, thus, a low frequency of quotatives. An interesting aspect, in this perspective, is the frequency of the quotatives in Service Encounters, which is in a somewhat medial position between Office Hours and Conversation. This may tentatively be explained with the fact that, although the Service Encounters corpus was designed to contain highly structured interaction, it actually also contains much casual conversation between co-workers, a fact that, as mentioned in Chapter Three may challenge the possibility of considering Service Encounters as a spoken register clearly distinct from conversation.
The distribution of the different quotatives across the three corpora is another interesting aspect of the quantitative analysis. I present the analysis of the distribution of the different quotatives in separate sections.

*Say*

The "standard" quotative *say* is by far the most frequent quotative used in Conversation and Office Hours, where it accounts for approximately 60% and 82% of the total quotative use respectively. *Say* is less frequent than *be like* in Service Encounters, where it accounts for approximately 27% of total quotative use, against the 59% of *be like*.

*Be like*

*Be like* is the second most frequent quotative both in Conversation and Office Hours, where it accounts for approximately 24% and 18% of the total quotative use respectively. *Be like* is by far the most frequent quotative in Service Encounters, where it accounts for approximately 60% of the total quotative use in the corpus, against the 27% of total quotative use represented by *say*. In all three corpora, *be like* is much more frequent (almost twice as frequent) in the simple present than in the simple past.

The finding that *be like* is by far the most frequent quotative in Service Encounters is notable: *be like* seems to occupy the "quotative space" that in corpora of general, casual conversation, such as Conversation, and in corpora representing more formal registers of interaction, such as Office Hours, is occupied by the traditional *say*. Why this should be so may very well have to do with the nature of the spoken discourse included in Service Encounters, which, as mentioned in Chapter Three and above in this chapter, seems to mainly consist in casual conversation among co-workers and/or colleges students. If much of
the interaction included in Service Encounters is simply very casual conversation among college students, indeed, then it is perhaps not so surprising that the quotative *be like*, which has traditionally been reported to be more frequent in the language of young speakers (Blyth et al. 1990, Romaine & Lange 1991, Ferrara & Bell 1995, Tagliamonte & Hudson 1999, Macaulay 2001) should be so frequent in Service Encounters.

Because previous studies of the new quotative *be like* (Blyth et al. 1990, Romaine & Lange 1991, Ferrara & Bell 1995) have not adopted a corpus-based approach, it is unfortunately not possible to compare these findings with previous findings. Blyth et al. (1990), indeed, do not provide the number and/or proportion of use of *be like*. Ferrara & Bell (1995) provide the proportions of use of *be like* in their 1990 corpus, distinguishing between females and males use. Because this present study did not adopt a sociolinguistic approach and did not look at issues of age and gender, this data cannot be compared with Ferrara & Bell’s data.

However, overall, these data clearly show that the quotative *be like* is a very common form of direct speech introduction in different registers of spoken interaction in present-day American English, and demonstrate that in certain registers it can even be twice as frequent as the traditional *say*.

**Go**

The quotative *go* is relatively frequent in Conversation and Service Encounters, where it accounts for 14.4% and 9.7% of total quotative use respectively, and absent in Office Hours. *Go* follows the same trend of the new quotative *be like* in being used more
often in the simple present. As a matter of fact, go is almost never used in the simple past tense form went.

**Be all**

The new quotative *be all* is the least frequent in Conversation and Service Encounters, and is absent in Office Hours. Interestingly, it is more frequent in Service Encounters, where it accounts for 4.3% of the total quotative use, than in Conversation, where it accounts for only 1.7% of the total quotative use. Like *be like* and *go*, the new quotative *be all* seems to be more often used in the simple present.

3. Analysis of frequency of use of the quotatives with grammatical person.

The analysis of the association of the use of the new quotatives with grammatical person (i.e.: first person singular *I*, second person singular *you*, third person singular *he, she, someone* and any other person’s name, etc.) is important because, in previous studies of the quotative *be like*, it has been considered as one of the “linguistic diagnostics for expansion” (Tagliamonte & Hudson 1999: 152), i.e., an internal linguistic factor that can be correlated with the progression of the linguistic change.

Blyth et al. (1990: 221) and Romaine & Lange (1991: 243) found that the quotatives *say* and *go* tended to be used in association with third person grammatical subjects to express the speech of others, while *be like* was rarely used in the third person singular, presumably due to its association with the reporting of speaker’s own internal or emotional states (Romaine & Lange 1991). In Ferrara & Bell’s most recent data sample from 1994, however, nearly half of all occurrences of *be like* were in the third person. More specifically, in the 1994 data sample, the use of *be like* with first person singular subjects accounted for 37% of
the total use of the quotative, while the use of *be like* with third person singular subjects accounted for 40% of the total use of the quotative. Ferrara & Bell interpreted this as evidence of the expansion of the discourse function of *be like* and of its increasing use within the quotative system of American English (Ferrara & Bell 1995).

Taken together, these studies provide a useful starting point for the analysis of the association of the quotatives with grammatical person. Here I look at the frequency of use of the quotatives in association with grammatical person in the three corpora under investigation, and compare the findings to previous studies, where possible. Because the focus of previous studies looking at association with grammatical person as a linguistic "diagnostics" for expansion (Blyth et al. 1990, Romaine & Lange 1991, Ferrara & Bell 1995) have focused on a main distinction between use of the quotative with the first person singular, and use of the quotative with the third person singular, in this analysis of the association of the use of the quotatives with grammatical person I will follow the same method. More specifically, I present frequency data of the use of the quotatives with the first person singular, the third person singular, and "other" subjects, where the latter include second person singular and plural (*you*), first person plural (*we, John and I*, etc.), and third person plural (*they*). The data are displayed in Tables 4.5, 4.6, and 4.7 below. Frequency data specifying association of simple present and simple past tense forms of the quotatives with all grammatical persons is provided in Appendix 1. I now briefly analyse the use of the quotatives in association with grammatical person in the three corpora used in the present study, with a particular focus on Conversation, which, being the largest corpus, has yielded the largest number of quotatives.
Belike

As shown in Table 4.5, in Conversation the use of *be like* in association with first person singular subjects accounts for almost 54% of the occurrences of the quotatives, while the use in the third person singular accounts for 41% of the total use of the quotative in the corpus. This finding strongly contradicts Blyth et al.’s (1990) claim that “*be like* is rarely used in the third person singular” (221), and instead confirms Ferrara & Bell’s (1995) finding that the use of *be like* in the third person singular is increasing.

The findings from the other two corpora used in the study, Service Encounters and Office Hours, corroborate this tendency, as in both of them *be like* is used with third person

---

**Table 4.5: Quotative use with grammatical person in Conversation**

<table>
<thead>
<tr>
<th>SAY</th>
<th>BE LIKE</th>
<th>GO</th>
<th>BE ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>% raw</td>
<td>%</td>
<td>raw %</td>
</tr>
<tr>
<td>1st ps</td>
<td>346</td>
<td>45.4</td>
<td>163</td>
</tr>
<tr>
<td>3rd ps</td>
<td>320</td>
<td>42.0</td>
<td>124</td>
</tr>
<tr>
<td>Other</td>
<td>95</td>
<td>12.4</td>
<td>15</td>
</tr>
<tr>
<td>total</td>
<td>761</td>
<td>100</td>
<td>302</td>
</tr>
</tbody>
</table>

**Table 4.6: Quotative use with grammatical person in Service Encounters**

<table>
<thead>
<tr>
<th>SAY</th>
<th>BE LIKE</th>
<th>GO</th>
<th>BE ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>raw</td>
<td>% raw</td>
<td>%</td>
<td>raw %</td>
</tr>
<tr>
<td>1st ps</td>
<td>23</td>
<td>41.8</td>
<td>60</td>
</tr>
<tr>
<td>3rd ps</td>
<td>15</td>
<td>27.2</td>
<td>54</td>
</tr>
<tr>
<td>Other</td>
<td>17</td>
<td>30.9</td>
<td>7</td>
</tr>
<tr>
<td>total</td>
<td>55</td>
<td>100</td>
<td>121</td>
</tr>
</tbody>
</table>

**Table 4.7: Quotative use with grammatical person in Office Hours**

<table>
<thead>
<tr>
<th>SAY</th>
<th>BE LIKE</th>
<th>GO</th>
<th>BE ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>raw</td>
<td>% raw</td>
<td>%</td>
<td>raw %</td>
</tr>
<tr>
<td>1st ps</td>
<td>10</td>
<td>22.2</td>
<td>3</td>
</tr>
<tr>
<td>3rd ps</td>
<td>7</td>
<td>15.5</td>
<td>6</td>
</tr>
<tr>
<td>Other</td>
<td>28</td>
<td>62.2</td>
<td>1</td>
</tr>
<tr>
<td>total</td>
<td>45</td>
<td>100</td>
<td>10</td>
</tr>
</tbody>
</table>
singular grammatical subjects in more than 40% of the cases: in Service Encounters the use of *be like* with third person singular subjects accounts for 44.6% of the total occurrences, while in Office Hours the use of *be like* accounts for 60% of the total use of the quotative.

In all three corpora, the use of *be like* with other grammatical persons accounts for a very low percentage of the total use of the quotative. For example, in Conversation *be like* is used with other subjects in a negligible 4.9% of the cases.

**Go**

In both the corpora in which it is accounted for, Conversation and Service Encounters, *go* is used much more frequently (almost twice as frequently) with third person singular subjects than with first person singular subjects: the use of *go* with third person subjects accounts for almost 58% of the total use of the quotative, as opposed to the almost 32% of the total use in Conversation, and for 65%, as opposed to the 35% of use with first person singular subjects, in Service Encounters.

This finding seems basically to reconfirm Blyth et al.'s claim that *go* is predominantly used with the third person singular. Unfortunately, there are no other quantitative studies of the use of quotative *go* in American English, and it is thus not possible to trace the evolution of the use of *go* with third person singular subjects. However, these findings clearly show that the new quotative *go* is more frequently used to quote the speech of others, rather than one's own speech.

**Be all**

The quotative *be all* is the least frequent of all, and, as mentioned above, it does not occur in Office Hours. From the little data available in Conversation and Service Encounters
it does also not seem possible to establish any pattern of use in association with different grammatical subjects: in Conversation be all is used more frequently in association with third person singular subjects, where it accounts for 56.5% of total quotative use, while in Service Encounters it occurs more frequently with first person singular subjects (66.6% of total use). Perhaps one interesting finding is that there were no occurrences of be all with plural subjects in the corpora under investigation.

Summary of quantitative analysis

The results of the analysis of the frequency of simple present and simple past tense forms of the new quotatives be like, go, be all and of the traditional quotative say in the three corpora Conversation, Service Encounters, and Office Hours have shown that the traditional quotative say accounts for a significant portion of the total direct speech quotative use across the three corpora. However, the proportion of frequency of occurrence of this traditional quotative varies considerably across the three corpora.

While say is the most frequent quotative in Conversation and Office Hours, the most frequent direct speech quotative in Service Encounters is be like. Be like stands out as a solid, stable, considerably frequent quotative across all three corpora. It is the only new quotative occurring in Office Hours, and it accounts for more than half and almost one quarter of total quotative use in Service Encounters and Conversation respectively.

The new quotatives go and be all occur only in Conversation and Service Encounters. However, while go appears as relatively frequent in that it accounts for about 15% and 10% of total quotative use respectively in Conversation and Service Encounters, be all occurs with a very low, almost negligible frequency.
Qualitative Analysis

In this section I present the results of the analysis of the discourse-pragmatic functions of the quotatives in the Conversation corpus. The analysis was based on the framework described in Chapter Three, and aims at capturing a main, rather broad distinction between plausible and improbable quotation, that is between quotation of speech which may plausibly have been uttered, and quotation of speech that was unlikely to have been uttered as it seems to simply represent the gist of some speech, the internal speech or thought of the quoted speaker. As mentioned in Chapter Three, underlying this framework is the assumption that, although direct quotation cannot be considered a truly verbatim report and is likely to be a reconstruction by the speaker, it has to be recognized that there are different levels of plausibility of the quotation. The level of plausibility of the quotation bears relevance to a study of the discourse function of the quotatives because the content of the quote is thought to determine which quotative will be used.

Because according to the framework used for this analysis all quotations attributed to plural grammatical subjects is improbable, I will only discuss quotations attributed to singular grammatical subjects, with a particular focus on first and third person singular subjects. There are two main reasons for circumscribing the discussion of the discourse function of the quotatives to quotatives in the first and third person singular: first, as shown in Tables 4.5, 4.6, 4.7, the use of the quotatives in the second person singular and in plural persons accounts for a negligible proportion of the total quotative use; secondly, in the literature on be like – the only new quotative that has been the object of quantitative studies – the use of the quotative with first and third person singular has been used to make claims
about the discourse function of the quotative (See Blyth et al 1990: 221; Ferrara & Bell 1995: 279 ff.).

The results of the discourse-pragmatic analysis are presented in Table 4.8 below:

Table 4.8.: Discourse-pragmatic functions of the quotatives.

<table>
<thead>
<tr>
<th></th>
<th>SAY</th>
<th>BE LIKE</th>
<th>GO</th>
<th>BE ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>raw</td>
<td>%</td>
<td>raw</td>
<td>%</td>
<td>raw</td>
</tr>
<tr>
<td>1ps</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plausible</td>
<td>255</td>
<td>73.6</td>
<td>45</td>
<td>27.6</td>
</tr>
<tr>
<td>Improbable</td>
<td>91</td>
<td>26.3</td>
<td>118</td>
<td>72.4</td>
</tr>
<tr>
<td>2ps</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plausible</td>
<td>17</td>
<td>36.1</td>
<td>2</td>
<td>66.6</td>
</tr>
<tr>
<td>Improbable</td>
<td>30</td>
<td>63.8</td>
<td>1</td>
<td>33.3</td>
</tr>
<tr>
<td>3ps</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plausible</td>
<td>222</td>
<td>69.3</td>
<td>84</td>
<td>67.7</td>
</tr>
<tr>
<td>Improbable</td>
<td>98</td>
<td>30.6</td>
<td>40</td>
<td>32.2</td>
</tr>
</tbody>
</table>

Table 4.8 shows that when used in the first person singular, the quotatives *say* and *go* generally introduce *plausible* quotation, which accounts for 73.6% and 66.6% of first person *say* and *go* respectively; while *be like* used with first person singular subjects generally introduces *improbable* quotation. Note also that, when used in association with the third person singular, *say*, *go* and *be like* all generally introduce *plausible* quotation, which accounts for 69.3%, 67.7%, and 89% of the total use in the third person singular. *Be all* seems to introduce plausible and improbable quotation with equal frequency both in the first and the third person, but any pattern must be taken as tentative due to the low frequency of occurrence of this quotative. Finally, quotative use in the second person singular is negligible for all quotatives and is not accounted for the quotative *be all*.

I now proceed to the discussion of the discourse-pragmatic function of the single quotatives across the first and third singular grammatical persons.
Say

The results displayed in Table 4.8 and plotted in Figure 4.4 below show that, when the traditional quotative *say* is used with first person singular subjects, it generally introduces a quotation that may be taken as plausible. Indeed, the use of *say* with first person singular subjects has been found to introduce *plausible* quotation in almost 74% of the cases, and *improbable* quotation in only 26% of the cases. This finding seems to substantiate Blyth et al.'s (1990) claim that "*say* is the most neutral verb, unambiguously marked for speech (direct or indirect)" (222). However, it also shows that the traditional quotative *say* may be used to introduce improbable quotations, as the example below shows:

(42) *<1945> [Was he] was he, quivering over the price or anything? <1914> Nope. <1945> Was he serious or just play [<unclear>] <1914> [I think] serious <1945> What's he want to do with the dog? <1914> Hunt. <1945> He's serious. <1914> And he said ... I'm gonna neuter the dog, I don't want anything to do with the training part I'm leaving that up to the experts, [[I said]], damn I wish there were more people like you. <1945> Yeah ... [<unclear> my mouth and then my gum gets stuck in my hair and <1914> But he was out there at putting my line in, he looked over and Marvin was out there cause Marvin, [<unclear>] <1945> [He is a <unclear>] <1914> Marvin's [<unclear> and that's like shutting off ... a part of Marvin <nv_laugh> <1945> <laughing> oh my g= <unclear> (Conversation)

The quotative *say* here introduces improbable quotation in that the quote "damn, I wish there were more people like you" seems to represent what the speaker thought, rather than what he/she told her/his interlocutor.

The results of the discourse-pragmatic analysis showed that the traditional quotative *say* introduces plausible speech in the vast majority of the cases in the first person, and also when used with third person singular subjects (69%), thus substantiating previous claims on the neutrality of this traditional quotative (Blyth et al. 1990: 220).
Belike

The pattern of use shown by be like is very different from the one displayed by say: when used with first person singular subjects, be like introduces plausible speech in only approximately 28% of the cases. In the rest of the cases (72%) it introduces improbable quotation, that is the inner speech or thought of the speaker, or a representation of the attitude of the speaker. This finding substantiates previous research on be like arguing that be like is used to introduce either direct speech and inner thought reports (Schourup 1985; Tannen 1986, 1989; Blyth et al. 1990; Romaine & Lange 1991; Ferrara & Bell 1995). The importance of this finding lies in the fact that it provides an empirical, quantitative basis for previous, non-corpus based claims about the discourse-pragmatic function of the new quotative be like. Examples of the use of be like with first person singular subjects to introduce improbable quotation are provided in Chapter Three (ex. 29, 30, 31, 32) and below:

(43) 2080> I thought that they really had breasts underneath them. <2072> I know it is too funny. <2080> [[I was like]] my god. <2078> Maybe we should call the holiday bake line and explain that to them. (Conversation)

(44) You know, I'm so mad. He had it halfway done. I was reading a magazine, I looked up and [[I was like]] oh, it's great. It's just what I wanted. I finished reading and I thought he was evening it up and he cut like another two inches off. (Conversation)

The use of be like with third person singular subjects follows a pattern that is the reverse of the pattern found with the use of the quotative with first person singular subjects: when used with the third person singular, be like typically introduces plausible quotation. The proportions are quite revealing: when be like is used with third person singular subjects, it introduces improbable quotation in approximately 32% of the cases, and plausible quotation in approximately 68% of the cases, against the 27.6% of first person plausible
quotation and the 72.4% of improbable quotation introduced by *be like* with first person singular subjects.

Finally, one important remark about the use of *be like* with first person singular subjects: the fact that in 28% of the cases of *be like* with first person singular subjects, the quotative was found to introduce *plausible* quotation challenges Ferrara & Bell’s (1995) claims on the impossibility of distinguishing speech from thought when *be like* is used with first person singular subjects. Ferrara & Bell argue that there is a difference between the use of *be like* in the third person singular and the use of *be like* in the first person singular: “When dialogue is reported in first person it is impossible to distinguish thought from speech”, while “when a third person’s dialogue is introduced (*he, she, they*), there is the possibility that actual words were spoken and that the quotation so introduced is a representation or demonstration of that speech.” (279). Clearly, the fact that, in 28% of the cases, first person *be like* was found to introduce plausible quotation challenges this claim. It was found that it is not always impossible to distinguish speech from thought when *be like* is used with first person singular: there are cases in which the quotation introduced by first person singular *be like* has propositional content that contributes to the advancement of the story (i.e. complicating action) – an element that Ferrara & Bell (1995) too recognize as an indicator of plausibility of the quote (Ferrara & Bell 1995: 279-280) – or is part of an utterance to which the protagonists responded. This is illustrated in examples (27) and (28) in Chapter Three, and in (45), (46), (47) below:

(45) And she had like eight things so Lois was waiting on her and of course you know, that lady wants to talk so the guy that the tube wore out on I said I'm sorry can I please have you to go the push out window so you can push through your money you could skip everybody, he had to wait behind her and Lois couldn't get her machine to work and she, I think she just got nervous and she kept screwing up so Andy had to finish it and then she was gone and [[I'm like]] I'm so sorry you had to wait he
said well, he goes, I guess things happen and stuff like that. <2091> That's very nice of him. <2090>
And then there was a girl behind him and I'm just like oh crap, what can I do. She was mad
our service sucked and she got ticked off because she had to wait in line again but there's not much
we can do lady. (Conversation)

(46) Like I was talking to this one, I was talking to this one girl right and she was telling me shut up,
shut up, shut up that means I like what you are saying keep talking to me. <1910> Well how did you
handle that? <1911> [[I'm like]] I thought you were telling me to shut she goes shut up, shut up,
shut up that means I like what you're saying keep talking to me. <?> <nv_laugh> <1910> Now that's
kind of crazy. <?> <unclear> <?> It's really wild talk and um <?> <nv_laugh> <?> That's good I
like that. (Conversation)

(47) She said will you answer the phone, will you answer the phone and [[I was like]] hello. Then it
was my Aunt and she's like it's me, answer the other line and so I kept pushing and there's nobody on
the other line so then she won't pick up the phone because she thinks it's like ... (Conversation)

Notice the difference, in (45), between the first I'm like (bolded), which is taken to
introduce plausible quotation on the grounds that it is part of an utterance to which the
interlocutor responds (he said well, he goes, I guess things happen and stuff like that...), and
the second I'm like (underlined), which is taken to introduce improbable quotation, as the
quote seems to represent the thought of the speaker, rather than plausibly uttered speech.

Go

The new quotative go seems to follow a pattern similar to the one of say, in that it
tends to be used to introduce plausible speech both when used with first person singular
subjects (66.6%) and when used with third person singular subjects (89%). Go is used to
introduce improbable quotation only in about 34% of the occurrences of use in the first
person singular, and 11% of the occurrences of use in the third person singular.

The finding that when used with first person singular subjects, go introduces
plausible speech in almost 67% of the cases may be considered to challenge previous claims
that go is used for dramatic effect and evaluation (Blyth et al. 1990: 220). The results of this
analysis of the discourse function of *go* suggests indeed that *go* is used in a way much similar
to the neutral *say*. The examples below illustrate the use of *go* to introduce *plausible*
quotations:

(48) *Last night I was supposed to call Stacey 'cause she was gonna give me a perm so I dialed this
number. This guy answers and I didn't recognize the voice but I thought well her roommate maybe
she knows somebody. [[I go]] is Stacey there and he goes no she dropped off her clothes and left
she'll be back about eight 0 clock. <laughing> And I'm like </laughing> she told me to come over at
seven 0 clock. So then I hung up and ... (Conversation)*

(49) *Nancy came to the door. <?> But it wasn't her. It was the Jehovah witnesses with the magazine
and [[I go]] no, no, no. I said I thought you knew, I tried to explain to her that I was waiting for
somebody that's why I called you Nancy or whatever and she's like, oh, anyway, and like someone
beeped and it (Conversation)*

(50) *... like I don't know I'm like another thirty feet away and I turn back and he's still there and he
pulls up and rolls up the window and he stands there saying are you going to get in or not? [[I go]]
no I'm walking. He goes get in here so I sat in and he goes you know for a college student you don't
have a lot of brains at all, yeah I know I was always wondering people always ask me while I'm
dating you <nv laugh>
<2091> Jeez. (Conversation)*

**Be all**

The new quotative *be all* seems to introduce plausible and improbable quotation with
equal frequency both when occurring in the first person singular, and when occurring in the
third person singular. However, the total number of occurrences of *be all* is so low that it
does not seem possible to make any strong claims about the discourse-pragmatic function of
this quotative.
Summary of discourse-pragmatic analysis of the quotatives

*First person singular subjects*

The proportion of use of the new quotatives *be like, go, be all* and of the traditional *say* with first person singular subjects to introduce plausible and improbable quotation is plotted in Figure 4.4 below:

Figure 4.4: Discourse-function of the quotatives with first person singular subjects

![Bar chart showing discourse function of quotatives with first person singular subjects](image)

Figure 4.4 shows that, when the traditional quotative *say* is used with the first person singular subject *I*, it generally introduces plausible speech (74%). Only in approximately one quarter of the cases (26%), does *say* with first person singular subjects introduce improbable quotation.

The new quotative *be like* follows a completely different pattern: in the vast majority of the cases (72%), first person singular *be like* introduces improbable speech, that is a representation of the thought or inner speech of the speaker; only in about one quarter of the cases (28%) does first person singular *be like* introduce plausible quotation.
The distribution of the discourse-functions of the new quotative go is closer to that of the traditional quotative say than to that of the new be like: first person singular go introduces plausible speech in almost 67% of the cases and improbable speech in about 33% of the cases.

When used with the first person singular subject, the new quotative be all seems to introduce plausible and improbable speech with equal frequency. However, conclusions are tentative in this case, due to the low number of occurrences found in the corpus.

**Third person singular subjects**

The proportion of use of the new quotatives be like, go, be all and of the traditional say with third person singular subjects to introduce plausible and improbable quotation is plotted in Figure 4.5 below:

Figure 4.5: Discourse function of the quotatives with third person singular subjects
As shown in Figure 4.5, the neutral quotative *say* introduces plausible speech in the majority of the cases (69%) in which it is used in association with third person singular subjects. Interestingly, this pattern is followed also by the new quotative *be like* and *go*, which, when used with third person singular subjects, introduce plausible quotation in almost 68% and 89% of the cases respectively. The new quotative *be all* introduces plausible quotation in a lower proportion (54%) but conclusions on the discourse-function of this quotative are still very tentative due to the low number of occurrences retrieved in the corpus.

**Summary**

In this chapter I have presented the results of the analysis of the frequency of occurrence of simple present and simple past tense forms the three new quotatives *be like*, *be all*, *go* and of the traditional quotative *say* in three corpora of spoken American English, e.g., Conversation, Service Encounters and Office Hours; I have compared their frequency across the different corpora; and I have analysed their patterns of occurrence in association with grammatical person. I have then discussed the results of the analysis of the discourse-pragmatic function of the quotatives, focusing on the discourse-pragmatic function of the quotatives used in association with first and third person singular subjects.

I now proceed to Chapter Five, where I will discuss the important findings of this study, the limitations of the study and implications for further research and teaching.
CHAPTER V: CONCLUSIONS

Major findings

The first research question that drove this study was concerned with the frequency of the new quotatives be like, go, be all in present-day spoken American English, and was divided into three closely related questions regarding the overall frequency of these quotatives in comparison with the traditional quotative say in three relatively small corpora of spoken American English, the comparison of their frequency across the corpora, and their frequency of occurrence in association with different grammatical subjects.

With respect to this first research question, this study has shown that while the traditional quotative say is still the most frequent direct speech introducer in spoken American English, the new quotative be like stands out as a stable, significantly frequent direct speech quotative, as attested by the fact that it occurs in all corpora used in this study, and the fact that it is by far the most frequent quotative in one of the corpora (Service Encounters). The new quotative go also appears as a relatively stable element in the American English quotative system, while the new quotative be all seems to be still rather infrequent, although it is attested in two of the corpora used in this study (Conversation and Service Encounters). Further, this study has shown that while all new quotatives appear to be used more frequently in the present tense but are used relatively frequently in the past tense too, the past tense quotative form went is hardly ever used. With respect to the question of frequency of occurrence in association with grammatical person, this study has shown that there are clear patterns of association: all new quotatives are used most often in association with first and third person singular subjects. More importantly, this study has shown that the
use of *be like* in association with third person singulars grammatical subjects accounts for more than 40% of the use of this quotative in all registers, thus suggesting that the use and the discourse-function of this quotative are in the process of expanding.

The second research question was concerned with the discourse-pragmatic function of the quotatives. The analysis carried out in this study aimed at establishing the degree to which the new quotatives are used to introduce plausible or improbable speech, that is whether they introduce quotation that represents the inner thought or speech of the reported speaker, or whether instead they introduce some speech that, while it may not be taken as a verbatim report, seems to have been plausibly uttered as it has prepositional content that contributes to the advancement of the story, or it is part of an utterance to which the reported interlocutor responded. The analysis of the discourse-pragmatic function has shown that there are some clear patterns of association between grammatical person of the quotative and discourse function of the quotative. The discourse-function of *go* is quite similar to that of *say*, in that *go* generally introduces plausible quotation both in the first person singular and the third person singular. On the contrary, the new quotative *be like* displays much more complex discourse-pragmatic functions: when used in the first person singular, it generally introduces improbable quotation, that is speech that represents the emotional state or the thoughts of the speaker at the time of the reported situation, rather than plausibly uttered speech; when used in the third person singular, *be like* generally introduces plausible speech. This finding is important in that it confirms previous non corpus-based studies arguing that the discourse function of this quotative is expanding from representation of the inner thoughts and of the emotional states of the speaker (Butters 1982, Tannen 1986, 1989) to quotation of plausible direct speech (Ferrara & Bell 1995). At the same time, this finding
also indicates that the quotative *be like* is used to introduce plausible direct speech also in the third person singular, just as the traditional quotative *say* and the new quotative *go*, thus contradicting previous non corpus-based studies claiming that it is impossible to distinguish thought from speech when *be like* is used in the first person singular (Ferrara & Bell 1995).

This study has also shown that type and frequency of the quotatives vary across corpora representing different registers of spoken interaction. This is clearly demonstrated by the overall low frequency of quotatives and by the negligible frequency of new quotatives in the Office Hours corpus. This suggests that, in consultative registers of spoken interaction such as one-on-one consultations between student and instructor, direct speech quotation is used less frequently than in casual registers of spoken interaction, such as conversation. We have also reasons to believe that in consultative registers such as Office Hours, not only is direct speech less frequent, but also it is used with different discourse functions from the ones identified by previous research (Tannen 1986, 1988, 1989; Macaulay 1987; Mayes 1990, Holt 1996) for the use of direct quotation in conversational storytelling, e.g. creating interpersonal involvement, evidential and evaluative functions.

**Limitations of the study**

This study has looked at simple present and simple past tense forms of the new quotatives *be like, go, be all,* and the anchor *say* in three corpora of spoken American English: Conversation, Service Encounters and Office Hours. While these forms have provided a fairly complete picture of the frequency of use and of the discourse function of the new quotatives, they certainly do not exhaust all possible forms that may occur. For
example, I have not looked at the combination of *be like* and *be all*, e.g. *be all like*; I have not
looked at the use of *be like* with adverbials, as in *be just like*; I have also not looked at
progressive forms such as *I am going*, *I was going*, etc. More complex searches could be
designed in order to capture these forms. Clearly, not having searched for these forms
implies that the total frequency of occurrence of the new quotatives as reported in this study
might be lower than it would be if these possible forms had been searched.

As mentioned in Chapter Three, I have excluded conditional, past progressive with
gonna, going to, future forms, and modalised forms, as in *you might say, she'd say*, etc.
These forms would have opened a whole new dimension – and arguably a very interesting
one – of the study, for two main reasons: 1) they would have called for searches of
“comparable” forms of the new quotatives, e.g. forms such as *I would be like, he would go,
he might be like*, etc.; 2) they would have called for a more complex framework for the
analysis of the discourse-pragmatic function, as hypothetical or unspoken discourse may
probably be classified neither as plausible nor as improbable.

This study aimed at looking at the use of the new quotatives in different registers of
spoken interaction, as represented in a corpus of casual conversation and two corpora of
consultations. While it has revealed some striking differences between the use of quotatives
in Conversation and Office Hours, this study has not allowed to identify any clearcut
differences between Conversation and Service Encounters. As mentioned in Chapter Three
and in Chapter Four, Service Encounters were indeed found to include much casual
conversation among college students and this might definitely have influenced the frequency
and use of new quotatives, as revealed by the high frequency of *be like* in the corpus. This
interpretation would also find possible evidence in the results of a study of register variation
in the TOEFL 2000 SWAL corpus, in which Service Encounters was found to be a highly involved, highly situation dependant and highly marked for nonimpersonal style register (Biber et al. 2002).

Finally, unlike previous quantitative research on the new quotatives (which, as mentioned in Chapter Two, is limited to the new quotative be like), this study has not adopted a sociolinguistic approach, in that it was not concerned with issues of age, gender, and social class of the speakers using the quotatives. This might be a limitation, especially considering that the literature on be like has shown that this quotative was most typical of the language of younger speakers (Blyth et al. 1990, Romaine & Lange 1991, Ferrara & Bell 1995, Tagliamonte & Hudson 1999, Macaulay 2001). Further, although the corpora used in this study are subcorpora of much larger and principled corpora (LSWE and TOEFL 2000 SWAL), a fact that should ensure a certain level of diversity in sex, age, and social status of the informants, some of the surprising findings reported in this present study on the distribution of the quotatives in the Service Encounters corpus point to the possibility that the Service Encounters corpus might be skewed toward younger informants, as suggested by the likelihood that the corpus mainly consists of casual conversation among college students.

**Implications for further research**

This study has raised as many questions as it has answered, and definitely points to new avenues of research. First of all, in order to obtain a clearer picture of the new quotatives and of their proportion of use within the quotative system of American English, forms in addition to simple present and simple past tense should be searched. These should
include the above mentioned combination of *be like* and *be all, be just all*, progressive forms, and quotative forms introducing hypothetical discourse.

A more comprehensive study of new quotative forms in spoken American English should include also other quotatives that have been reported in the literature, such as *be* and zero quotatives. This might require the use of tools for corpus analysis more sophisticated than concordancers, or, at least, more complex searches in order to capture forms that the researcher may not be aware of when running the searches.

Secondly, more research into the discourse function of the quotatives is needed, especially if forms introducing hypothetical or unspoken discourse are included.

Thirdly, this study points to the need for more qualitative research into the interrelationship between discourse context and direct speech. Myers (1999) and Baynham (1996) call for more research into the functions of direct quotation in different discourse contexts, and, thus, in different registers of spoken interaction. Now that corpora of spoken Academic English (e.g. TOEFL 2000 SWAL, Michigan Corpus of Academic Spoken English) have been completed, it will certainly be interesting to look at the use of the new quotatives in different registers of spoken Academic English.

Finally, some of the surprising findings on the frequency and distribution of the quotatives in the corpora used in the present study (particularly in Service Encounters) call for large-scale studies on the new quotatives of American English that include the investigation of sociolinguistic variables such as sex, age and gender, based on larger and more representative corpora of spoken American English.
Conclusion

This study has shown that the new quotatives be like, go, be all, and particularly be like and go represent a solid, stable and frequent presence in the quotative system of present-day American English. Speakers of American English use these quotatives with significant frequency in the everyday business of making speech reports in casual conversation.

The importance of this finding lies in its potential for grammatical description. As has been shown in Chapter Two, several studies have noted the inadequacy of ELT materials and traditional, non corpus-based grammars in accounting for the various choices that are open to speakers when making speech reports in everyday life. ELT materials and grammars should account for the use of forms such as the ones analyzed in this study, and should provide information about the discourse-pragmatic use of these forms. For example, the new quotative be like should be described as a direct speech introducer that when used in association with first person singular subjects may introduce inner speech or thought, and thus may be used as a substitute of the verb to think.

This study has also shown the potential of small corpora for studies of spoken data involving qualitative discourse analysis. It has shown that even relatively small corpora can yield numerous examples of new language features. Researchers and teachers can exploit these resources for research purposes, but also for the development of more authentic materials showing the use of features of the spoken language in context.

Finally, the study has shown the potential of corpus linguistics and computerized techniques for large scale discourse analysis studies. This study would indeed not have been feasible without the availability of computerized corpora, corpus analysis tools such as
concordancers, and computerized databases for the storage of retrieved and selected examples.
### APPENDIX A: OCCURRENCES IN CONVERSATION

<table>
<thead>
<tr>
<th>Grammatical Person</th>
<th>Simple Present</th>
<th>Simple Past</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SAY</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ps</td>
<td>37</td>
<td>309</td>
<td>346</td>
</tr>
<tr>
<td>2ps &amp; 2pp</td>
<td>28</td>
<td>19</td>
<td>47</td>
</tr>
<tr>
<td>3ps</td>
<td>124</td>
<td>196</td>
<td>320</td>
</tr>
<tr>
<td>1pp</td>
<td>-</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>3pp</td>
<td>17</td>
<td>25</td>
<td>42</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>207</strong></td>
<td><strong>554</strong></td>
<td><strong>761</strong></td>
</tr>
<tr>
<td><strong>BE LIKE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ps</td>
<td>92</td>
<td>71</td>
<td>163</td>
</tr>
<tr>
<td>2ps &amp; 2pp</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>3ps</td>
<td>92</td>
<td>32</td>
<td>124</td>
</tr>
<tr>
<td>1pp</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>3pp</td>
<td>3</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>193</strong></td>
<td><strong>109</strong></td>
<td><strong>302</strong></td>
</tr>
<tr>
<td><strong>GO</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ps</td>
<td>59</td>
<td>1</td>
<td>60</td>
</tr>
<tr>
<td>2ps &amp; 2pp</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>3ps</td>
<td>109</td>
<td>-</td>
<td>109</td>
</tr>
<tr>
<td>1pp</td>
<td>3</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>3pp</td>
<td>13</td>
<td>-</td>
<td>13</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>188</strong></td>
<td><strong>2</strong></td>
<td><strong>190</strong></td>
</tr>
<tr>
<td><strong>BE ALL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ps</td>
<td>2</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>2ps &amp; 2pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3ps</td>
<td>11</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>1pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>13</strong></td>
<td><strong>10</strong></td>
<td><strong>23</strong></td>
</tr>
</tbody>
</table>
## APPENDIX B: OCCURRENCES IN SERVICE ENCOUNTERS

<table>
<thead>
<tr>
<th>Grammatical Person</th>
<th>Simple Person Present</th>
<th>Simple Past</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SAY</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ps</td>
<td>-</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>2ps &amp; 2pp</td>
<td>2</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>3ps</td>
<td>3</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>1pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3pp</td>
<td>1</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>6</td>
<td>49</td>
<td>55</td>
</tr>
<tr>
<td><strong>BE LIKE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ps</td>
<td>38</td>
<td>22</td>
<td>60</td>
</tr>
<tr>
<td>2ps &amp; 2pp</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>3ps</td>
<td>38</td>
<td>16</td>
<td>54</td>
</tr>
<tr>
<td>1pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3pp</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>81</td>
<td>40</td>
<td>121</td>
</tr>
<tr>
<td><strong>GO</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ps</td>
<td>6</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>2ps &amp; 2pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3ps</td>
<td>13</td>
<td>-</td>
<td>13</td>
</tr>
<tr>
<td>1pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>19</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td><strong>BE ALL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ps</td>
<td>6</td>
<td>-</td>
<td>6</td>
</tr>
<tr>
<td>2ps &amp; 2pp</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>3ps</td>
<td>2</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>1pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>9</td>
<td>-</td>
<td>9</td>
</tr>
</tbody>
</table>
### APPENDIX C: OCCURRENCES IN OFFICE HOURS

<table>
<thead>
<tr>
<th>Grammatical Person</th>
<th>Simple Present</th>
<th>Simple Past</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SAY</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ps</td>
<td>1</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>2ps &amp; 2pp</td>
<td>10</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>3ps</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>1pp</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3pp</td>
<td>6</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>22</td>
<td>23</td>
<td>45</td>
</tr>
<tr>
<td><strong>BE LIKE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ps</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2ps &amp; 2pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3ps</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>1pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3pp</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>6</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td><strong>GO</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ps</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2ps &amp; 2pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3ps</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>BE ALL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1ps</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2ps &amp; 2pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3ps</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3pp</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
REFERENCES


ACKNOWLEDGMENTS

This study would not have been possible without permission to use the copyrighted corpora it is based on. Thank you ETS: Permission to use the TOEFL 2000 SWAL Corpus was granted by Educational Testing Service, the copyright owner. Thank you also to Pearson Education Ltd., for granting the use of corpus files of the LSWE Corpus for research purposes.

I also owe thanks to many people who have helped me in various ways to carry out this study.

First and foremost, thanks to my major professor, Dan Douglas, for being a patient, caring and always available presence. Thank you, Dan, for help with the write up of this thesis, for valuable comments on earlier drafts, and for encouraging me throughout this work.

Very special thanks to Susan Conrad, who has been much more than a “long distance” committee member and has continued to be a real mentor even after moving to Portland. Susan’s own commitment, devotion, and enthusiasm for research in Corpus Linguistics have spurred my interest in Corpus Linguistics and moulded it into a more rigorous and aware attitude towards language research. Thank you, Susan, for providing me with the corpus files from the Longman Corpus and encouraging me to venture into the study of spoken language. Thank you for helping me develop a few of the ideas at the heart of this study and challenging my thinking in many ways.

Many thanks to Carol Chapelle for thorough comments and challenging questions that have greatly helped me improve this study. Thank you also to Fred Lorenz, whose
“outsider’s” outlook has helped me see some aspects of this study from a different perspective.

Thank you to Anne O’Keeffe for sending me data on quotative use in the Limerick Corpus of Irish English, and to Sylvia Chou for help with some references and very helpful discussions on reported speech.

Last but not least, thank you to all those friends who have encouraged me and supported me in various ways in the past months. In particular, I wish to thank Silvia Spanoyannis, Joon Sang, Giulia Cipressi and Suzanne Eckhardt.