Studies in discourse and grammar

Point of View and Grammar

Structural patterns of subjectivity in American English conversation

Joanne Scheibman

Point of View and Grammar

Studies in Discourse and Grammar

Studies in Discourse and Grammar is a monograph series providing a forum for research on grammar as it emerges from and is accounted for by discourse contexts. The assumption underlying the series is that corpora reflecting language as it is actually used are necessary, not only for the verification of grammatical analyses, but also for understanding how the regularities we think of as grammar emerge from communicative needs.

Research in discourse and grammar draws upon both spoken and written corpora, and it is typically, though not necessarily, quantitative. Monographs in the series propose explanations for grammatical regularities in terms of recurrent discourse patterns, which reflect communicative needs, both informational and socio-cultural.

Volume 11

Point of View and Grammar: Structural patterns of subjectivity
in American English conversation
by Joanne Scheibman

Point of View and Grammar

Structural patterns of subjectivity in American English conversation

Joanne Scheibman
Old Dominion University

John Benjamins Publishing Company
Amsterdam/Philadelphia

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Scheibman, Joanne

Point of View and Grammar : Structural patterns of subjectivity in American English conversation / Joanne Scheibman.

p. cm. (Studies in Discourse and Grammar, ISSN 0928–8929 ; v. 11)

Includes bibliographical references and indexes.

1. English language--Spoken English--United States. 2. English language--Social aspects--United States. 3. English language--United States--Grammar. 4. English language--Discourse analysis. 5. Speech acts (Linguistics) 6. Interpersonal relations. 7. Conversation. I. Title. II. Series.

PE2808.8 S34 2002

401'.41-dc21 2002074558

ISBN 90 272 2621 0 (Eur.) / 1 58811 232 2 (US) (Hb; alk. paper)

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands

John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

To Miss Vick

Contents

Acknowledgements

I am grateful to my teachers, colleagues, and friends for invaluable discussions and support throughout the various stages of this project. In particular I am indebted to Joan Bybee, Sandy Thompson, Paul Hopper, Anna Gordon, Melissa Axelrod, Sherman Wilcox, and Scott Williams. Portions of this book appeared previously in an article published in *Frequency and the Emergence of Linguistic Structure* under the title "Local Patterns of Subjectivity in Person and Verb Type in American English Conversation", © 2001 by John Benjamins Publishing Co.

List of tables

Chapter 1

Linguistic subjectivity and usage-based linguistics

> "[T]he subjectivity of utterance" might seem to be neither novel nor interesting … it is something we have all been operating with throughout life without having a specialized technical term for it. When we come to investigate locutionary subjectivity within the framework of modern linguistics, however, we soon discover that it is far from being as simple and straightforward as it might appear to be at first sight. (Lyons 1994: 13–14)

1.1. Introduction to the study

A common observation that occurs to analysts transcribing or analyzing English conversational data is how little *information* seems to be passed among participants during the course of their interactions. Where in English interactive discourse are the semantically- and lexically-substantive propositions that speakers have been assumed to be trading? And why is it that in spontaneous conversation participants use a limited repertoire of linguistic constructions when linguists know, based on years of good generative thinking, that language users have access to an infinite number of structural combinations?

In English conversation, participants are not simply, nor very often really, reporting on events that take place in the world in an unmediated manner. It is the case that what English speakers are consistently conveying in their talk are evaluations, opinions, and attitudes—in short, their points of view—constrained by the exigencies of face-to-face communication. Furthermore, this ubiquitous expression of speaker point of view and the concomitant personalization of utterances has resulted in conventionalized structural patterns that may be seen in the distribution of grammatical and lexical elements in English conversation. It is these frequent structural patterns in interactive discourse—those that express subjective stance—that form the basis for the analyses offered in the present work.

This study, then, takes as its starting point the notion that language—in particular casual conversation—is subjective to varying degrees. This idea diverges from a point of view implicit in many linguistic approaches that assumes the most common function of language to be propositional, or referential. Within such orientations, linguistic expressions are tacitly treated as independent meaning packages separate in function and structure from context and speakers, and they are assumed to be used by participants to provide faithful descriptions of events and actors in the world. When we look at the actual occurrence and combination of lexical and grammatical elements in conversational English, however, what we find is that expression of speaker stance and its impact on the structure of interaction is not separate from the reporting of happenings and actors in the world.

The theory and practice that inform this study are usage-based. Generally speaking, usage-based approaches in linguistics view language structure as rooted in actual usage, and the frequency/repetition of linguistic expressions in natural discourse is regarded as an important factor in the conventionalization of linguistic form (see section 1.3.). Investigation of language use in naturalistic contexts—in this case, in conversation—necessarily focuses attention on the functions and context of linguistic expressions, and consequently, on discourse participants themselves. Analytical consideration of speaker stance, then, is theoretically important when working within a framework that views usage as a key influence on the conventionalization of linguistic form.
The intersection of the two general theoretical issues referred to above—(1) that language (in particular, spontaneous conversation) is subjective in that it is fundamentally used by speakers to express their perceptions, feelings, and opinions, and (2) that conventionalized linguistic structure, or grammar, emerges from repetition, or frequency of use, of (sequences of) lexical and grammatical elements in natural discourse—suggests a general hypothesis that linguistic items (constructions of all sizes) that most frequently appear in conversation are those that participate in subjective expression. In other words, we would expect greater cooccurrence of elements in conversation whose combinations lend themselves to conveying speaker point of view than those whose combinations do not. It is these two central claims, then, that motivate the analyses in this study, which proceed by examining the occurrence and combinations of lexical and grammatical elements found most frequently in a database of English conversational utterances.

The rest of chapter 1 provides a theoretical foundation for the study by

situating it with respect to previous research on linguistic subjectivity and other usage-based investigations, in particular, within the area of conversation and grammar. Chapter 2 presents discussion of data management and coding procedures used in this work. It outlines in detail the classifying and coding of structural, functional, and semantic properties of utterances in the conversational database. The coding used for this study—including its implementation and justification—is a theoretically important part of the investigation itself. Because linguistic categories have typically been derived without discourse data (Hopper 1997), the fitness, or sometimes lack of fitness, between analytical categories and the conversational utterances they are applied to provides a potential test of the bases of grammatical and lexical classes within interactive discourse. A benefit of a classification system such as this one—one that includes many formal and functional categories—is that the broad range of coding fields allows an analyst to look at a variety of structural combinations in a given context. This flexibility facilitates investigation of more local distributional patterns found in the data.

Chapters 3 and 4 present the analyses of the conversational data, and the division between these two chapters is based on semantic verb type. Chapter 3 is an examination of the most commonly occurring lexical verb types by subject of the utterance, excluding relational predicates. Conversely, chapter 4 focuses exclusively on relational utterances (copular clauses) which are the most frequent clause type found in the conversational corpus. And finally, chapter 5 offers summaries and conclusions and presents implications for further research in linguistic subjectivity and for the study of language structure in general.

1.2. Subjectivity, propositionality, and linguistic theory and practice

1.2.1 Linguistic subjectivity

In his introduction to the volume, *Subjecthood and subjectivity*, John Lyons (1994) remarks that interest in linguistic subjectivity is currently fashionable, and indeed, in the last several years, there have been several studies highlighting ways in which speakers use language to express their perceptions, feelings, and opinions in discourse (subjectivity) and how such expressive motivations and strategies conventionalize and interact with linguistic structure (subjectification). Verhagen (1995:116) suggests that the most usual mode of

expression in language is in fact a subjective one—that "the presence of some argumentational orientation is the default situation in natural language", and he notes that even lexical items like *expensive* and *tall* orient an addressee towards a particular conclusion and are not solely informative.

In a discussion of linguistic subjectivity, Benveniste (1971: 225) writes: "Language is marked so deeply by the expression of subjectivity that one might ask if it could still function and be called language if it were constructed otherwise". He characterizes subjectivity as the ability of speakers to view themselves as subjects, and he discusses ways in which common grammatical categories—in particular, person (in the form of personal pronouns)—contribute to this expressive capacity of speakers. Benveniste notes that the first person singular pronoun is a rich source of subjectivity in language in that it explicitly refers to the speaker. He discusses meaning distinctions present in utterances that contain the same verb but contrast in subject. For example, *I* with verbs such as *feel*, *believe,* and *suppose* typically express the speaker's attitude regarding a subsequent piece of discourse or an event in the current context; on the other hand, when these same verbs occur with the third person singular subjects *she* or *he*, there is an impression that what is conveyed is descriptive or informative. Benveniste also identifies deictic terms and tense as subjective in that they make spatial and temporal reference to the speaker.

Like Benveniste, Lyons includes tense and deixis in his characterization of subjective phenomena because they make reference to the discourse context, and consequently, to the speaker as well. In a discussion of *locutionary subjectivity*, Lyons (1982) disputes the notion that linguistic self-expression may be equated with the assertion of a set of propositions. He claims that there is a distinction in language between subjective utterances (the speaker's expression of herself) and objective utterances (a set of communicable propositions), and that the distinction between these types is gradual and not absolute.[1] To illustrate the contrast in English between subjective and objective utterances, Lyons (1982: 107) provides the following minimal pair. According to Lyons, example (1) conveys a subjective point of view in that it reveals the speaker's consciousness, and (2) is more objective since it is interpreted as a report of an event in which the speaker mentions herself as a participant—a more distancing rendition.

1. I remember switching off the light.
2. I remember *myself* switching off the light.

Working in the theory of Cognitive Grammar, Langacker (1990; 1991) also suggests that sentences illustrated by (1) are more subjective than those like (2). Langacker (1990; 1991) defines subjectivity and objectivity as viewing relations between a perceiver and an object of perception. Maximally subjective viewing characterizes a situation in which the perceived object is "implicit and hence non-salient—to use the theater metaphor, it remains *offstage* in the audience [e.g. (1)]—whereas the objectively-construed entity is salient by virtue of being placed *onstage* as the focus of attention [e.g. (2)]" (Langacker 1990: 7). Similar to Lyon's analysis above, Langacker's characterization of subjectivity is primarily a cognitive one and does not make explicit reference to subjective expression in interactive contexts.

There are also usage-based studies that have uncovered strong associations between linguistic structure and speaker expression. For example, in a study of American English conversation, Thompson and Mulac (1991) report on the grammaticization of first person singular subjects with the main clause predicates, *think* and *guess (I think, I guess)* as epistemic parentheticals. Iwasaki (1993) demonstrates how speakers' perspective influences tense and clause chaining in Japanese narratives. Scheibman (2000) reports that the negative auxiliary *don't* consistently appears in a reduced form when it occurs in its most frequent context—in the collocation *I don't know*—where it primarily functions as an epistemic downtoner, or politeness marker (see also Tsui 1991). And based on a large corpora study, Tao (2001) reports that English *remember* most frequently appears with *I + don't*, and, similar to *I don't know*, is mainly used epistemically in English conversation.

In an analysis of English written narrative, Hopper (1991; 1997) observes that simple verbs (single lexical items) are relatively rare in discourse, and that much more common are dispersed predicates, or *MAVES* (Multiply Articulated Verbal Expressions), especially in language in which the speaker is highly involved. In an investigation of egophoric expressions (linguistic elements that refer to speech act participants) and generic pronouns in Swedish and English conversation, Dahl (2000) finds that linguistic expressions making reference to speech act participants, particularly the speaker, occur frequently in conversation, and that utterances with first person singular subjects, for example, show different distributions with respect to lexical verb type in comparison to those with third person subjects.

1.2.2 Subjectification in grammaticization

A particularly important and well-documented contribution to work in linguistic subjectivity, proposed and elaborated by Elizabeth Traugott, is the phenomenon of subjectification in grammaticization processes whereby over time meanings become increasingly based in speakers' attitudes towards what they are saying (e.g. Traugott 1988; Traugott 1989; Traugott 1995; Traugott & König 1991). Traugott finds evidence for subjectification in her diachronic analyses of English modals, connectives, epistemic speech act verbs, modal adverbs, and evidentials. Traugott includes as subjectification cases in which meanings have taken on metalinguistic, or textual, functions. For example, in Middle English, *while* went from coding temporal simultaneity to functioning as a connective; Traugott (1995: 40) notes that this newer function is more subjective because "instead of profiling a specific time, it now profiles discourse structure". Traugott includes as subjective not only expressions that specifically convey speaker point of view, but also metalinguistic elements that make general reference to the speech context and frequently have a textual function.

Other authors have explored subjective influences on grammaticizing elements. In her study of the English Progressive, Wright (1995) examines an epistolary portion of a 17th century novel and finds that the experiential Progressive occurs frequently in the present tense with first person subjects. Wright notes that the choice of the progressive draws attention to the speaker's emotional experience regarding what is being talked about (see also Hatcher 1951). Similarly, in her analysis of the development of the English Present Perfect from the Old English Resultative construction, Carey (1995) reports that in late Old English, the most frequent verb classes that occur with the *have* + participle construction are verbs of communication and perception. Her data show that almost all of the communication verbs appear with first person subjects (*Now I have (often) said/explained x)* and nearly all of the perception verbs occur with second person singular subjects (*Now you have often heard about x*) (Carey 1995: 92). Carey claims that since perfect constructions relate a past event to the current speech situation, it makes sense that notions of current relevance would develop with verbs that directly relate to the ongoing discourse. In later "hot news" stages of perfect constructions, Carey suggests the event's relevance to the current discourse is no longer tied to the semantics of the participle, but is wholly contingent on the speaker's judgment.

In this selected review of research on linguistic subjectivity, we see that characterizations of subjective phenomena are diverse. What all of these cases demonstrate, however, is the important influence speaker point of view has on linguistic structure. This is not surprising given the undeniable connection between speakers and their utterances in all forms of discourse. In this study I show that the varied relations between linguistic structure and speaker point of view are neither exceptional nor unique, but rather the norm in conversational English.

1.2.3 Functions of language and the privileged role of propositional transmission in linguistic theory

Notwithstanding the ubiquity of subjective expression in discourse (the conveying of mental states, affect, preference, evaluation) the *propositional*, or *referential*, function of language (the communication of information and ideas) has been promoted in the practice of linguistic analysis and elevated in discussions of the functions of language. Historically, much linguistic and philological inquiry has relied on the metaphorical assumption that linguistic structure in its assorted chunks contains meaning independent of context or participant stance. These form-meaning packets, then, are combined and exchanged by speakers in the shape of words, phrases, and clauses in order to transmit information from one reflective mind to another—this being the fundamental act of communication: the transfer of propositional material encoded in linguistic form from one human being to another (see Reddy 1993, e.g., for a critique of this theoretical point of view). This function of language—the communication of ideas—has variously been called *descriptive, referential, propositional, ideational,* and *designative* (Lyons 1977), and it has been elevated in many discussions of the functions of language.

An important contribution to the literature on language functions is Jakobson's (1960) typology of the functions of language which he derives from a model of a communicative event. The elements of the model consist of an addresser who sends a message (which makes reference to something in the world) to an addressee; a code shared by both participants; psychological and social reasons for the participants to be in contact, and a physical channel through which the message is transmitted. These aspects of the communicative event determine the different functions of language (e.g. focus on the addresser=*emotive* function, focus on the context=*referential* function).[2]

Jakobson (1960: 342) notes that though it is the *referential* function that is "the leading task of numerous messages", "accessory participation" of other language functions should be taken into account. The *emotive*, or expressive, function in Jakobson's scheme characterizes the speaker's point of view and emotions, and Jakobson allows that expression of speaker attitude "conveys ostensible information". He adds, however, that in their purest forms, expressive utterances tend to be phonologically unique (e.g. *tsk tsk* in English, ideophones in Niger-Congo languages). This notion—that pure expressivity is a response to internal or external stimuli and is outside of the linguistic system—is discussed from various points of view in other works as well (e.g. Haiman 1998; Martinet 1991). These examples illustrate a tradition in linguistic theory of assuming a crisp division between the expression of information (accomplished with language) and the expression of emotion (structurally and functionally on the margins of language).

In his analysis of indexicals and reference, Silverstein (1976: 14) claims that pure reference (descriptive language) likely plays a minor communicative role cross-culturally, though it has "formed the basis for linguistic theory and linguistic analysis in the Western tradition" (e.g. Lyons 1977; Ogden & Richards 1923). To illustrate: the notion of compositionality—that grammatical and lexical bits contribute to the meanings of propositions (e.g. English *-s* contributes the meaning of 'more than one' to some argument of a proposition)—reflects the referential mode that has predominantly characterized linguistic analysis. We assign meaning to linguistic material separate from social and linguistic context. Similarly, if we label an aspectual category *progressive*, we are saying that a particular segment adds the sense of 'being in progress' to an event described in a proposition. In other words, as analysts, we segment and slot linguistic *substance* based on its structural and semantic contribution to what we envision as the expression of a proposition. Not only has linguistic theory been biased by this referential view of grammatical and semantic categorization, but most material that has served as data for such analyses has itself been descriptive, or referential, in nature (e.g. hypothetical sentences of the *John kissed Mary* variety, written language). In fact, it is likely that the use of this sort of propositionally-full data in linguistic research strongly influences the referential bias in analysis (and vice versa).

In a discussion of linguistic signs of self-alienation, John Haiman (1995: 214) points out that "speakers using language in general are alienated from the emotions they describe if they control them sufficiently to use

language to describe them". So in some sense language presents a paradox—human expression is inherently subjective, but in using a conventional code, we objectify this expression. Given, then, that language always represents a speaker's interpretation of the world, investigating linguistic structure from the perspective that all language is subjective to varying degrees (as opposed to assuming fundamental propositionality) may not only contribute to the characterization of linguistic subjectivity itself, but may enhance what we know about the conventionalization of meaning and interaction in context.

1.2.4 Beyond propositionality

Before closing section 1.2., it is important to make clear that while there is a strong tradition of viewing speaker expressivity as a phenomenon separate from and marginal to a more dominant referential function within linguistic theory, there is also important work—as well as whole subfields of linguistics—that rely on the understanding that language and communication are much more than the exchange of formal propositional material for purposes of conveying information.

In contrast to theoretical positions that privilege referentiality within linguistic analysis, Vološinov (1973: 99–106) argues unequivocally against an analytical division between referential and evaluative meaning in language. He suggests that scholars' focus on language as an abstract system "is what led most linguists to divorce evaluation from meaning and to consider evaluation an accessory factor of meaning, the expression of a speaker's individual attitude toward the subject matter of his discourse" (Vološinov 1973: 105). Vološinov continues:

> This sort of disjuncture between referential meaning and evaluation is totally inadmissible. It stems from failure to note the more profound functions of evaluation in speech. Referential meaning is molded by evaluation; it is evaluation, after all, which determines that a particular referential meaning may enter the purview of speakers—both the immediate purview and the broader social purview of the particular social group. Furthermore, with respect to changes of meaning, it is precisely evaluation that plays the creative role. A change in meaning is, essentially, always a *reevaluation*: the transposition of some particular word from one evaluative context to another. (Vološinov 1973: 105)

Vološinov contends that meanings of words and utterances are not separate from value judgments associated with them. He discusses several ways in which evaluation is integral—not marginal—to (the study of) language and

meaning. For example, at the level of local interaction, it is individual speakers' value judgments that in part determine which utterances and elements of utterances are actually said and heard. At a more macro sociolinguistic level, Vološinov observes that meaning itself is always associated with the values of a particular social group; thus it has political import (1973: 106). And finally, Vološinov attributes creative motivation for language change itself to speakers' evaluations.

Malinowski (1923) claims that it is not possible to adequately describe meaning detached from the social and expressive exigencies of actual usage. He suggests that without consideration of the *context of situation*—the realities of culture, custom, and situation in which language is naturally embedded—an utterance has no meaning. Rommetveit (1974) argues that the propositional meaning of an utterance varies from context to context, and he observes that so-called incomplete linguistic forms (e.g. ellipsis) and nonliteral language have meaning in context due to the *sharing of social worlds* (the establishment of an intersubjective framework) that occurs among participants in communicative settings. Similarly, Clark (1996) discusses the importance of building *common ground* between speaker and addressee for the interpretation of meaning in English conversation.

And finally, there are entire areas of language study that focus on the nonpropositional functions of language. For example, speech act theorists discuss *illocutionary force* to underscore speakers' acts and intentions in discourse. Conversation Analysts organize their investigations of social interaction around a turn-taking mechanism and sequential, alternating pairs of utterances. And interactional sociolinguists are concerned with the situatedness of language, social activity, and context (e.g. Schiffrin 1994). The present study also aims to treat linguistic structure as contextually complex and attempts to de-emphasize the analytical divisions that often exist between participants and linguistic material.[3]

1.3. Emergence of linguistic structure and usage-based linguistics

1.3.1 Overview

One of the most important contributions to recent linguistic scholarship is the set of theories and practices that view language structure as rooted in usage—

thus treating it as dynamic in both its manifestation and in its development and change. In his classic paper, "Emergent grammar", Paul Hopper (1987) argues that linguistic form is neither fixed nor aprioristically determined; rather structure is shaped (often in prefabricated chunks) by discourse use, and this process is itself ongoing. Hopper notes that "the more useful a construction is, the more it will tend to become structuralized, in the sense of achieving cross-textual consistency, and serving as a basis for variation and extension" (1987: 150).

This general usage-based framework illustrated by Hopper's work and formulated and elaborated by many linguists (e.g. Bybee, To appear; Haiman 1998; Langacker 1987), emphasizes the role of frequency, or repetition, in the formation of what we call grammar. This theoretical orientation subsumes several overlapping lines of linguistic research, such as: diachronic and synchronic investigations of grammaticizing material (e.g. Bybee, Perkins, & Pagliuca 1994; Thompson & Mulac 1991), repetition and the ritualization of linguistic form (e.g. Bybee, To appear; Haiman 1998), the impact of social interaction on grammatical structure (e.g. Ford 1993; Fox 1987; Ono & Thompson 1995), and studies of cognitive processing mechanisms affiliated with frequency and storage leading to the conventionalization of form (e.g. Boyland 1996; Bybee & Scheibman 1999; Krug 1998).[4] What appears to unite these various approaches is a common goal of connecting larger, analytically delineated patterns (=linguistic structure) with the sounds, gestures, and timing characteristic of local usage activities.

1.3.2 Conversation and grammar

Usage-based investigations that are particularly relevant to the foundation and method of the present study are those that endeavor to describe the grammar of conversation, and in doing so, are able to test and elaborate extant linguistic units and categories. Conversation is the most frequent type of linguistic activity that people engage in, and as such, is considered to be the site from which conventionalized structure emerges. This recently established field of conversation and grammar, in practice, forces a merging of a taxonomic analytical approach (characteristic of traditional structural analysis) with more pragmatic investigations of context and participant interaction. Researchers of conversation and grammar come from several disciplines: functional linguistics (e.g. Kärkkäinen 1998; Ono & Thompson 1995), anthro-

pology (e.g. Duranti 1994), and sociology, specifically Conversation Analysis (CA) (e.g. Schegloff 1996).

The interactive, dynamic, and prosodic characteristics of spontaneous conversation provide unique challenges for the study of grammar. Examples of topics that have received attention from scholars in this area are (1) formal units and categories in conversation and the role of intonation on structure, (2) form-function relations in interactive discourse, and (3) the co-construction of utterances.

Chafe (1993; 1994) and others (e.g. Du Bois, Schuetze-Coburn, Cumming, & Paolino 1993) have argued for the *intonation unit*—linguistic material uttered under one intonation contour—as a basic structure in conversation (see chapter 2, section 2.3.1.). Sacks, Schegloff, and Jefferson (1974: 702), however, posit the existence of turn-constructional units in English conversation (e.g. lexical items or syntactic categories), instances of which "so usable allow a projection of the unit-type under way, and what, roughly, it will take for an instance of that unit-type to be completed". The idea is that turn changes occur at these completion points. Ford and Thompson (1996) test the claim put forth in Sacks et al. (1974) that speaker changes occur at points of syntactic completion. In a twenty minute conversation, Ford and Thompson find that 71 percent of speaker changes occur at places where both grammatical and intonational closures (IU boundaries) coincide. They also report that syntactic boundaries on their own are the *least* reliable indicator of any sort of closure (intonational, semantic) or turn change. Results from Ford and Thompson indicate that if we use grammar to extend and change conversational turns, then intonation is an important part of grammar.

In a study of questions in English conversation, Weber (1993) finds that 41 percent of the questions in her corpus do not occur as interrogative forms; instead questions are expressed (and interpreted) in a variety of ways (e.g. as clauses, phrases, or lexical items). For example, Weber suggests that the pronoun *you* is often a marker of question function because it appears most often in interrogative forms and is therefore associated with questioning. Typically in structural analysis analysts look for one (or perhaps two) mechanisms that mark questions in a particular language: a morpheme, a lexical item or particle, or a change in word order. Weber's study contends that in English conversation there are a variety of conventional ways that participants code questions (e.g. interrogative, declarative, and nonclausal forms).

In a study based on English conversation, Thompson and Hopper (2001) find that only 27 percent of clauses in their corpus have two participants, and in those, transitivity is very low. Thompson and Hopper comment:

> We would be the first to acknowledge the range of facts that have emerged from assuming highly Transitive clauses to be the starting point for the study of grammatical relations and voice and valency-changing operations; at the same time, we are struck by the fact that most of the attention in current functional linguistics, not to mention more structurally oriented research, is currently being given to patterns whose instantiations are vanishingly rare in ordinary talk. (Thompson & Hopper 2001:52)

Similar to the investigations of grammar and conversation summarized in this section, the present study also illustrates the kinds of disparities or "clashes" that can occur in linguistic analysis when conversational data are introduced into the analytic process (Hopper 1997).

There is a tacit assumption, even among functional and cognitive linguists, that producing and interpreting linguistic material occurs in the separate heads of communicators—that messages are formulated by a speaker then transmitted to an addressee. Because of this assumption, structural descriptions do not typically make reference to an interactional context. But in spontaneous conversation, much depends on relationships among the participants: speakers negotiate and ratify one another's contributions based on individual and sociocultural expectations, conversational participants attend to the effects their utterances have on their interlocutors, and such cooperation manifests linguistically.[5] In their search for evidence of constructional schemas in conversation—schematic structural categories with cognitive status derived from actual instantiations of those categories in discourse—Ono and Thompson (1995) examine utterances which are co-constructed by two participants. In the following example, B expands A's unit by contributing a relative clause.

(3) A: ..actually,
they just went out to <% Chisera= %>,
.. to go [out to the river].
B: [**which is a hundred miles**],
in the [**2 bush 2**].
A: [2 it's 2] about a hundred miles away,
.. and they w- were just going to go up to the river.
(Ono & Thompson 1995:228)

These co-constructions are not infrequent in English conversation and they provide a striking illustration of collaborative activity in discourse. Though the primary goal of the present study is the analysis of the structure of speaker stance, the demands of an interactive context are also found to impact linguistic expression of subjectivity in a variety of ways.

1.3.3 Local patterning

Given the idea that both lexical and discursive substance in context form the basis for grammatical organization (i.e. that what we know as grammar arises from both the conventionalization and creativity of human interaction), then how successful may analysts expect to be in explaining global patterns of linguistic form and function when confronted with the often interdependent and opaque elements of complex local activities? Relative to more formal (aprioristic) theoretical approaches, one wonders whether the practice of usage-based linguistics brings with it—aside from obvious differences in classification and analysis of naturally-occurring data that this kind of work entails—a less glamorous, or less global, type of linguistic generalization, or perhaps a shift in expectation regarding our capacity to generalize, simply because of the natural complexity of the material serving as input to (the description of) linguistic patterns (cf. Hopper 1997).

If interactive and lexical material direct grammar, and conventionalization of form occurs within specific social and linguistic contexts (e.g. Bybee 1998), then we should expect to find differences in behavior (distribution, function) among members of the same grammatical and lexical categories (e.g. we shouldn't be surprised to find that the subject pronouns in English occur in distinct linguistic and interactional environments and fulfill different functions in discourse, even given their common membership in the grammatical class of person) (cf. Bybee 1985). If grammar emerges from linguistic and interactive contexts, then tracking linguistic categories within these individual contexts becomes important.

Local patterns, then, contribute to the formation of more global patterns. However, it might not always be possible to find reliable, unified explanations for global structural organization of linguistic form for several reasons: (1) because of the unique nature of component contributions (i.e. that the whole is neither equal to nor derivable from the sum of its parts), (2) because of limitations of extant analytical categories that have typically been delineated

without having been derived from natural discourse data, and (3) because there are contextual factors that are invariably unaccounted for in any analysis. Another potential obstacle to fully accounting for global linguistic patterns is suggested by recent work on the emergence of complex adaptive systems. Simply stated, complex adaptive systems characterize the behavior of biological organisms (along with their attendant social and cognitive organizations and structures; e.g. political systems, financial markets, thinking, learning, and language) as collective entities that function by schematizing and adapting to experiences in order to function and persist in the world. And, notably, complex adaptive systems themselves have the capacity to engender greater complexity (Gell-Mann 1994). With respect to language structure, then, it is possible that some global patterns have themselves emerged from the complexity of more local systems and may not be transparently reducible to their input (tokens of usage). One of the goals of the present study is to look at local (often, lexical) contexts in order to account for global frequency distributions found in the data.

1.4. Summary

This chapter provides discussion of the theoretical concepts and practices that inform this work. The primary focus of this investigation is grounded in the notion that speakers primarily use language subjectively, to express their points of view—not to present unmediated descriptions of the world. This is especially true for interactive discourse. Furthermore, given that the theoretical foundation of this study is one that assumes that linguistic structure emerges from processes of conventionalization of form and function based on frequency of use, it follows that the most common elements and constructions occurring in English conversation should be those that reflect broad-based subjectivity. Distributional, semantic, and pragmatic analyses offered in subsequent chapters demonstrate that the most frequently occurring conversational structures are indeed those that that are used to express speaker point of view. Additionally, because the data used for this study are from actual conversations, the coding and analyses presented here contribute to the larger effort of analyzing grammar as contextually complex.

Notes

1. Lyons adds, however, 'that what is here described as objective is, in origin, intersubjective, so that language is even more deeply imbued with subjectivity than I am supposing' (1982: 105).
2. By *context*, Jakobson means referent.
3. It is the case, however, that the methods used in this study (e.g. audio-taped conversation, linear graphic transcripts, structural coding of utterances) invariably distance the data and analyses from the speakers.
4. As is true within any theoretical framework, not all researchers in these related fields uniformly agree on the various factors claimed to influence the development of grammatical and lexical constructions (nor, if they do, on the ranking of these factors).
5. See Clark (1992) for experimental work in collaborative construction of meaning and issues related to audience design.

Chapter 2

Classification and coding of conversational data

2.1. Conversation as linguistic data

Fundamental to usage-based analyses of linguistic structure is the requirement that investigations be based on naturally occurring data. Conversation, as the most common site of language use, is an important source of data for studying linguistic contexts in which structural patterns emerge and conventionalize. Current research suggests that there are not always reliable correspondences between grammatical categories and constituents as they are traditionally characterized and their appearance and function in natural discourse (Hopper 1997). For example, in a study based on analyses of audio- and video-taped face-to-face and telephone conversations, Fox (1987) argues that general anaphoric patterns of English are a result of local management of talk on the part of conversational participants. The traditional claim that a speaker's decision to use a particular anaphor is correlated with the distance between coreferential devices (which can be measured, e.g., by counting the number of clauses between mentions (e.g. Givón 1990)) does not adequately account for the distribution of these forms in conversation. Fox demonstrates that the basic pattern of anaphora in English conversation is that the first mention of a referent occurs with a full noun phrase; after that, a pronoun is used to indicate that the speaker understands that an interactional sequence is not closed.

In a study based on conversational Japanese, Fujii and Ono (2000) find that 70 percent of object noun phrases are not expressed with the accusative marker *-o*. In interactive discourse, *-o* only appears when objects are discourse referential (e.g. demonstrative + NP) or when the object is highly specified (when it occurs after the verb). These findings, however, are inconsistent with previous grammatical descriptions that characterize Japanese case marking as obligatory (e.g. Shibatani 1990) and indicate that the decision to use *-o* or not depends on the conversational context.

Fox 1987 and Fujii and Ono 2000 provide illustrations of how decontextualized descriptions of linguistic structure may overgeneralize the distribution or properties of grammatical expressions. Additionally, without the benefit of natural discourse data, hypotheses concerning the emergence, or formation, of grammatical and lexical forms and sequences of structural elements run the risk of being descriptively incomplete.

An advantage, then, of introducing interactive discourse into formal investigations of language is that these data allow analysts to view details of local usage activities that contribute to the formation of more general linguistic patterns. However, there are also difficulties in using conversational utterances to investigate theoretically extant grammatical classes. In part this is because units of grammar are not necessarily units in conversation. Even when traditional grammatical elements do appear regularly in interactive discourse, it can be the case that the identification and distributions of these forms are not consistent with those found in traditional linguistic descriptions. This dilemma may be particularly relevant to researchers working in the area of conversation and grammar. Other examinations of language based on recorded, interactive data—for example, those conducted by Conversation Analysts—are not typically geared towards quantitative investigations of morphosyntax. In line with CA goals of "discover[ing] the systematic properties of the sequential organization of talk" (Levinson 1983: 287), Conversation Analyst researchers tend to focus on interactive units (e.g. turns, sequences of turns) in their analyses, not grammatical units.

Imposing on conversational utterances classifications that were largely developed to categorize linguistic elements without benefit of looking at actual usage highlights disparities between two general theoretical orientations: *referentially descriptive* versus *usage-based*. *Referentially descriptive* refers to the more traditional type of classification of linguistic elements—analyses based, for example, on the construction and modeling of paradigms. In *usage-based* approaches, extant analytical categories may present difficulties for the structural and semantic coding of lexical and grammatical material as it is found in interactive discourse. For example, casual conversation is replete with repetitions and restarts which can make labeling of an utterance's major constituents (e.g. subject, main verb) difficult because there may be more than one of these elements in a locution. Additionally, depending on the utterance, the same form may alternately represent two (or more) semantic types (e.g.

the lexical verb *tell* may be a verb of speaking, as in *she told me*, or a verb of perception when it occurs with the modal auxiliaries, *can* and *can't*, as in *I can('t) tell*). Unlike traditional grammatical discussions of the coordinating conjunction *and*, this element rarely (if ever) functions to conjoin clauses of pragmatically equal status in English conversation (Schiffrin 1987). And finally, it is often the case that the semantic and syntactic distinctions between auxiliary-like and main verbs (e.g. *keep, get, afford, start*) are not at all easy to determine in interactive discourse. These kinds of discontinuities found between *referentially descriptive* categories and conversational elements will be discussed in detail in section 2.4.

2.2. Sources of data

The data for this study consist of nine audiotaped informal conversations among friends and/or family members. Eight of the nine tapes and their transcripts were provided courtesy of the *Santa Barbara Corpus of Spoken American English*.[1] One tape (*Girls*) was recorded and transcribed by me. In all, eighty minutes of conversation were coded for analysis—ten minutes from each of seven tapes and five minutes from each of two tapes—yielding a total of 2,425 utterances. Except in one case, where it was not possible, conversational utterances were not coded until participants were two minutes or more into their talk, thus allowing speakers a little time to become used to being recorded.

Represented in this corpus is the language of 33 adult speakers of American English, 21 women and 12 men. The rationale for coding short sections from several tapes as opposed to coding longer episodes from fewer tapes was to maximize both the number of speakers and conversations. General details of the database are summarized in Table 2.1.

Since the data are audio, gestural and other kinesthetic information do not contribute to the analyses presented in this work. There is little doubt that nonlinguistic aspects of conversational contexts, such as eye gaze, movement, facial expression, and arrangement and composition of the physical context, are important to interaction (e.g. Goodwin 1981) as well as to the expression of speaker stance. Though the absence of this information is a limitation of the present study, omission of these details does not impede the specific goal of

Table 2.1. Data sources.

Conversation (Abbreviation)	Number of Speakers	Women	Men	Minutes coded	Number of Utterances
Appease the Monster (A)	5	3	2	10	345
Cafe Virago (C)	3	2	1	10	296
Deadly Diseases (D)	3	2	1	10	373
Farm Talk (F)	2	0	2	10	284
Girls (G)	4	4	0	10	259
Hypochondria (H)	3	2	1	5	112
Lambada (L)	4	1	3	10	309
Mable (M)	4	3	1	5	130
Raging Bureaucracy (R)	5	4	1	10	317
Totals (9 conversations)	33	21	12	80	2425

this work which is to investigate the construction of speaker point of view by looking at the frequency and cooccurrence of lexical and grammatical elements.

2.3. Transcription

2.3.1 Intonation units

All conversations included in the corpus were transcribed using the Du Bois et al. (1993) system. This transcription system is unique in its segmentation of discourse into intonation units (IUs). Intonation units are conversational chunks uttered under one intonation contour and identified using a convergence of prosodic cues, such as anacrusis (a pattern of acceleration-deceleration), pauses preceding and following linguistic material, an overall decline in pitch level, and a terminal pitch contour at the end of each intonation unit (e.g. falling, rising) (Chafe 1993; Chafe 1994). In this transcription system, each IU is placed on a separate line. With respect to the theoretical importance of IUs, Chafe (1994) suggests that we perceive speech broken up into intonation units, and that, cognitively, each intonation unit represents the information that is active in the speaker's mind at the moment.

Efforts to relate intonation units to syntactic categories described for English reveal a high association between IU and grammatical unit, in particular, clauses. In a study based on English oral narratives, Croft (1995) finds

that 97 percent of the intonation units in his corpus are clauses or phrases, and approximately 77 percent of these IUs are clauses. Drawing from English conversational data, Chafe (1994) reports that in his sample 60 percent of the substantive intonation units—contentful IUs that convey information about ideas, events, and states—are clauses. Chafe (1993; 1994) posits three main types of intonation units: substantive (described above), regulatory (IUs that regulate the flow of information in a discourse, e.g. *well, maybe, I think*), and fragmentary (IUs that were truncated by the speaker before finishing their projected contours, e.g. *I mean she*). These distinctions in IU types, however, are not coded for this study.

Croft's more frequent association between IU and clause than Chafe's may be due to differences in their data sources. It is likely that creators of (*Pear Story*) narratives, as in Croft's data, include more linguistic depictions of objects and events (prototypical functions of clauses) than do the conversational participants studied by Chafe, whose reporting on situations is integrated in interactive expression.

2.3.2 Scope of the transcription system[2]

Aside from segmenting speech into intonation units, the Du Bois et al. (1993) system marks other prosodic and phonetic phenomena. Such elements include pauses (within and between IUs), terminal pitch direction, transitional continuity (the function of the terminal pitch, e.g. final, continuing, appeal), primary accents, and word or syllable lengthening. Vocal noises such as paralinguistic glottal stops, laughter, inhalation, and exhalation are often coded as well.

The system also marks overlapping speech and provides a series of symbols to describe shifts in voice quality, such as changes in loudness, pitch, and rhythm, crying, yawning, and breathy voice. In general, standard English spellings are used with the exception of several conventionalized forms (e.g. *gonna, uh, wanna*) (Chafe 1993), though analysts sometimes supplement standard orthography with phonetic transcription to enhance understanding of the exchange. The system also provides symbols to insert researchers' comments (e.g. nonlinguistic events and noises in the environment and indecipherable material).

2.3.3 Transcription and theory

Researchers use transcription systems that are consistent with their theoretical goals. For example, in accordance with their focus on the organization of conversational turns, Conversation Analysts rely on transcripts in which the basic unit of transcription and analysis is speaker turn (e.g. Sacks et al. 1974; Schegloff 1972). In contrast, for Edelsky's (1981) investigation of collaboratively-developed and singly-developed floors, a CA transcript which highlights sequential speaker changes is inadequate. Instead, Edelsky graphically positions utterances of the speaker holding the floor in the center of the page and situates the conversational contributions of other speakers around this main participant's. In her study of the cooperative and often simultaneous character of all women talk, Coates (1996) uses a system similar to a musical staff notation because this kind of transcript permits the utterances of several speakers to have equal perceptual status.

In the last several years, the Du Bois et al. transcription system has been used by analysts concerned with the grammatical structure of language as it is found in conversation (e.g. Ford & Thompson 1996; Ono & Thompson 1995). The appeal of this system is related to the idea that prosody and prosodic units (e.g. intonation units) participate in the formation of elements and groups of elements that we refer to as grammar. The decision to use the Du Bois et al. transcription system for this project was both a principled and expedient one. This method of transcription is being used by many researchers working in the area of conversation and grammar, and, significantly, it is the system used to transcribe conversations included in the *Santa Barbara Corpus of Spoken American English*. Though the IU is not the primary unit of analysis in the present study, its interactions with traditionally defined grammatical categories is of interest. Moreover, because this system focuses on the structural aspects of speakers' utterances, it is compatible with a study of the contribution of linguistic form to the expression of speaker point of view.

The practical and theoretical consequences of transforming recorded data to a written record are well known to analysts of natural discourse. Over 20 years ago, Elinor Ochs (1979: 44) observed, "the problems of selective observation are not eliminated with the use of recording equipment. They are simply **delayed** until the moment at which the researcher sits down to transcribe the material from the audio- or videotape [bolding in original]". Transcription of conversation is always a selective process; what ap-

pears on a transcript in layout (e.g. a down-the-page, line-by-line organization) or in detail (e.g. the linguistic or nonlinguistic features that are coded versus those that are not) biases both the analysis and the generalizations that spring from that analysis (Edelsky 1981; Ochs 1979). With respect to the present study of speaker point of view, this derivative, or selective, relationship between transcript and recorded activity parallels the relationship between language use and the larger context of human (inter)action in which language occurs. In the same way that a written transcript presents a selective view of a recorded discourse, what human beings *say* selects from and objectifies what we *know* and *do* (including of course what we say) (cf. Hanks 1995). This property of selectivity with respect to language use is the most basic form of linguistic subjectivity.

However, it is not simply the method of transcription that affects the results of a given study, but also the classifications and categories chosen to systematize the transcribed material. The next section extends the present discussion by examining the coding system used to organize and manipulate the data for the study.

2.4. Coding

2.4.1 Introduction

Coding is the process of segmenting natural and experimental events (e.g. conversation) into discrete labeled chunks for purposes of counting, analysis, or statistical manipulation. Classifying social activities into analytically useful bits without transforming the data into something dangerously disparate from the "original" events is a complex task, itself not discrete, but ongoing (Lampert & Ervin-Tripp 1993). Lampert and Ervin-Tripp (1993) provide a cogent discussion of the construction of coding systems for linguistic data. Typically, researchers construct categories (and values for those categories) based on theoretical orientations and research hypotheses, but categories may also emerge from the data themselves.[3] Lampert and Ervin-Tripp note that it is advantageous for investigators to make their coding systems selective (i.e. to construct them around their research goals) because this constitutes a more productive use of time. They also suggest, however, that a disadvantage of a selective approach is that in failing to describe the full database, one runs the

risk of overlooking important and interesting interactions that may be present in the data.

The coding system used for the present work includes elements both suggested by research hypotheses based on a pilot study (e.g. the coding of subject, main verb type, and tense) as well as those that surfaced during the coding process (e.g. the coding of connectives, more detailed coding of predicate types than originally planned). Coding values vary from category to category. They may be semantic (e.g. verb and adjective types), syntactic (e.g. subject and clause type), or functional (e.g. animacy and referentiality of the subject). The basic unit of coding for this study is the *utterance*. All in all, there are 37 categories, or fields, in the database.

The data were entered from transcripts into a spreadsheet application. Utterances were typed in the first column, and the other 36 coding categories appear in subsequent columns. That is, individual utterances appear in spreadsheet rows, and the coding values for each utterance appear in columns. Discussion of the details of the coding system by field follows.

2.4.2 Utterance

Selection of a basic unit for this study was a difficult process. Given that the transcribed data are segmented into intonation units, using IUs as basic units for counting would have been a harmonious choice. However, an important goal of this study is to look at the frequency of subjective structures in conversation by studying relationships between predicates and subjects, and many IUs simply do not contain subjects and verbs, as illustrated in (1) and (2), which are intonation units consisting of a prepositional phrase and a noun phrase, respectively.

(1) … from Indonesia. L11–29[4]

(2) this kid, R12–38

Additionally, some intonation units contain more than one subject and/or predicate, as in (3). This situation makes coding of subject and predicate per unit problematic.

(3) he said you've been= <MRC blessed with a .. great bod=y MRC>, D20–44

Because intonation units are not always coextensive with clauses, the IU was

not selected as a basic unit for this study. In order to be able to code both subject and verb for a single unit, a decision was made to use a finite clause—called an *utterance* in this work—as the basic unit of analysis (see (4) below). The relatively high correlation between intonation unit and clause in English conversation supports using clause as a structural unit (see discussion in section 2.3.1.).

(4) and I gave stickers to the kids, R19–3

For as much intuitive consensus as there seems to be among linguists regarding what a prototypical clause is in English (e.g. a constituent with a finite verb and optionally a subject), and as elevated a category as the clause is in English grammar (e.g. in both descriptive and prescriptive treatments of subordination and coordination), there does not appear to be in the literature a standard, explicit definition of this important syntactic category. Quirk et al. (1985: 47) define English clauses and their components by illustration only, and they note that the clause "particularly the independent clause, is in many ways a more clearly defined unit than the sentence". Additionally, in the initial chapters of *A Comprehensive Grammar Of The English Language*, focus is on the simple sentence (a single independent clause) "as the most central part of grammar" (Quirk et al. 1985: 47). Jespersen (1965: 103) defines a clause "as a member of a sentence which has in itself the form of a sentence (as a rule it contains a finite verb)". This, then, brings us to consider why *sentence* was not chosen as a basic unit of analysis for this study.

Both Lyons (1977) and Levinson (1983) make a distinction between *sentence* as an actual piece of language in a discourse and *sentence* as an abstract construct in theoretical models. The *sentence* also has been used to identify a unit of written language (e.g. Lyons 1977) and to describe a syntactic configuration composed of a matrix clause plus all other dependent clauses that are affiliated with it (e.g. Halliday 1994; Quirk et al. 1985; Schiffrin 1987). Therefore, *sentence* was considered too theoretically abstract, too often associated with written registers, and too structurally large, or inclusive, to serve as the base unit for this analysis.

Though the basic unit for coding in this study is a structural element whose prototype is a finite clause, recall that its label is *utterance*, not *clause*. Like *sentence*, the word *utterance* has been used in a variety of contexts. Harris (1951, cited in Lyons 1977) defines it as a speaker's stretch of talk surrounded by pauses (silence), and Levinson (1983) suggests that an *utterance* is the

actual expression of a theoretically abstract sentential unit. Additionally, *utterance* has been used as a unit by speech act theorists (Lyons 1977). The rationale for using *utterance* instead of *clause* in the present study is to highlight the central meaning of *utterance*—its designation of a chunk of language produced by real speakers—while downplaying the heterogeneous structural classifications that characterize English clauses (e.g. finite? nonfinite? subjectless? coordinate? complement? embedded?).

2.4.2.1 *Composition of an utterance*

The *utterance* category is a text field in the database used for this study.[5] Since intonation unit boundaries are not factors in delineating these units, the question of what material is considered part of the utterance requires explanation. In cases where subject and predicate information occur in different IUs, then those two IUs are combined as a single utterance as in example (5), where the commas indicate an intonation unit boundary with continuing intonation contour.[6]

(5) .. the so=ngs they were playing on the radio, were the exact same so=ngs they were singing at Bahia, D15–29

Similarly, prepositional phrases and other adverbials occurring in separate intonation units are grouped in one utterance with their accompanying subject and predicate, as in (6) and (7).

(6) he's doing it, **all over the ya=rd**.[7] G8–41

(7) **all his life**, he's been like that, D13–31

Sentential adverbs (e.g. *normally* in (8)), discourse markers (e.g. *well* in (9)) and connectives (e.g. *because* in (10)) that appear in separate intonation units preceding a clause are grouped in the same utterance with the main clause.

(8) **normally**, it used to storm along in fifth ^gea=r. F7–9

(9) **well**, I ^don't normally sound .. like ^Lucille Ball. H3–51

(10) **because**, … they'll get kind of surly, R18–43

2.4.2.2 *Uncoded clauses*

Nonfinite clauses are not coded as separate units in the database. They may, however, appear (uncoded) in utterances with finite verbs, as in (11).

(11) it's m- .. less socially acceptable here **to make, … an a rude remark about women**. C13–15

Adnominal relative clauses, or "the central type of relative clause" in English, (Quirk et al. 1985: 1245) are not coded as separate utterances. Example (12) illustrates an adnominal relative that is not coded as a separate unit in the database, but, rather, occurs as part of the predicate nominal.

(12) this is the knife, **that you use for birthday=s,** A14–13

Similarly, in (5) above, the relative clause, *they were playing on the radio*, is coded as part of the subject noun phrase, *the songs they were playing on the radio*. The decision not to code adnominal relative clauses as separate units (even when both subject and predicate information are available as in (12)) emerged during the coding process primarily for practical reasons. Due to the sheer volume of utterances produced by speakers in only 80 minutes of casual conversation, it became a practical necessity to set limits on units to be coded.

Unlike adnominal relative clauses, however, nominal relative clauses are coded as single units in the corpus. Quirk et al. (1985) describe nominal relatives as noun phrases modified by adnominal relatives, and they resemble *wh*-interrogative clauses. Significantly, "the nominal relative clause is more self-contained than the adnominal relative clause and can function as an element in a superordinate clause" (Quirk et al. 1985: 1056). Examples follow below.

(13) oh I don't know **what the radio station is.** L17–20

(14) I can see **how, .. that would vary, from region to region**, C9–49

These structures are more like independent clauses, and they typically function as complements to main clauses. This impression is supported by Quirk et al. (1985: 1059) who remark in a note, "Nominal relative clauses have also been called 'independent' or 'free' relative clauses".

2.4.2.3 *Utterances without expressed subjects*

Utterances with finite verbs without expressed subjects are coded in the database. Because frequency relationships between particular subjects and verb types are of interest, the lack of an overt subject in an utterance is potentially informative. There are three general cases in which an utterance without an expressed subject is coded as a unit in the database: (1) impera-

tives, (2) coordinate predicates (utterances with one subject but two or more predicates), and (3) utterances that consist of common expressions in which the subject is clear from either the context or the expression itself. These three situations are illustrated in (15), (16), and (17), respectively.

(15) imperative:
chill babe. A10–31

(16) coordinate predicates:
and she went out there, got on her knee=s, L23–17

(17) common expressions:
sure does. C9–52
looks like it. D13–46

2.4.2.4 *Utterances without main verbs*

There are a group of coded utterances in the corpus that do not have main verbs. These clauses consist of a subject plus an auxiliary (e.g. *is, don't*), a modal auxiliary (e.g. *can, wouldn't*), an intermediate function verb (e.g. *have to, be going to*) (see Quirk et al. 1985: 136 ff. and section 2.4.18.), or simply the adverb, *just*.

2.4.2.5 *Subject-verb combinations with discourse function*

It is well known that particular combinations of English subjects and verbs function in conversation more often as pragmatic expressions than as main clauses. For example, Schiffrin (1987) analyzes the discourse uses of *you know*, and *I mean*; Thompson and Mulac (1991) track the development of epistemic parentheticals in English from the subject-verb combinations, *I think* and *I guess*; and a range of pragmatic functions beyond a literal 'not knowing' has been attributed to the expression *I don't know* in English conversation (Scheibman 2000; Tsui 1991). Quirk et al. (1985: 1112) refer to such expressions and other similar ones as *comment clauses* defined as "…parenthetical disjuncts. … [that] may occur initially, finally, or medially, and thus generally have a separate tone unit". Thompson and Hopper (2001) call these constructions epistemic/evidential clauses. However, because the coding of subject and verb for the present study is based on formal properties, all subjects and finite verbs—including these frequent discourse markers—are coded as subject-verb combinations regardless of their functions.

2.4.3 Data source

Each utterance in the corpus is given a unique number. It consists of a letter (the abbreviation of the taped conversation, see Table 2.1.), the page number of the printed transcript where the utterance is found, and a number marking the line on the page of the transcript where the utterance begins. If more than one utterance appears on a given line, utterances are distinguished with lower case letters (e.g. a, b, c). A hyphen appears between page and line numbers. For example, an utterance from the conversation *Mable* that appears on page nine, line 35 of the transcript is assigned the identifying number M9–35.

2.4.4 Speaker

The name or initial of the speaker of the utterance as indicated on the transcript is marked in this field. In cases where a full name appears, it is typically a pseudonym. In transcription, these abbreviations or pseudonyms are also substituted for speakers' real names that appear in the conversations when speakers make reference to one another. Du Bois et al. (1993: 49) suggest that "pseudonyms should retain some flavor of the actual names".

2.4.5 Sex of speaker

The values for this category are *F* (female) and *M* (male).

2.4.6 Syntactic type

The values for this category are formally determined based on the syntactic shape of the utterance. They are: *declarative* (SV(O or C) constituent ordering), *interrogative* (the auxiliary, or a *wh*-element + auxiliary, occurs in front of the subject), *imperative* (no overt grammatical subject; also included are *let* and *let's* imperatives). Other codes used for this field are *be-passive* (*be* + past participle of the main verb), and *get-passive* (*get* + past participle of the main verb). There are also several tokens that are double-coded as interrogatives and *be*-passives and interrogatives and *get*-passives. All truncated utterances in the database contain enough information to be assigned syntactic type values.

2.4.7 Clause type

Establishing a breakdown of clause types was not a goal at the outset of planning the study. However, early in the coding process it became clear that relationships among clauses might be useful to an investigation of linguistic subjectivity. For example, given Thompson and Mulac's (1991) findings that the "main clauses" *I think* and *I guess* are grammaticizing as epistemic expressions in English, then it is of interest to investigate what other combinations of subject + verb (e.g. first person singular subjects with verbs of *perception*, *cognition*, *speaking*, *feeling* or predicate adjective constructions) regularly express speaker point of view. Also of interest in looking at clausal relations are the ways in which speakers evaluatively structure their linguistic presentations (in their choice, e.g., of connectives). Coding clause types for this study, then, was not for purposes of creating a sensitive taxonomy based on details of structures or functions, but, rather, to mark general distinctions that permit investigation of the issues mentioned above and also allow for open-ended exploration.

Clauses in the corpus are coded as one of four general types: *main, coordinate, subordinate,* and *complement*. Using traditional linguistic categories to classify clauses is in keeping with the methods used in this study which allow implicit testing of the fitness between *referentially described* grammatical categories and spontaneous conversational data.

Recent investigations of traditional classifications of clause types and clause combinations by discourse-oriented linguists reveal that "the motivations for clause creation and combination are semantically and pragmatically heterogeneous, [and] that correlations between formal indices of subordination … and pragmatic function are inconsistent" (Haiman & Thompson 1988: xii). In general, this literature suggests that theoretical dichotomies such as parataxis vs. hypotaxis or, more specifically, coordination vs. subordination, are continuous notions rather than categorical ones (see, e.g., Chafe 1988 for English conversation and Matthiessen and Thompson 1988 for written English). When applied to actual discourse, traditional descriptions and distinctions often require reformulation or even refutation (e.g. complex embedding might be a property of formal written genres.)

Each clause in the database is coded relative to its relationship to preceding and following clauses; this means that one utterance may be coded with more than one value for clause type. For example, in (18), the clause *I fell heir,*

to that painting, is both a complement (object) of the clause *I think* and a main clause to the subordinate clause *cause it's -s a huge big thing*; thus it is coded *complement/main*.

(18) I think **I fell heir, to that painting**,
cause it's –s a huge big thing, M11–45–47

2.4.7.1 *Main*

Main is the unmarked member of the *clause* category. In general, utterances that do not serve as major constituents of other clauses are tagged as main clauses; often clauses that do not contain connectives (coordinators, subordinators, or complementizers) are coded *main*. Examples of main clauses are illustrated in (19), (20), and (21).

(19) he's having a good time. M7–40

(20) **I wonder** why that would be. C8–47a

(21) if her skin weren't like .. really dark brown,
she'd probably be bright red, R9–49

2.4.7.2 *Coordinate*

In their description of coordination, Quirk et al. (1985) note that distinctions between coordinators and subordinators are not always discrete; though they consider *and, or,* and *but* coordinators in English, they suggest that *and* and *or* are more central members of this category than *but*. Following Quirk et al., many clauses in the database that are linked with *and*, and the few linked with *or*, are tagged as coordinate whether the subject is expressed in both clauses or not. For example, both of the utterances in (22) are coded coordinate.

(22) this is .. a raging bureaucracy,
and there's nothing that I can do. R13–6 – 7

There are a few pairs of coordinate clauses that illustrate asyndetic coordination (=no linking word), as in (23).

(23) he wor=ks ha=rd,
Ø he plays hard, D19–33 – 34

Also included in this group are utterances that exemplify what Quirk et al. (1985:978) call *pseudo-coordination*—a situation in which a small class of verbs (exemplified by *go* and *come in* in (24) and (25), respectively) precede *and*, and they "have an idiomatic function which is similar to the function of catenative constructions".

(24) and I **went and sprayed** it twice. F11–14

(25) and all these teachers **are coming in and saying**, R16–9

2.4.7.3 *Subordinate*

All clauses that contain a connective that is neither *and* or *or* and do not function as a major constituent of another clause are coded as *subordinate*. The most frequent subordinate clauses are *but, because/cause*, and *so* clauses. Other examples of subordinate clauses in the database are conditionals (*if*), temporal clauses (e.g. *when*), and comparatives (e.g. *like*). This group is illustrated in (26) and (27).

(26) it finished **when she got married to Todd.** A9–27

(27) we couldn't remember,
so we made something up. G5–40

2.4.7.4 *Complement*

Though *complement* is a term that is used in a variety of ways by linguists (e.g. to refer to a predicate adjective or predicate nominal), in this study the coding of a unit as a complement clause follows McCawley (1988) who defines it as a predicate that fulfills a major syntactic role in a matrix sentence (e.g. subject, direct object). Dixon (1991) notes that complement clauses often follow verbs of speaking but also verbs of other semantic types as well (e.g. wanting, thinking, liking, deciding[8]). Additionally, *wh*-clauses (structures referred to as nominal relative clauses in section 2.4.2.2. and exemplified by (13) and (14) above) function as complements to main clauses. Main clauses that have been characterized as conventionalized pragmatic expressions, such as *I mean, I guess, I don't know*—even *you know*—(Chafe 1988; Scheibman 2000; Schiffrin 1987; Thompson & Mulac 1991) frequently precede complement clauses in conversation. Example (28) illustrates a typical case of a complement clause following a verb of cognition.

(28) I think **I can afford it.** A15–9b

2.4.7.5 *Coding challenges*

The difficulty of having to stipulate one-to-one relations between unit and value when the data did not easily lend themselves to such linear organization was an issue that was prominent during coding of clause types and apparent throughout the entire coding process. For example, though the criteria laid out in this section indicate that a coordinate clause can be identified by the presence of the coordinating conjunction *and*, nonetheless, there are utterances in the corpus that begin with *and* that are coded as main clauses. The reason for this is that in listening to the tapes and reading the transcripts, it was not possible to locate a coordinate clause that was semantically, or propositionally, affiliated with the *and* clause in these cases. In English conversation speakers often begin their utterances with *and* (also with *so* or *but*) without linking these connectives to immediately preceding linguistic material, but rather to more distant topics or even to their own as yet unarticulated (and unrecoverable) perspectives. In (29), the topic of the episode in which this utterance occurs concerns one of the participant's consistently getting sick when he travels in Mexico. In this example, the speaker's *and* is making reference to the whole discourse, not to a specific preceding utterance.

(29) **and** you both eat the same things? D12–4

Another challenge in coding clause type had to do with indicating multiple complex relationships among clauses. For example, there are cases in which a series of two (or more) coordinate clauses together stand in a main clause relationship to a following subordinate or complement clause, as illustrated in (30). In this episode, speakers are discussing how the spelling of one of their family names changed during immigration of an ancestor.

(30) (a) they ... moved the E=, ... (1.3) after the O=,
(b) and took it off the en=d.
(c) when they came over here. M10–10 –16

In (30), it is not clear which of the first two coordinate utterances, (a) or (b), should be coded as the main clause affiliated with the *when* subordinate clause (c). Grammatically speaking, the coordinate clauses, (a) and (b), together make up the matrix clause, but this coding is unwieldy both for data entry and analysis. Both (a) and (c), then are double-coded *coordinate/main*. Similarly,

in situations in which speakers are using direct quotation, there are sequences of clauses that may have multiple relations with one another. Consider (31).

(31) (a) because then they're like,
(b) <Q what are we gonna learn today Q>.
(c) <Q see I'm originally a .. PE teacher Q>,
(d) <Q but I guess I can teach math Q>? R11–16 – 24

In this utterance group, (a), which is itself subordinate to a preceding (also subordinate) clause, stands in a main clause relationship to clauses (b), (c), and (d), which are all complements of the quotative verb, *be like*. Note too that utterance (d) contains two clauses: a main clause, *but I guess*, and the complement to that clause, *I can teach math*. In such cases, then, in which clauses maintain complex grammatical associations with one another, items are double-coded based on their relationships with immediately preceding and/or following utterances.

Though the focus of this study is not English clause structure, much time and attention were paid to situating clause type and utterance type. This was necessary because these categories are directly relevant to establishing a basic unit of analysis for the study, and it proved to be an interesting process leading to questions for further research on the nonlinear structure of clauses in English interactive discourse.

2.4.8 Connective

Any coordinator, subordinator, or complementizer affiliated with an utterance is entered in this field. Usually these elements appear utterance-initially, but not always (cf. Chafe 1988 who discusses the placement of *and* in intonation units). Also entered in this category are pragmatic items that tend to occur at the beginnings of utterances (e.g. *well, anyway, yeah but*). The rationale for including these discourse markers with more traditionally (syntactically) defined connectives in this category is based on research by discourse analysts who observe that, in conversation, even syntactically classifiable elements such as *and* and *so* do not typically link propositional material encoded in linguistic utterances, but rather signal interactive phenomena (e.g. Schiffrin 1987 suggests that *and* is often a marker of speaker continuation).

2.4.9 Other clause

This is a text field that contains the main, coordinate, or subordinate counterpart clause(s) of the coded item, if there is one. For example, if a given utterance is a subordinate clause, then the text of the affiliated main clause will appear in this field. Similarly, if a clause is, for example, simultaneously a coordinate clause with respect to a preceding utterance and a main clause in relation to a following complement, then this field will contain the texts of both the preceding and following clauses.

2.4.10 Polarity

The two values for this category are *affirmative* and *negative*. Utterances that are negated with *never* are additionally coded in one of the adverbial fields for that element, since *never* not only changes polarity but also functions as an adverbial of time and an intensifier.

2.4.11 Subject

The values for this category are *1s, 2s, 3s, 1p, 2p, 3p* and three coding values used to differentiate clauses without an expressed subject: *none-imperative, none-coordinate predicate,* and *none specified*. Coding is based on the form of the subject, so in cases where there is a mismatch between structure and meaning/function, it is the form of the subject that is coded. For example, English speakers often use a second person singular subject generically to refer to their own experience, as in (32). In such cases, the subject is coded *2s* with a comment about first person usage in a notes field. Additionally, generic referentiality of the subject is coded in the category, *referentiality of subject* (see section 2.4.16.)

(32) I had a terrible flu.
I keep having like,
feeling good, X-wise,
it's just,
you feel it shift in your body. G16–3–8

Second person plural subjects are *you guys*, *you both*, and one case of *you* where it was clear that the speaker was referring to a married couple. There were a few cases where it was not immediately obvious whether *they* had a

singular or plural referent, and this ambiguity was noted in a notes field.

In utterances with unexpressed subjects, the referent was typically recoverable from the discourse. However, because the values for this category are based on formal properties, these clauses are coded as having no subjects. Clauses without overt subjects are coded in three possible ways. Both second person imperatives and *let's* imperatives are coded *none-imperative*. Utterances with coordinate predicates in which the subject is not repeated with the second predicate, as in (33), are coded *none-coordinate predicate*. These constructions tend to occur in conversational narratives. During recitation of events, for example, speakers do not consistently pronounce the subject of each clause when there are several of them in a row.

(33) first she hires me n- like, .. the Friday before school starts.
and expects me to get my room ready, R15–3,5

The most general coding value assigned to unexpressed subjects is *none specified;* the majority of these items are conventionalized expressions (see section 2.4.2.3.).

2.4.12 Third person singular subject type

Utterances coded as having a 3s subject in the *subject* category are coded more specifically in this field. The values for 3s subject type are: *s/he, they, it, one, that, this there, what, who*, and *lexical noun phrase*.

2.4.13 Third person plural subject type

Utterances coded as having a 3p subject in the *subject* category are coded more specifically in this field. The values for 3p subject type are: *they, these, those,* and *lexical noun phrase*.

2.4.14 Lexical noun phrase subject

The text of third person singular and third person plural lexical noun phrase subjects are entered in this field.

2.4.15 Animacy of subject

The motivation for coding *animacy* and *referentiality* (see below) of subjects came from a pilot study conducted in preparation for this investigation. The findings from the preliminary analysis suggested that animacy and referentiality may be relevant factors in the expression of speaker evaluation, specifically in third person singular copular constructions with *it, this,* and *that* subjects (e.g. *that* more frequently than *it* or *this* has an abstract referent—an event, a piece of discourse, a generalization—than do *it* or *this* which more often tend to make reference to individuated entities).[9]

The values assigned to the category of animacy are: *human, nonhuman animate, inanimate, abstract*. These codes are not intended to exhaustively characterize animacy in English; the primary objective is to maintain a distinction between human and nonhuman subjects, with some refinement. Coding for both *human* and *nonhuman animate* was straightforward except for several utterances with *that* subjects (and a few with *it* subjects) in which it was ambiguous whether the subject was referring to a particular person or to a situation in which a person is designated (the latter case would be coded *abstract*). In (34) and (35), subjects are coded *human*, though it is not completely clear whether the *it* subjects in (35) refer to a person or to the previous piece of discourse about the person (an inquiry).

(**34**) **that** w- would be, John .. Eaton's mother-in-law, F9–43

(35) (a) **it**'s not a woman,
(b) and **it**'s not a male lesbian. G3–35 –36

Some explanation regarding the coding values *inanimate* and *abstract* is also in order. Subjects referring to things (e.g. jacket, audiotape, vitamins), body parts, places, and specific, well delineated events (e.g. test, semester, Christmas) are coded *inanimate*. The most general group, then, is *abstract*, which includes all other events and states (e.g. aging, getting a DWI, being sick), discourse events, ideas, temperaments, and so-called nonreferential subjects (e.g. *it, there*).

2.4.16 Referentiality of subject

Of all the categories in the database, referentiality of subject was perhaps the most time-consuming to code; it required going back to the audiotapes more

often than any other category. Though there are several reasons why identifying referents of noun phrases in interactive discourse is a difficult task, most challenging is making a distinction between reference to a real-world event or state and to the construction of that event or state by speakers in conversation. Making such decisions rests on a relatively modernist stance that presupposes an objective, nondiscursively bound reality—a point of view that introduces a bias into the analysis of the data. Additionally, having to hypothesize how speakers construe events based on their language use highlights some of the rich and thorny issues facing cognitive linguists, such as the characterization of the phenomena *events* and *entities*.

The values for *referentiality* were determined in a pilot study and were further delimited during the process of coding for the current investigation. The values are: *speech act participant (SAP), entity, generic, event-state, discourse, nonreferential*, and the double codes: *SAP/entity, SAP/generic*, and *event-state/discourse.* Table 2.2. presents these codes with examples.

2.4.16.1 *Speech act participant (SAP)*

Referents of subject noun phrases who are themselves participants in the conversation are coded SAP. Naturally, first and second person subjects figure prominently in this group, but there are also third person subjects (primarily, singular) coded *SAP*. It is interesting that the majority of *s/he* subjects referring to SAPs are used by couples to refer to one another.

Table 2.2. Subject referentiality codes.

Referentiality Type	Example Utterance	Data Source
SAP	<@ **you**'re quick tonight @>.	G6–23
entity	man, **he** threatened little kid=s.	R20–27
SAP/entity	**we** got back to the shop last night,	F2–47
generic	when **you** had this immediate image of <X bra-burning X>,	C12–27
SAP/generic	**you** have to be Catholic first.	L16–12
event-state	K: yeah but at that point, she was still on an n- . in= ... unending ... money ... streak.	
	M: yeah but has **that** finished?	A9–22
discourse	maybe **that** should be his last name.	G6–1
event-state/discourse	C: that's not that young in terms of forming,	C14–46
	A: **that**'s true.	C14–48
nonreferential	**there**'s no word?	H7–28

2.4.16.2 *Entity*

This value is assigned to subjects whose referents are people (but not SAPs), animals, or things. Typically these are third person subjects, but first and second person subjects whose referents are not speech act participants fall into this group. For example, in (36), *I* is part of the speaker's direct quote, thus its referent is coded as *entity*, not as *SAP*.

(36) and the woman's like <Q **I** can't do that,
there's a privacy code Q> or whatever, R13–26

2.4.16.3 *SAP/entity*

These subjects make reference to both a speech act participant and another person (or several). In this corpus, only utterances with first person plural subjects have this coding value. These are situations in which the speaker is speaking for herself and others (e.g. colleagues, friends) who are not present during the conversation. This is conventionally called the *exclusive* usage.

2.4.16.4 *Generic*

Subjects coded as *generic* are those that designate a class of referents as opposed to an individual referent. There is a range in this group with respect to specificity of the class. Referents can be quite general, as in (37); this utterance is taken from a conversation in which the participants are incredulous that people living in San Francisco are careless about protecting themselves from the HIV virus, given the availability of information. *People* refers to an undifferentiated group who have access to educational materials about HIV and, even so, the speaker claims, are having unprotected sex.

(37) **people** have the information they need. L15–51

Similarly, in (38), *the ceviche* refers to ceviche as a type of food (served in Mexico, in this conversation), not as a particular serving of the food.

(38) **the ceviche**'s full of vegetables=, D12–25

Many utterances coded *generic* have a *they* subject (e.g. institutional *they*) that refers to groups or classes of people and things at varying levels of specificity. In (39), *they* refers to commercial birthday cards in general. In (40), *they* designates an entire nationality, or region. In (41), *they* is more

specifically delineated in that it makes reference to the speaker's own elementary school class.

(39) **they**'re always dorky, A11–10

(40) I mean the French are basically a northern people,
and **they**'re, .. **they**'re formal, C9–5b

(41) and **they** get up out of the chair, R17–17

Also coded *generic* are second person singular subjects that could be glossed as 'one', as in (42). As noted above, often these generic *you* subjects have a first person reading.

(42) but **you** still see the marks in the la=nd. C10–40

You subjects whose reference is ambiguous between a generic reading and a participant in the conversation are coded *SAP/generic* (below).

2.4.16.5 *SAP/generic*

This coding value groups referents that may be viewed as both speech act participants and as generic referents. The subjects in this group are primarily second person singular, typically appearing in conventionalized expressions such as *you know* or *you know what* in which *you* maintains both a general reading as well as one that, in actual usage, makes reference to the addressee(s). Other subjects that fit the criteria for the *SAP/generic* code are first person plural subjects where *we* is construed as simultaneously referring to the conversational participants and also to a larger, more generic *we*. In (43), notice *we* appears in the rhetorical *we're talking*.

(43) **we**'re talking=, .. Cookie Monster. A12–32

2.4.16.6 *Event-state*

What may be characterized as referring to a linguistic construal of events and states rests on definitions of the terms *event* and *state* themselves. For purposes of coding referents of subject noun phrases for this value, events and states referred to in discourse may be as specific as an act of spraying growing peas with a chemical or not shaving one's legs to more general opinions about, and evaluations of, both real and hypothetical situations and experiences (e.g.

what it means to have a peculiar reaction to the odor of some vitamins or disagreeing with a particular point of view). The most straightforward references to events and states to code are those that are easiest to construe as singular activities, such as exams, or dance classes, as illustrated with *that* in (44) and *this* in (45).

(44) for most people, .. Christmas is a ti=me, when .. you — you put up
o=rnaments, and you celebrate,
and you drink,
and smo=ke,
and, have friends over,
that's what Christmas is for. H5–46 – H6–2

(45) **this** is a beginning lambada class, L20–47

There are also many utterances in this group where the subject noun phrase—usually a third person pronoun with a nonhuman referent—refers to more abstract phenomena such as ideas and opinions about events and states. In (46), *it* refers to the speaker's conception of French people's perception of German people and English people, and in (47), *that* refers to the previous speaker's delight that she gets a break from her studies during Thanksgiving holiday.

(46) **it**'s curious. C7–28

(47) **that**'s nice. G14–38

2.4.16.7 *Discourse*

Discourse referential subject noun phrases are designated as those that explicitly refer to linguistic elements (e.g. names, a word, a clue) or to bigger stretches of the discourse (e.g. what a speaker says)—mainly, metalinguistic phenomena. In (48), *it* refers to a conversational item that was on a list to be discussed during dinner, and *that* in (49) makes reference to a pronunciation of a chemical name.

(48) and **it** was supposed to be Z's work. G17–11

(49) **that** s- sounds like cow's balls? D17–7

2.4.16.8 *Event-state/discourse*

Note that even in comparatively straightforward examples of references to events and states such as (44) that *that* not only refers to prototypical Christmas activities but also to the speaker's discursive elaboration of those activities. Similarly, *this* in (45) not only refers to the lambada class but also to the situating of the discussion of the class within the discourse. In many utterances coded as *event-state*, there is a potential reading that what is being referred to is not simply a situation in the external and internal worlds, but also to the textual constructions of these situations.

At the outset of coding for referentiality it became clear that it was not a simple matter to distinguish between an *event-state* and an *event-state* as it is represented in discourse. Aside from the important considerations concerning what elements and processes contribute to the construal of what we call events and states in general, how does one determine which aspects of a given situation should be linguistically encoded in an utterance in order to designate a noun phrase as referring to an *event-state* as opposed to its referring to a piece of *discourse*? That is, how in principle does one separate linguistic and nonlinguistic input (e.g. frames, scripts, inferences) that contribute to the understanding of a variety of situations? This inseparability of real worlds and linguistic representations of those worlds led to a decision in this project to code as *event-state* any situation that could be construed as having some coherent organization (see (44) and (45) above). It also contributed to the decision to restrict coding of utterances as *discourse* referential to very specific cases in which it was the actual language that was referred to. Even given these conservative guidelines, there remained several utterances whose subject noun phrase referents straddled the line between being *event-state* and *discourse* referential, and these are double-coded as *event-state/discourse*.

The most general characteristic that may be said to be shared by the majority of these double-coded items is their expression of strong commitment to the opinion of a previous speaker. For example, in utterances such as *that's true, that's right, that's for sure* there exists an ambiguity in the referentiality of *that* between referring to the propositional *content* of the preceding contribution and to the *expression* of that information by the preceding participant. As illustrated in (50), *that* in *that's true* simultaneously refers to the content of the speaker's opinion and serves as a positive reaction to his assessment.

(50) Miles: and the- — only way to learn how is to dance with them mo=re.
Jamie: **that**'s true. L19–11–14

2.4.16.9 *Nonreferential*
In many grammatical descriptions, some usages of the subjects *it* and *there* are characterized as empty, or nonreferential (e.g. Jespersen 1965; Quirk et al. 1985). However, in coding utterances with *it* subjects, for example, it became apparent that these forms are not as empty as has been claimed. In (51), the referent of *it* may be characterized as something like 'the time, which in this case is now'. The referentiality of this subject is coded *event-state*, not *nonreferential*.

(51) **it**'s not the end of Chanukah. H7–10

A particularly bleached sense of *it* is one that refers to the discursive situation in focus and these uses are also coded *event-state*. For example, in (52) participants are talking about the effects of World War II on French attitudes, and *it* refers to the speaker's assessment of the existence of rifts between the French and German people.

(52) **it**'s not a gorge that <@ goes century deep @>. C10–43

There are, however, several utterances with *it* subjects for which it is more difficult to locate even these well grammaticized readings of ambient *it* (Bolinger 1977). These subjects are coded *nonreferential*, though admittedly, even these noun phrases are not completely without meaning. Examples follow in (53) and (54).

(53) if **it** wasn't for my back. L21–19

(54) no no, **it**'s just me. G8–20

The bulk of the utterances coded *nonreferential* in the corpus, however, are those with *there* subjects. Bolinger (1977:90) notes that in *there* expressions (e.g. *There is money in the drawer*), *there* acts as a subject pronoun, and "[t]he real subject becomes the complement". Both Bolinger (1977) and Lakoff (1987) discuss the influence of a locative or deictic meaning in these existential constructions, and Bolinger suggests that the current general meaning of *there* in these utterances is something like 'awareness'. As noted above for *it*

utterances that are coded *nonreferential*, *there* clauses coded *nonreferential* also retain some meaning, even if it is not easily characterized.

2.4.17 Central modal

The English central modals are entered in this field (*can, can't, could, couldn't, might, might not, must, shall, should, shouldn't, will/'ll, won't, would'/d, wouldn't*).

2.4.18 Intermediate function verb

With respect to their semantic and morphosyntactic properties, there are a variety of English verbs that fall somewhere between the broad categories of auxiliary and main verb. Applying grammatical and semantic criteria used to identify (modal) auxiliary verbs in English (e.g. inversion, abnormal time reference, negative contraction), Quirk et al. (1985) present an auxiliary–main verb scale. Central modal auxiliaries are at one endpoint of the continuum (as types of predicates with one verb phrase; e.g. *I can go*) and main verb plus nonfinite clause constructions are at the other end (as types of predicates with two verb phrases; e.g. *He hopes to do it*). Between these points, Quirk et al. situate several sets of verbs, together called *verbs of intermediate function*. These include marginal modals (e.g. *need, ought to*), modal idioms (e.g. *would rather, gotta*), semi-auxiliaries (e.g. *be going to, be supposed to*), and catenatives (e.g. *keep+-ing* participle, *start+-ing* participle) (Quirk et al. 1985: 137 ff.).

In the present study, Quirk et al.'s identification of verbs of intermediate function is used as a guide in coding certain English verbs as *intermediate function verbs* (IFVs), though no attempt is made to classify these elements more specifically, e.g. as marginal modals, semi-auxiliaries, etc. The goal here is to separate true modal auxiliaries from other modal/auxiliary-like expressions from main verbs.

A list of the IFVs in the database may be found in Appendix B. Included in this group are Quirk et al.'s marginal modals, modal idioms, and semi-auxiliaries. From the catenative group (chainlike structures with several nonfinite form) are verbs that precede an *–ing* participle, such as *begin, continue (on), start,* and *stop*. Other items coded as IFVs are verbs that behave as "pragmatic particles" (Quirk et al. 1985: 147–148) such as *want to, let,* and *let's*. There are

also a few other verbs identified as intermediate function verbs that are not specifically mentioned in Quirk et al; these items—*afford to, have* (NP), *like to,* and *make* (NP)—express modal meaning and typically occur in catenative expressions (e.g. *I can't afford <@ to take @>.*).

2.4.19 Predicate type

Predicates are largely represented in the database by listing the major elements that comprise them. These elements include: v*erb, auxiliary (aux), modal, intermediate function verb (IFV), copula, nonfinite verb (NF)*, and the copular complements *NP, ADJP, PP*. The major predicate types in the corpus are listed in Table 2.3.. These types, however, do not represent the complete inventory of coding values for predicate type. Other less frequent types are derived from the basic types listed in this table. For example, a copula + NP construction may also contain a modal or an auxiliary as in the utterance, *you guys could be partners*., and this is coded modal + copula + NP. Similarly, an utterance may contain two intermediate function verbs (IFVs) as in *my sister* ***used to*** *always* ***try and*** *give me papaya enzy=mes*, coded IFV + IFV + NF. In other words, the coding of more complex predicate types is based on the coding of the major types shown in Table 2.3.

Note that not all predicate types are represented by a literal ordering of constituents found in the utterance; rather, some predicates are coded more schematically in the sense that not every element is represented. For example, there are no slots in the predicate type templates for adverbs; these items are coded in separate fields in the database. Additionally, copula plus complement types are represented as copula + NP (ADJP, PP), not as aux + NP (ADJP, PP). There is in principle no compelling motivation to prefer the more functional label (copula) over the formal designation (aux); the decision comes from interest in comparing subjects appearing with copular constructions with those that more frequently occur with lexical predicates.

Anisomorphy between the formal composition of predicates and the designation of predicate types is also found in the coding of several intermediate function verbs (IFVs). Notice in Table 2.3. that an example of the predicate type IFV + NF is found in the utterance *so she's just gonna watch.* The IFV in this clause is the entire construction *be gonna*; that is, the IFVs *be gonna* and *be going to* (also *be supposed to* and *be able to*) do not have their auxiliaries coded as separate elements; instead, the entire expression is coded as an IFV. There

Table 2.3. Major predicate types.

Predicate Type	Example Utterances	Data Source
verb	I **bet**.	L19–10
	he **lives** in Pullman,	F3–38
aux + NF	**I'm telling** you.	F8–39
	I **don't think** so.	A18–4
modal + NF	you**'d see** like, flat line.	G6–38
	you **can use** this for your muffins.	A13–20
copula + NP	and that **was one thing I noticed**,	C8–8
	this **is a beginning lambada class**,	L20–47
copula + ADJP	it's **really horrible**.	D13–33
	Vienna**'s awfully @conservative**.	C14–33
copula + PP	it's **from her mothers lover**.	G3–29
	peas **are in em** so,	F5–12
IFV + NF	I **tried to tell** Myra,	A8–20
	so she**'s** just **gonna watch**.	L2–28
aux + IFV + NF	you **don't have to spend** money on goodies.	R19–15
	that**'s gotta ... hurt**.	F4–22
modal + IFV + NF	you**'ll keep pulling** it,	G10–26
	I **might be able to** find **ou=t** for you.	M11–35
aux/modal only	but they **don't**.	C7–27
	yeah, maybe you **shouldn't**.	L25–12
no verb	because they just,	R11–2
	what about your work, Z?	G17–22

are also cases in which a particular element of a predicate might logically be coded in more than one way. For example, the utterance *I was just trying to think,* is coded IFV + NF; *try to* in the progressive is the intermediate function verb, and *think* is the nonfinite (main) verb. While this coding does not mark the presence of the auxiliary *be* in the progressive construction, it provides a solution to the problem of there being in fact two nonfinite forms in this utterance: *trying* and *to think*.

All IFVs, then, regardless of their aspectual marking or composition, are coded as single units. Predicates with *be gonna/going to* and similar expressions are not coded as containing auxiliaries, even though progressive predicates that occur without IFVs are coded as aux + NF (e.g. *he was just talking.*). Similarly, anterior progressive predicates (e.g. *w- in fact we've been using them for several years.*) are coded aux (*'ve*) + aux (*been*) + NF (*using*); *been* is coded as aux even though it is also a nonfinite form (past participle). For the sake of consistency and to keep the number of predicate types low enough to allow this category to be sensitive to general distributional patterns, decisions were

made in ambiguous cases so that, minimally, coding would be constant. In general, such decisions entailed choosing to code a bigger, or more functional, category schematically, such as a construction, over the delineation of individual elements, when both possibilities existed.

2.4.20 MAVE/Unit verb

Though treated as central in many analyses of English structure, Hopper (1991; 1997) suggests that in fact simple verbs (single lexical items) are relatively rare in discourse. Much more common, especially in high involvement discourse, are dispersed predicates, or *MAVES* (Multiply Articulated Verbal Expressions). In his analysis of vernacular written narrative, Hopper (1991) illustrates different types of dispersed predicates in English, such as phrasal verbs; predicates composed of auxiliaries, modals, IFVs, and adverbial phrases; anacrustic coordinated clauses (e.g. ***wake up and** remember*; cf. Quirk et al.'s *pseudo-coordination* discussed in section 2.4.7.2.); collocations; and verbal groups containing a basic English verb (e.g. *make, do, give, get, take, put, have*) plus a nominal (e.g. *get a **glimpse** of, take a **look***). Quirk et al. (1985: 750) describe this last type of predicate as verb plus *eventive object*, and they note that "this EVENTIVE object is semantically an extension of the verb and bears the major part of the meaning" (e.g. *They **are having an argument*** vs. *They **are arguing***).

There are three coding values for this category: *MAVE*, *unit*, and *N/A* (not applicable); *N/A* is assigned to predicates that have no verb at all or those consisting of an auxiliary or modal alone. Note, too, that some of the information in this category duplicates what is found in the *predicate type* field.

Utterances in the corpus assigned the code *MAVE* are those with multipart predicates: phrasal verbs (e.g. *so we **made** something **up**.* G5–40), verbal groups containing auxiliaries, modals, or intermediate function verbs plus main verbs (e.g. *and **I'm thinking**,*), verb plus complement/object combinations (e.g. *on Friday I **have eighteen kids**,*; *that's **his attitude**,*), fixed expressions (e.g. *Peter .. **went .. all the way***; ***open your eyes***), predicates with adverbs (e.g. *and she **just** giggles like a goofball.*), and other linking verb constructions (e.g. *it **looks like the host***; *it **seems to me**,*).

Relative to *MAVEs, unit* verbs in the database are those predicates that are expressed as a single lexical verb (e.g. *cats **love** those*=; *you **know**,*; *and she* **said**,). This distinction between *MAVE* and unit verb is continuous and rela-

tive rather than categorical and absolute and does not reflect a simple difference between a verb pronounced as one word and a verb pronounced as more than one word. For example, many predicates coded as *MAVEs* are in fact single word verbs such as those found in copular constructions (e.g. *be*) and other verbal groups composed of a basic verb (e.g. *do, have*) plus a complement or object, but these verbs do not contribute much semantic information to the predicate as a whole; in other words, they are far from the prototype of what grammarians take to be lexical verbs.[10] In some cases it is not easy to decide whether the predicate should include the verb's object or complement. For example, in (55) the verb *say* is coded as a *unit* verb, but the decision is less clear in (56) in which the complement of the verb appears to be much more part of the predicate than the direct quote complement in (55). Notice that (55) is composed of two intonation units, whereas (56) consists of one IU. The decisions was made to code (56) as a *MAVE*.

(55) and I **said**,
<Q well, ... I don't kno=w Q>. M6–47

(56) if you **say something like tha=t,** C13–17

2.4.21 Nonfinite verb

Infinitives and participles following auxiliaries, central modals, or intermediate function verbs are entered in this field in whatever form they occur in the utterance. Infinitives are not entered with a *to* marker because of difficulties assigning the *to* to the nonfinite verb versus the IFV, since there is phonological evidence that the infinitive marker often adheres to the preceding verb and not the nonfinite element (e.g. *gonna, wanna*). As discussed in section 2.4.19. not all nonfinite verbs are entered in this field; the nonfinite forms that do not appear in this field are those that are part of IFVs or the verb *be* when it is both nonfinite and functioning as an auxiliary (*been*).

2.4.22 Main verb

The main verb of the utterance is entered in this text field. In a single verb predicate, it is the only verb in the clause. In MAVES, or predicates with auxiliaries, modals, and IFVs, the main verb is the nonfinite form that is last in the chain. In (57), the main verb is *worry*.

(57) I won't have to **worry** about it, D13–22

No verb is entered in this category (i.e. *none specified*) for predicates that consist of various combinations of auxiliaries, modals, and IFVs without a nonfinite form (e.g. *of course, you'd have to.*) or when there is no verb at all (e.g. *you pregnant?*).

2.4.23 Main verb type

In order to assess the subjectivity of subject-verb combinations in English conversation, it was necessary to classify main verbs by semantic type because it is not possible, for example, to discuss the frequent cooccurrence of mental verbs (verbs of cognition) with first person singular subjects without delineating these semantic classes in advance. As noted previously for clause type, the purpose in coding verb type was not to create an exhaustive taxonomy of English verbs; such a task alone could fill the pages of several volumes. Rather, the intent was to establish a usable number of semantic groupings of verbs based on extant systems in the literature to be able to discern frequent patterns of subject-verb cooccurrence in conversation.

Studies of English verbs and verb classes have been based on semantic and structural properties, such as the lexical and morphological expression of temporal processes (e.g. Langacker 1987; Quirk et al. 1985; Vendler 1967) and on the semantic roles of verb arguments or voice alternations (Dixon 1991; Halliday 1994; Levin 1993). The construction of a coding system for this category is based on Halliday's taxonomy of verbal processes in English which models three general processes of human experience: *being*, *sensing*, and *doing* (Halliday 1994). Table 2.4. provides a summary of the 11 coding values for main verb type. The values *cognition, existential, feeling, material, perception, relational,* and *verbal* are taken directly from Halliday; the *corporeal* class is from Dixon (1991), and the *possessive* value was adopted during the coding process. In keeping with the goal of tracking the subjective expression in conversation, this system allows for more sensitive distinctions among verbs having to do with speakers' states and processes (e.g. *feeling*, *thinking*, *speaking*, *perceiving*, *corporeal*) than for more general external actions which are grouped under one class, *material*, subsuming a diverse collection of verbs expressing both abstract and concrete activities. The rest of this section will discuss and illustrate some of the challenges that surfaced during coding of this category.

Table 2.4. Main verb type codes.

Verb Type	Description	Examples
cognition	cognitive activity	*know, think, remember, figure out*
corporeal	bodily gestures, bodily interaction	*eat, drink, sleep, live, smoke*
existential	exist, happen	*be, have, sit, stay, happen*
feeling	emotion, wanting	*like, want, feel, need, bother, enjoy*
material	concrete and abstract doings and happenings	*do, go, take, teach, work, use, play, come*
perception	perception, attention	*look, see, hear, find, notice*
perception/relational	perception (subject not senser)	*look, smell, sound*
possessive/relational	possession (x has a)	*have, get*
relational	processes of being (x is a, x is AT a)	*be, get, be like, become*
verbal	saying, symbolic exchange of meaning	*say, talk, mean, tell, ask, go* (quotative), *be like* (quotative)
N/A (not applicable)	no main verb in utterance	

Many of the verb type values presented in Table 2.4. are unambiguous; however several of them might benefit from explanation. Processes of *cognition* and *feeling* and *corporeal* and *verbal* processes are relatively unproblematic codes except to note that *verbal* includes symbolic types of communication and representation (e.g. *mean, propose, remind*) as well as more prototypical verbs of speaking. The difference between *perception* and *perception/relational* rests on the presence or absence of an experiencer; both Halliday and Dixon make this distinction. In Halliday's scheme, mental verbs of all types require a sentient agent; therefore, when verbs such as *see, hear,* and *taste* occur with a third person inanimate subject, they are considered *relational* in that a relation is set up between two entities. In the present study, however, this last group is coded *perception/relational* to mark the speaker's sensory experience implicit in these subject predicate combinations, as in (58).

(58) that **smells** like ga=rbage, D18–25

Similarly, *relational* predicates expressing possession are coded *possessive/relational*.

Existential processes indicate that something exists or happens, illustrated in (59) with the presence of *there* and the verb *to be*.

(59) anyway, **there was** this guy in his fifties, L23–10

Also coded as *existential* are verbs of rest (e.g. *sit, stay* (Dixon 1991))[11] and existentials with *have* (e.g. *cause they **have** like buses to catch and stuff,*).

The largest and least differentiated verb type value designates *material* processes, which "express the notion that some entity 'does' something" (Halliday 1994: 110). Depending on the perspective of the discourse entity, the process might be one of 'happening' or 'doing'. The verbs in the corpus coded as *material* may be as concrete as *break* or as abstract as *compromise*.

In several cases, a given lexical verb might be logically placed in more than one category. For example, the verb *think* could appear in both the *cognition* and *corporeal* classes (see Dixon 1991); *think*, however, is coded as a verb of *cognition* in the database. Similarly, Dixon labels *smell* as *corporeal*; however, it is tagged *perception/relational* in the corpus. All predicate adjective and predicate nominal constructions are coded as *relational* even though utterances like (60) and (61), whose complements contribute to the expression of emotion, might also have been coded as a predicates of *feeling*. A decision was made, however, to code the verb type, not the semantics of the entire predicate; therefore, utterances (60) and (61) were coded *relational*.

(60) and she's totally ha=ppy now. D23–20

(61) well I'm glad (they mowed,) F5–44a

Though one of the goals of coding main verb type was to hold lexical items steady while looking at their variable uses in discourse, it was simply not possible to consistently assign a single lexical item the same code. Polysemy—whether due to variation in linguistic contexts, pragmatic contexts, or shifts in semantic roles of arguments in the clause—occasionally required coding the same phonological form as more than one verb type. For example, as described above, the presence or absence of an experiencer subject determined whether the verb *feel* is coded as a verb of *feeling* or *perception/relational*, respectively. Similarly, the verb *to be* may be coded as *relational* or *existential*. The verb *be like* may function descriptively (coded *relational*) as in (62), or as a quotative (coded *verbal*) as in (63).

(62) MILES: but their hips **are like**,
JAMIE: are they going --
are they going like, really fa=st? L19–43

(63) I'd **be like**, ... yes you can, R9–27

Other examples include: *have* which may be coded as *existential* or *possessive/relational*; *tell* which primarily represents a *verbal* process but is coded as a verb of *perception* when it follows the modals *can* and *can't*; and *call* is a *material* process in (64) but a *verbal* one in (65).

(64) and he **called** looking for Jeannie and, F3–44

(65) what's the other one **called**. H7–21

2.4.24 Tense

The morphological marking of tense on the finite verb in the utterance (not the more complex semantics of time reference) is coded in this field. The values are *present, past, modal* (for clauses with central modals where formal marking of time reference is typically abnormal), and *N/A* (*not applicable*). *N/A* is assigned to utterances without verbs.

2.4.25 Aspect

Formal marking of aspect on the verb phrase is entered in this field. The values are *simple, progressive, anterior, anterior progressive, habitual* (past), *modal,* and *N/A* (*not applicable*). Parallel to the tense category, *modal* as an aspectual code marks clauses with central modals. The value *N/A* in this field codes utterances without verbs and also those for which there is insufficient information; the latter cases are truncated clauses with the auxiliary *be* (e.g. *but that's* –). The codes *simple, progressive,* and *anterior* (and *anterior progressive*) represent the major aspectual oppositions found in English (Comrie 1976). *Simple* is the unmarked value with respect to *progressive* and *anterior*. Progressives are typically formed with the auxiliary *be*, but there are instances of other auxiliary-like verbs in this construction (e.g. *keep, start*). Utterances with the IFVs *be gonna* and *be going to* are coded as *progressive (gonna)*, a subset of *progressive*. The only utterances coded as *habitual* are past tense predicates containing the IFV, *used to*. Aspectual meanings unaccompanied by a change in form are not coded (e.g. the expression of the habitual present by the simple present).

2.4.26 Predicate nominal

Predicate nominals (illustrated in (66)) appearing in noun phrase complements of copular clauses in the corpus are entered in this text field.

(66) it is a different sensation entirely. L19–6

2.4.27 Noun type in predicate nominal

In predicate nominal expressions, the noun that appears in the complement contributes to the meaning of the entire predicate. For this reason, it is of interest to classify these nouns by semantic type in order to evaluate the expressivity of utterances that contain these copular constructions. This is analogous to the task of establishing semantic classes for main verb types.

Unlike verb classification schemes which are not infrequent in linguistic investigations, there are far fewer taxonomies of semantic types of nouns in the literature. However, Dixon (1991) presents a sketch of semantic types associated with the category noun in English, a subset of which serves as a basis for coding noun type in the present study. Table 2.5. summarizes the values used to code this category.

The *clause/event* subcategory largely consists of nominal relative clauses. Recall from section 2.4.2. that nominal relative clauses are coded as separate

Table 2.5. Noun type codes.

Noun Type	Examples
abstract	*connection, alternative, way, problem*
activity	*fight, work, situation, celebration*
animal	*turtle, chicken*
artifact	*motor, diaper, bag, book*
body part	*ass, liver, hair, skin*
clause/event	*what's so peculiar, what guys have been telling me*
empty noun	*thing, nothing, something*
human	*teenager, mother-in-law, writer*
language	*word, name, clue*
natural world	*acres, iron* (mineral), *weather*
proform	*it, me, one, this, that*
quantity	*part, unit, piece, age*
states and properties	*sensation, attitude, fun, inflammation*
time/place	*November, midterm, town, time*
none (truncated)	*the very first —, but there certainly is a .. strong —*

utterances in the database (see (13) and (14)). However, for cases in which the nominal relative follows the verb *to be*, the predicate of the entire construction is coded as copula + NP with the nominal relative serving as the complement NP. Examples follow in (67) and (68).

(67) that's **what's so peculiar**. C8–48
(68) this is **what guy=s have been telling me,** L15–40

2.4.28 Adjective in predicate nominal

In this text field is entered an adjective if one appears in a copular construction with a noun phrase complement. For example, in (66) above, the lexical item *different* is written in this column.

2.4.29 Predicate adjective

For copular constructions in the corpus with adjective phrase complements, the predicate adjective is entered in this text field, as illustrated in (69).

(69) and his face was **all dirty**, F6–28

2.4.30 Adjective type in predicate nominal and predicate adjective

There are two adjective type fields in the database: one codes adjectives that appear in predicate nominal constructions (e.g. *it's just* **dry** *skin.*), and another codes adjectives in predicate adjective expressions (e.g. *you'll just be* **disgusted,**). With a few modifications, the semantic classes used to code these adjectives are a subset of those proposed by Dixon (1977; 1991) to characterize adjectives in English. Table 2.6. presents the semantic types used to code adjectives occurring in predicates in the corpus.

In a discussion of lexical categories, Dixon (1991) observes that semantic types are not mutually exclusive. For example, should the adjectives *strong* and *dirty* be coded as *physical property* or *value* types? In the discussion of verb type above, we also saw overlap in assigning lexical items to semantic classes. A challenge in stipulating semantic types for adjectives was having to choose between *value* and another, more propositionally defined, code, such as *physical property, human propensity,* or *difficulty*. This is because labeling something as *healthy, dorky,* or *hard* is not only descriptive but also encodes a

Table 2.6. Adjective type codes.

Adjective Type	Examples
age	*old, young*
difficulty (coded for pred adj only)	*hard, difficult, easier*
dimension (includes numbers, speed)	*big, small, full, far, long, one, seven, quick*
function (includes nouns acting as modifiers)	*smoking, GI, test, injector, privacy, medical, national, microwavable, beta, annual, PE*
human propensity	*happy, famous, miserable, interested, dorky*
physical property (includes colors)	*dirty, dry, strong, plastic, healthy, pregnant, blue, sick*
qualification	*probable, true, sure, exact, basic*
similarity	*like, different, same*
value	*good, overwhelming, horrible, great*
other	*own, local, other, temporary, any, recent, complex, done*
none (coded for pred adj only)	*it's pretty u- --, but you can still be ... quite --*

value judgment (cf. Verhagen 1995).[12] Therefore, whenever possible, coding of adjectives for this study follows the details of Dixon's classification of adjectives in English.

The adjective types *function* and *other* supplement Dixon's scheme. *Function* identifies a group of words in the corpus that may be said to designate the kind of activity or action appropriate to an entity (e.g. person, thing, institution). For example, in (70), *smoking* characterizes the activity that is appropriate in a given seating area. Similarly, *microwavable* in (71), describes the function of a particular food container. Half of the words coded *function* are, in fact, nouns functioning as modifiers (e.g. ***injector*** *pumps,* ***privacy*** *code*).

(70) that's the **smoking** section. G4–32

(71) that is **microwavable**. A15–12

2.4.31 Adverbial

Lexical and phrasal expressions traditionally labeled adverbials are entered in this field. This functionally diverse group consists of lexical and phrasal adverbs and discourse markers (e.g. *just, really, tomorrow, maybe, negatively, all his life, like that*) and prepositional phrases (e.g. *at one point, in second gear, out here, for birthdays*). Quirk et al. (1985) exclude as adverbials those items

that are grammatically part of another constituent. For example, *really* in *not only that, but they have* **really** *lewd billboards all over the Castro district* is part of the noun phrase and not considered an adverbial element. However, in the present study, all adverbials and prepositional phrases are entered in this field regardless of the degree of their grammatical integration.

2.4.32 Adverbial type

Identification and classification of adverbs and adverbials in English is a complex undertaking. Quirk et al. (1985) categorize these elements into four main types based on their relations in the sentence: *adjuncts* and *subjuncts* (e.g. obligatory or optional expressions of time or place, manner or degree) are relatively more integrated in the clause structure of the utterance, and *disjuncts and conjuncts* (e.g. items with sentential scope or textual functions) are more peripheral in the sentence. Because the goal of this study is to look at the contribution of these forms to speaker stance, a taxonomy based on the grammatical functions of adverbials is not suitable. Instead, with one exception (*other PP*), all adverbs and adverbials (lexical items, phrases) in the corpus are categorized by semantic, or functional, type. The eight codes used for this field are *exclamation, intensifier–amplifier, intensifier-downtoner, manner, modality, space, time,* and *other PP* (prepositional phrase). These values are outlined in Table 2.7.

Exclamations and *space* and *time* expressions were relatively straightforward types to identify in the corpus. The delineation of *manner* as means, quality, and comparison is based on descriptions in Halliday (1994) and Quirk (1985). The rationale for creating the miscellaneous code, *other PP*, was simply to keep track of a structural group (prepositional phrases) that subsumes expressions such as locatives, comitatives, and temporal phrases. These oblique complements are important to the analysis of relational clauses in chapter 4.

Once coding values for this category were established, the most challenging aspect of coding adverbial type was distinguishing between intensifiers and modality expressions. An intensifier is an element "that scales a quality, whether up or down or somewhere between the two" (Bolinger 1972: 17). Expressions involved in upward scaling are coded *intensifier-amplifier*, illustrated by *so* in (72). Those that are concerned with scaling downward are coded *intensifier-downtoner*, exemplified by *almost* in (73).

Table 2.7. Adverbial type codes.

Adverbial Type	Description	Examples
exclamation	miscellaneous exclamations and expletives	*oh my god, shit, damn, man, no wonder*
intensifier – amplifier	scale upwards from an assumed norm	*very, right, awful, so, entirely, more, extremely*
intensifier – downtowner	scale downwards from an assumed norm	*pretty, kind of, rather, about, fairly*
manner	means, quality, comparison	*like this, socially, like lead, harder, without slurring, differently*
modality	modification of the force or truth value of an utterance using emphasis, focus, or approximation	*just, probably, really, necessarily, apparently, definitely*
space	place, position, direction	*home, here, there, in a rental car, up your nostrils, away*
time	temporal relations (fixed position in time, duration)	*now, this year, already, never, still, anymore, again, on Thursday, at that point*
other PP	miscellaneous prepositional phrases incompatible with any of the other codable functions	*from Bill, about it, with her woman friend, of Jews, for children*

(72) he's asking **so** nicely. C7–12

(73) so it's **almost** impervious to destruction. A19–6

Modality, as described in relation to adverbial expressions, includes the enhancement or diminishing of the truth value or force of an utterance by emphasizing either the positive or negative pole of the statement (e.g. *one's **certainly** more sensitive to them.*) or approximating, or hedging, speaker commitment (e.g. *you **probably** bought that crop of peas,*) (Quirk et al. 1985:485). Also included as *modality* are Quirk et al.'s *focusing subjuncts*—expressions that stipulate what parts of an utterance to pay attention to by restricting or specifying focus (e.g. *so it was **just** whatever station it was on.*) or that indicate that the focus of the utterance is true with respect to added information (e.g. ***even** in Berkeley most women probably don't,*; *they have a form in Spanish **al=so.***).

As is true for the majority of categories in this study, codes for adverbial type are not mutually exclusive. For example, it is difficult to say whether the lexical item *just* is best described as an *intensifier-downtoner*, an emphasizer

(a *modality* function), or simply an expression of time. Similarly, should *really* be coded as a marker of modality or as an intensifier? Bolinger (1972: 20) writes that "[i]ntensification is the linguistic expression of *exaggeration* and *depreciation* [italics added]". But notice that the related notions—*emphasis* and *restriction*—are considered modality functions. In a discussion of emphasizing modifiers of adjectives (e.g. *Mister Samuel must have had a* ***really*** *bad day.*), Quirk et al. (1985: 447) suggest that, in contrast to intensifiers, emphasizers "add to the force (as distinct from the degree) of the adjective". As promising as this succinct statement is, the semantic distinction between the "force" and the "degree" of what is being modified becomes murky when used to classify adverbs in conversation such as *really, just, only,* or *so* as either intensifiers or modality expressions. Bolinger (1972) suggests that, diachronically, truth-identifying adverbs (e.g. *really, truly, honestly, entirely*) are a rich source of intensifiers. Given the natural subjectivity of expression in English conversation, it is not surprising that there is polysemy between *identifying* (e.g. emphasizing, particularizing) aspects of the *truth* and *assessing* these aspects using degree words. For purposes of coding adverbial type, then, all lexical intensifiers and modality items in the corpus were checked against classifications in Bolinger (1972) and Quirk et al. (1985).

2.4.33 Other first person singular expressions

Entered in this text field are first person singular expressions in the utterance aside from the subject (*I*), such as object pronouns (e.g. *tell* ***me****, bothered* ***me****, with* ***me***), possessives (e.g. ***my*** *test,* ***mine***), verbs that implicitly make reference to the experience of the speaker (e.g. *sound* as in *that sounds like Schmidt*,), and adnominal relative clauses containing first person singular expressions (e.g. *and that was one thing* **I** *noticed,*). Also noted in this field is when a speaker uses a second person singular subject to refer to her own experience (e.g. *and they look at* ***you*** *and they go, ... <Q the what Q>?*).

2.4.34 Number of intonation units

The number of intonation units in each utterance is entered in this field. The values are *one* (=the utterance is coextensive with the IU), *two*, *more than two, part of IU* (=more than one utterance within an IU), and *truncated*.

2.4.35 Transitional continuity

As noted in section 2.3.2., the Du Bois et al. transcription system codes transitional continuity for each intonation unit. Transitional continuity is a functional description of intonation and may be described as "the marking of the degree of continuity that occurs at the transition point between one intonation unit and the next" (Du Bois et al. 1993: 53). The values for this category are *final* (often realized in English as a lowering of pitch at the end of an intonation unit), *continuing* (in English this often manifests as a slight rise in pitch at the end of an IU), *appeal* (in English this is frequently signaled as a high rise in pitch at the end of an IU), and *none* (this value is used for utterances that are part of an intonation unit or belong to a truncated intonation unit). For utterances that include more than one IU, transitional continuity of the last IU in the utterance is coded.

2.5. Summary

This chapter describes the source, transcription, and coding of the data used for the study. The bulk of the chapter, however, deals with the last task—delineating the methods used to segment and label conversational utterances according to their grammatical and lexical substance. This coding system is complex in that several elements of each clause are tagged for a variety of semantic and syntactic properties; however such extensive classification is warranted because of the nature of the topic under investigation. The claim made in this study—that linguistic subjectivity in American English conversation is a phenomenon that unites lexical and grammatical material, and that the structures and meanings contributing to this ubiquitous expression of speaker point of view correlate with their higher frequency in natural discourse—requires a coding system that is potentially sensitive to a wide range of structural and semantic material.

Ascertaining the high frequency of a single expressive lexical item (e.g. an intensifier, a value adjective) contributes to the characterization of subjectivity in English without a doubt, but such an endeavor does not on its own describe this phenomenon. Similarly, identification of a highly used predicate type (e.g. copular construction) or subject-verb combination (e.g. first person singular pronoun plus verb of cognition), which handily mark speakers' opinions and

attitudes, enhances our understanding of linguistic subjectivity, but does not in itself provide a broad view of it. Subjectivity is a general notion that is not tied to one particular linguistic expression or category. Therefore, providing evidence for the conventionalization of subjective forms in English interactive discourse requires attending to a range of combinations of grammatical and discursive elements and constructions that appear frequently in conversation. The coding system underlying this study, then, is constructed both to test hypotheses concerning the subjective nature of linguistic expression in English and also to allow for more open-ended exploration. Subsequent chapters present results of these inquiries.

Notes

1. I would like to acknowledge the *Corpus of Spoken American English* at the University of California, Santa Barbara for permission to use these recorded conversations and transcripts. In particular, I would like to thank SandraThompson and Margaret Field for making the materials available to me.

2. Transcription symbols are listed in Appendix A.

3. In the present study, the words *category* and *field* are used to refer to the actual topics used to describe utterances (e.g. *utterance, clause type, main verb, subject*). The terms *value* and *code* refer to the possible dimensions within a category that are assigned to a given utterance (e.g. the values, or codes, for the category *clause type* are *main, coordinate, subordinate*, and *complement*).

4. Numbers following linguistic examples locate utterances in the database.

5. A text field is one in which what is entered is simply text; that is, it is not a controlled classification system. This is in contrast to other fields for which predetermined values are assigned, and coding of a given utterance for those fields is limited to one of the values (e.g. syntactic type, main verb type).

6. IU boundaries in these data are indicated in four ways: a period (.) indicates a final intonation contour; a comma (,) indicates continuing contour; a question mark (?) indicates an appeal; and a double hyphen (--) marks a truncated IU. See Appendix A for a key to all transcription symbols.

7. Bolding highlights the parts of utterances under discussion.

8. These semantic verb classes are from Dixon (1991).

9. See chapter 4, section 4.2. and throughout for discussion and examples.

10. Paul Hopper, personal communication.

11. Quirk et al. (1985) call the situation type expressed by this class of verbs *stance*.

12. See chapter 4 for discussion of this issue.

Chapter 3

Patterns of subjectivity in person and predicate

3.1. Introduction

Two theoretical assumptions discussed in chapter 1 are worth restating in introducing the analyses presented in the next two chapters: (1) language—in particular, spontaneous conversation—is subjective in that it is fundamentally used by speakers to express their perceptions, feelings, and opinions, and (2) conventionalized linguistic structure, or grammar, emerges from repetition, or frequency of use, of sequences of lexical and grammatical elements in natural discourse. The union of these two ideas suggests a general hypothesis that linguistic elements that commonly appear in conversation should be those that participate in subjective expression (e.g. the frequent use of value adjectives such as *great* or *good*). There should also be greater cooccurrence of items whose combinations lend themselves to conveying speaker point of view than those whose combinations do not (e.g. after Benveniste 1971, verbs of cognition would more frequently appear with a first person singular subject than with a third person singular). In other words, with respect to both units and combinations of units (constructions of all sizes), there should be associations between commonly occurring conversational material and semantic and pragmatic expression of subjectivity.

This chapter examines relations between subjects and predicates in order to target the most frequent subject-verb combinations in American English conversation. Distributions of properties of subjects and verb types represented in these data are skewed in the sense that subjects do not equally cooccur with all verb lexemes in fact or in frequency. Structural and functional patterning of grammatical and lexical elements reveals that the most common combinations of subjects and predicates in the conversational database are precisely those that express speaker point of view—those that allow participants to personalize their conversational contributions, evaluate, and index attitude and situation. In contrast to these prominent trends in the corpus, the

more propositional acts of reporting found in interactive discourse are constrained to a much greater degree in distribution (i.e. they tend to occur during conversational narratives), and are themselves frequently mediated by subjective and interactive linguistic elements.

Highlighted in this portion of the study are properties of the subject (person, animacy, referentiality) and the predicate (main verb, main verb type, tense, aspect, modal elements). Only utterances with expressed subjects and main verbs are considered for the analyses presented in this chapter. These items account for 2,172 of the total 2,425 items in the database. Not included are clauses with unexpressed subjects (e.g. *looks good, and got a ... DWI that night*,) and the majority of those without main verbs (e.g. *he couldn't*).

The general organization of this chapter is by subject (1s, 2s, 3s, 1p, 3p),[1] and within each of these major sections, the most frequently occurring predicates are discussed. The largest group of third person singular subjects—those that occur with relational predicates (primarily copular clauses)—serves as the main focus of chapter 4. To ground specific discussions of person-predicate combinations in this chapter, section 3.2. outlines general distributional trends found in the corpus as a whole.

3.2. Global frequency patterns in the data

Table 3.1. summarizes the most frequently occurring subjects in the database. Third person singular occurs as the subject in 930 utterances out of the total 2,172, or in 43 percent of the tokens; the next most common subject is first person singular in 28 percent of the items. Second person singular subjects account for 15 percent of the data; 3 percent of all clauses have a first person plural subject, and 10 percent appear with third person plural subjects. Notably, 64 percent of all utterances in the corpus are present tense, and this pattern persists for all subjects, though the ratio of present to past (and to modal) varies by person. For example, only 14 percent of 2s utterances occur in the past tense (48/335)—proportionally fewer past tense predicates than for any other subject. Predicates with 2s subjects also more frequently contain central modals than do clauses with other subjects (58/335, or 17 percent).

Table 3.2. outlines the distribution of semantic verb types with respect to subject. The most frequent verb class is relational; these utterances account for

Table 3.1. Utterances by person and tense (n=2,172).

	Present	Past	Modal	Total	Percent
1s	348	191	78	617	28.41%
2s	229	48	58	335	15.42%
3s	626	240	64	930	42.82%
1p	35	21	10	66	3.04%
2p	3	1	1	5	0.23%
3p	147	50	22	219	10.08%
Total	1388	551	233	2172	100.00%
Percent	63.90%	25.37%	10.73%	100.00%	

approximately 30 percent of all predicates, and 497 of these relational verbs, or 78 percent (497/641), have third person singular subjects. The next most frequent verb class is the material type—a large heterogeneous group of lexical items whose meanings fall under the general processes of 'doing' and 'happening' (Halliday 1994); these items make up 25 percent of all utterances. The third most frequent verb type in the corpus are verbs of cognition (16 percent of the total): notably, 57 percent of these tokens cooccur with first person singular subjects (195/340), and 32 percent (110/340), with second person singular subjects. Finally, predicates designating verbal processes account for 11 percent of the main verb types, and over half of these (52 percent) appear with first person singular subjects.

3.3. First person singular subjects

3.3.1 Introduction

First person singular is the prototypical site for expression of speaker point of view and, as indicated by the distributions reported in Table 3.2., the second most frequently occurring subject in the corpus. As summarized in Table 3.3., *I* most often appears with verbs of cognition (32 percent of 1s subjects), material verbs (23 percent of 1s subjects), and verbal process predicates (20 percent of 1s subjects). The remaining 25 percent of first person singular subjects are spread among the other verb types (except perception/relational, which only occurs with third person subjects). Not only are verbs of cognition the most frequently occurring verb class with first person singular subjects in the conversational corpus, but, as shown in the right-hand most column of

Table 3.2. Utterances by person and verb type (n=2,172).*

	1s	2s	3s	1p	2p	3p	Total
cognition	**195**	**110**	**15**	**6**	**0**	**14**	**340**
	31.60%	**32.84%**	**1.61%**	**9.09%**	**0.00%**	**6.39%**	**15.65%**
corporeal	24	7	30	1	1	3	66
	3.89%	2.09%	3.23%	1.52%	20.00%	1.37%	3.04%
existential	12	6	62	3	0	8	91
	1.94%	1.79%	6.67%	4.55%	0.00%	3.65%	4.19%
feeling	19	9	10	2	0	5	45
	3.08%	2.69%	1.08%	3.03%	0.00%	2.28%	2.07%
material	**143**	**90**	**176**	**30**	**2**	**100**	**541**
	23.18%	**26.87%**	**18.92%**	**45.45%**	**40.00%**	**45.66%**	**24.91%**
perception	27	19	6	10	0	2	64
	4.38%	5.67%	0.65%	15.15%	0.00%	0.91%	2.95%
perception/rel	0	0	35	0	0	4	39
	0.00%	0.00%	3.76%	0.00%	0.00%	1.83%	1.80%
possessive/rel	21	31	28	5	0	17	102
	3.40%	9.25%	3.01%	7.58%	0.00%	7.76%	4.70%
relational	**50**	**41**	**497**	**6**	**2**	**45**	**641**
	8.10%	**12.24%**	**53.44%**	**9.09%**	**40.00%**	**20.55%**	**29.51%**
verbal	**126**	**22**	**71**	**3**	**0**	**21**	**243**
	20.42%	**6.57%**	**7.63%**	**4.55%**	**0.00%**	**9.59%**	**11.19%**
Total	617	335	930	66	5	219	2172
	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

*Individual percentages below utterance counts refer to person (e.g. 31.60% of items with 1s subjects occur with verbs of cognition). The right-hand column shows the percentage of each verb type in the corpus (e.g. 11.19% of all semantic verb types are verbal).

Table 3.3., predicates with 1s subjects account for 57 percent of all verbs of cognition in the database.

3.3.2 Verbs of cognition with 1s subjects

Table 3.4. presents counts of verbs of cognition with 1s subjects by tense and aspect; for these data, aspect values are simple, progressive, and anterior. The three most frequent lexical verbs in this group are *think* (84/195 or 43 percent), *know* (56/195 or 29 percent), and *guess* (17/195 or 9 percent).[2] Note, too, that 131/195, or 67 percent, of these predicates are present tense. Also of interest is that 31/50, or 62 percent, of the past tense tokens are usages of the verb *think*. And finally, worth noting is that there are only 17 predicates

Table 3.3. First person singular subjects by tense and verb type (n=617).

	Present	Past	Modal	Total	% Verb Type
cognition	**131**	**50**	**14**	**195**	57.35%
	37.64%	**26.18%**	**17.95%**	**31.60%**	
corporeal	17	5	2	24	36.36%
	4.89%	2.62%	2.56%	3.89%	
existential	5	3	4	12	13.19%
	1.44%	1.57%	5.13%	1.94%	
feeling	10	6	3	19	42.22%
	2.87%	3.14%	3.85%	3.08%	
material	**46**	**61**	**36**	**143**	26.43%
	13.22%	**31.94%**	**46.15%**	**23.18%**	
perception	9	12	6	27	42.19%
	2.59%	6.28%	7.69%	4.38%	
possessive/relational	12	7	2	21	20.59%
	3.45%	3.66%	2.56%	3.40%	
relational	23	22	5	50	7.80%
	6.61%	11.52%	6.41%	8.10%	
verbal	**95**	**25**	**6**	**126**	51.85%
	27.30%	**13.09%**	**7.69%**	**20.42%**	
Total	348	191	78	617	28.41%
	100.00%	100.00%	100.00%	100.00%	

marked progressive in both present and past tense, and 10 of these (59 percent) are also instances of the verb *think*. The rest of the analysis in this section proceeds by tense.

3.3.2.1 *Present tense verbs of cognition with 1s subjects*

The most frequent lexical verb in the present tense group is *know* (n=52), and 77 percent (40/52) of these uses occur in the construction *I don't know*, which exhibits a range of semantic and pragmatic functions in conversation (Scheibman 2000; Tsui 1991). For example, in (1), the speaker is not conveying her lack of knowledge about an entity or event; rather, *I don't know* in this context allows the speaker to avoid expressing too strong an opinion, in this case about the influence of Paris on French culture.

(1) …(1.5) but --
.. I don't know.
I'm not saying that Paris is the whole --

Table 3.4. Verbs of cognition with first person singular subjects by tense and aspect (n=195).

	Present				Past				Modal	Total
	Simple	Prog	Anter	Pres Total	Simple	Prog	Anter	Past Total		
believe	1			1					3	4
catch					1			1		1
figure	1			1	1			1		2
figure out		1		1	2	3		5		6
find out	2		1	3		1		1	2	6
forget	2			2	2			2		4
guess	**17**			**17**						**17**
imagine									1	1
know	**52**			**52**	**2**		**1**	**3**	**1**	**56**
learn					1			1	1	2
realize	1			1	2			2		3
remember	2			2	1	1		2	1	5
suppose	1			1						1
think	**44**	**4**		**48**	**23**	**6**	**2**	**31**	**5**	**84**
wonder	2			2		1		1		3
Total	125	5	1	131	35	12	3	50	14	195

…(1.5) everything that –
I mean as represented by everything in France,
but there certainly is a .. strong -- C10–6

The second and third most frequent combinations of first person simple present tense subjects with verbs of cognition are tokens of the collocations *I think* (n=44) and *I guess* (n=17), respectively; the majority of these function as epistemic clauses (e.g. *I guess Scott's making some good bucks.*) or epistemic parentheticals (e.g. *he lives in Pullman, I think,*) (Kärkkäinen 1998; Thompson & Hopper 2001; Thompson & Mulac 1991). There are also four present tense *think* predicates in the progressive, and these were all uttered by the same speaker to preface expression of incredulity. For example, in (2), the speaker is expressing amazement that, in his opinion, people are not protecting themselves from HIV disease by consistently using condoms.

(2) I mean these guy=s,
… say these women don't care if they use rubbers or not.
this is what guys have been telling me,
… you know,

and I'm thinking,
I- I can't believe that, L15–43

Eighty-six percent, then, of all present tense 1s+verb of cognition combinations in the database are formulaic expressions such as, *I think, I don't know*, and *I guess*.[3] These constructions function epistemically or act to mitigate assertion or disagreement in conversation (Scheibman 2000). That is, these frequent subject-verb combinations organize expression of the speaker's point of view or carry out negative politeness strategies in interactive discourse. Notably, these constructions are not used to inform participants of the speakers' cognitive activities. Moreover, the frequency of these first person singular expressions with the verbs *think* and *know* contributes greatly to the frequency of affiliated grammatical and semantic categories; that is, the high use of these conventionalized expressions is in large part responsible for the frequency of the categories first person singular subjects, verbs of cognition, as well as for the individual verb lexemes *think* and *know* in the corpus.

3.3.2.2 *Past tense verbs of cognition with 1s subjects*

Table 3.4. shows that only 50/195 of cognition verbs with 1s subjects, or 26 percent, are morphologically marked past tense. The most (only) frequent lexical verb in this group is *think* (n=31) which accounts for 62 percent of these past tense tokens. With respect to aspect, 23 of these *think* tokens are in the simple past, 6 are marked past progressive, and there are 2 anterior predicates. As it true for the frequent 1s+verb of cognition combinations in the present tense, past tense tokens also personalize speakers' contributions as opposed to conveying propositional information about participants' cognitive states. Regarding these past tense *think* tokens, though they are morphologically marked past tense, expressions such as *I thought* or *I was thinking* do not typically indicate past time reference. For example, the speaker uttered the intonation unit in (3) at the moment an overhead light spontaneously went on in the room. This use of past tense, then, has present relevance, since the speaker was not conveying something she *thought* in the past except as it framed the current situation.

(3) **I thought** we got our electricity fixed. L16–49a

Similarly, in (4), Marci's *I was thinking* refers to her trying to remember—during the present conversation, not in the past—who it was who told her Trish was pregnant.

(4) WENDY: you know,
Trish is pregnant again.
KENDRA: yeah.
WENDY: Trish Rhodes.
WENDY: .. [yeah].
MARCI: [(H)]= .. was it Trish=,
… t- who told me she was [pregnan=t,
WENDY: [she looked r=eally good].
MARCI: .. at],
.. I was thinking it was -- A7–47

3.3.2.3 *Verbs of cognition with non–1s and non–2s subjects*
There are only 35 tokens, or 10 percent, of verbs of cognition that appear with non-first person and non-second person singular subjects (see Table 3.2.). Fifteen of these occur with third person singular human subjects (*s/he* or lexical NPs), and often these assertions that relate to another person's mental states are tempered by the speaker, as in (5), in which the speaker uses the epistemic parenthetical *I think* to soften his statement.

(5) **he thought I think** he'd have to pull it in second gea=r. F6–35a

Six utterances with cognitive predicates have first person plural subjects, and 14 cooccur with the third person plural subject *they*. In this latter group, all but two of the utterances have subjects that are referentially generic (see section 3.7.) In (6), for example, *they* refers to an imprecisely delineated group of people. It seems to be the case that speakers hesitate to make assertions about other people's awareness or opinions, unless these statements are mediated (hedged) or the subject referents are generically construed.

(6) **they don't know** what they're doing? R12–17a

3.3.3 Material verbs with 1s subjects

Recall from the discussion of main verb types in chapter 2 (section 2.4.23.) that material verbs designate processes of 'doing' and 'happening', or *processes of the external world* (Halliday 1994). In this corpus, material verbs are the second largest class of verbs and also the group containing the most lexemes (168 different lexical verbs). Furthermore, material verbs constitute the second largest verb class for utterances with first person singular subjects (see

Table 3.3.): 23 percent of 1s clauses contain material verbs (n=143). With regard to linguistic subjectivity, an interesting question surfaces related to the high frequency of this verb class with 1s subjects: If *I* is the prototypical site for a speaker's evaluative and organizing expression, then why does this subject pronoun so frequently occur with predicates referring to propositional events and activities? The broad delineation of the material verb category certainly accounts for its high frequency with all subjects. Even given the heterogeneity of the class, looking at the low frequency of individual verbs in this group is informative, as will be illustrated below.

In the case of *I*+verbs of cognition, the high frequency of the group is related to the high frequency of *I* plus particular lexical items in the group (e.g. *think, (don't) know, guess, thought*); the situation for *I*+material verbs is different. The most frequently occurring lexical verbs in the 1s+material group are neither terribly frequent nor terrifically lexical: *go* (n=19), *do* (n=15) *get* (n=13). Although these three verbs as a group account for 33 percent of 1s+material items, there are no individual, highly repeated expressions in this set. There are 65 lexical material verbs and 143 individual tokens in the 1s+material set of verbs (2.2 tokens to each verb lexeme), compared to the 1s+verb of cognition group which contains 15 lexical verbs and 195 individual items (13 tokens to each verb lexeme).

3.3.3.1 *Central modals with material verbs with 1s subjects*

Of the 78 first person singular predicates containing modal auxiliaries, 36, or, 46 percent, occur with material verbs. This high percentage is in part due to the frequency of subjects with human referents in these utterances because it is typically people who are discussed in terms of processes of 'doing' and 'happening'. However, in interactive discourse, it appears that assertions about activities (material processes) are frequently accompanied by mediating elements.

Table 3.5. is identical to Table 3.3., except that in Table 3.5. percentages are given for tense instead of verb type for first person singular utterances (e.g. for 1s cognition predicates: 67% are present tense, 26% are past tense, and 7% contain modals). Note that the proportion of central modals is greater for material predicates than for the other frequent verb types in the corpus (relational, material, cognition, verbal).[4] That is, 25 percent of 1s+ material predicates contain modal auxiliaries. Compare this to other frequent verb types that typically occur with human subjects as well, such as verbs of

Table 3.5. First person singular subjects by verb type, breakdown by tense (n=617).

	Present	Past	Modal	Total
cognition	**131**	**50**	**14**	**195**
	67.18%	**25.64%**	**7.18%**	**100.00%**
corporeal	17	5	2	24
	70.83%	20.83%	8.33%	100.00%
existential	5	3	4	12
	41.67%	25.00%	33.33%	100.00%
feeling	10	6	3	19
	52.63%	31.58%	15.79%	100.00%
material	**46**	**61**	**36**	**143**
	32.17%	**42.66%**	**25.17%**	**100.00%**
perception	9	12	6	27
	33.33%	44.44%	22.22%	100.00%
possessive/rel	12	7	2	21
	57.14%	33.33%	9.52%	100.00%
relational	23	22	5	50
	46.00%	44.00%	10.00%	100.00%
verbal	**95**	**25**	**6**	**126**
	75.40%	**19.84%**	**4.76%**	**100.00%**
Total	348	191	78	617
	56.40%	30.96%	12.64%	100.00%

cognition or verbal process predicates; the modal proportions for these two groups are 7 percent and 5 percent, respectively. This contrast suggests that it is not humanness of subject referent alone that conditions the greater presence of modals with this verb class.

The greater proportion of central modals with material predicates is a trend for all subjects throughout the database (see, in particular, section 3.4.3.1. below). This fact suggests that in interactive discourse in English even the most potentially informative, or propositional, of predicate types are structurally integrated into the speaker's point of view, marked in this case by the presence of modal expressions. For example, in (7), the speaker's description of her activity (accessing the file) is mediated by the modal *can't* which subjectifies the material predicate; that is, this utterance is not *about* a process of 'accessing', but rather about the speaker's role in the process (in this case, *not being able to* accomplish the process).

(7) and **I can't access** the file. R13–30

3.3.3.2 *Present tense material verbs with 1s subjects*

Table 3.5. also demonstrates that 1s+material verb is one of the few subject-verb type combinations in the corpus for which there are more tokens in the past tense (the typical site of reporting on events and activities) than in the present. Forty-three percent of the items in this group (n=61) are past tense and 32 percent are present tense (n=46).[5]

Of the 46 present tense material verbs with *I* subjects in the corpus, 26 of these (57 percent) contain an intermediate function verb in their predicates (e.g. *gonna, want to, need, try to*). Similar to (7) above, notice that (8) and (9) do not concern activities related to *sleeping* or *dancing*, respectively. As discussed for central modals with material verbs, these modal-like elements also mitigate the propositionality of the predicate by personalizing the event, or activity, designated by the main verb—in essence, making the predicate more subjective.

(8) **I need to get** sleep over the weeke=nd. A6–41b

(9) **I wanna go out** lambada dancing with you=. L18–45

Another subset of the present tense 1s+material verb group are those predicates that express habitual meaning (11/46).[6] In (10), the speaker is discussing the general state of her relationship with her boss at work.

(10) and **I don't play** that game.[7] G18–8

While habitual expressions certainly convey information (e.g. in 1s utterances with material verbs, about what the speaker usually does), they also function as generalizations. And the act of generalizing is itself evaluative.

3.3.3.3 *Past tense material verbs with 1s subjects*

First person singular subjects with material verbs constitute one of the few subject-verb groups in the database that reflects a reversal of the general trend of present tense predicates being more frequently occurring than past tense. Forty-three percent of material verbs with 1s subjects are morphologically marked as past tense (61/143) compared to 32 percent marked as present tense (46/143). Only 10, or 16 percent, of these past tense utterances contain some kind of intermediate function verb; in comparison, 57 percent of present tense 1s+material verb utterances have an IFV (see section 3.3.3.2. above).

These 1s+material verb past tense utterances are used by speakers to report on past experiences, as illustrated in (11), (12), and (13). However, given that the most frequent verbs in this group are basic English verbs (e.g. *do, get*), the level of propositionality, or referential informativity, of these unmediated material predicates is relatively general.

(11) **I** just **went** to some local doctor, D10–42

(12) **I did** well, on the first block. G13–38

(13) **I put** their papers up on the board. R19–5

3.3.4 Verbs of verbal process with 1s subjects

The fourth most common verb type in the corpus is the group of verbal processes. This class not only includes verbs of saying but also other symbolic processes such as meaning (Halliday 1994). Notably, first person singular subjects account for over half of all verbal process predicates in the corpus (126/243, or 52 percent) (see Table 3.2.), and 75 percent of these (95/126) occur in the present tense (see Table 3.3.).[8] Table 3.6. presents the distribution of verbal process predicates occurring with first person singular subjects in the database.

Table 3.6. Verbs of verbal process with first person singular subjects (n=126).

	Present	**Past**	**Modal**	**Total**
ask		2	1	3
be like (quotative)	4		1	5
bet	3		1	4
go			2	2
mean	77			77
propose		1		1
put		1		1
say	4	8		12
show	1			1
sound	1			1
suggest			1	1
swear	2			2
talk		5		5
tell	3	8		11
Total	95	25	6	126

Most striking about this group is that 61 percent of all tokens (77/126) and 81 percent of present tense tokens (77/95) are instances of the collocation *I mean*. Schiffrin (1987) characterizes *I mean* as a discourse marker that prefaces expansion or explanation of speakers' contributions in conversation. As is true for *I think* and *I guess*, *I mean* also functions epistemically (cf. Thompson & Hopper 2001) as in (14).

(14) **I mean** Rene and Anne were very sweet. C11–38a

The lexical verb *mean* only occurs with non-first person singular subjects seven times in the database: four times with the subjects *this* and *that*, and three times with *you* (e.g. *what do you mean,*). Its frequency, then, is due to its participation in the discourse collocation *I mean*.

3.3.5 Summary of utterances with first person singular subjects

In this corpus of spoken American English conversation, the first person singular pronoun occurs as the subject of 28 percent of all utterances. Clauses with 1s subjects account for 57 percent of the verbs of cognition, 52 percent of verbal process predicates, and 26 percent of material verbs. Within the category of utterances with 1s subjects itself, 32 percent of the items appear with verbs of cognition, 23 percent with material verbs, and 20 percent with verbal process predicates; that is, 75 percent of 1s subjects (464/617) occur with these three verb types. The high frequency of *I* with present tense verbs of cognition and verbal processes is attributable to the high frequency of individual routinized expressions (e.g. *I think, I mean*) that personalize and organize the speaker's contribution. Even for the group of past tense cognitive verbs, the most frequent item is *I thought* which neither conveys information about past events nor even typically designates past time reference.

With respect to material verbs, the situation differs from the cognition and verbal types. The high frequency of this verb class is in part related to the broad delineation of the category—that this group includes many lexical verbs. Even so, it is notable that there are no highly frequent verbs in this particular group, and that the most common lexical items are basic verbs in English that form low content, nonreferential verb-object combinations (Thompson & Hopper 2001), or dispersed predicates (Hopper 1991). The majority of present tense material verb predicates with first person singular subjects contain intermediate function verbs (modal elements) (57 percent),

and another 24 percent express habitual meaning, which was suggested to be generalizing, or evaluative in nature. On the other hand, past tense material verbs with first person singular subjects contain proportionally fewer intermediate function words than do their present tense counterparts.

In general, the data suggest that material verbs cooccur more frequently with expressions of modality (central modals and intermediate function verbs) than do other frequent verb types with subjects whose referents are typically human as well (e.g. cognitive, verbal). This argues against the idea that it is the presence of a human or speech act participant subject alone that conditions greater use of modal elements. Instead, it appears as though when participants are relating events in the world (material processes), the information is mediated by the speaker's subjective, or evaluative, stance.

3.4. Second person singular subjects

3.4.1 Introduction

Utterances with second person singular subjects are the third most frequent subject type in the database, and they account for 15 percent of the total (n=335). Notably, only 14 percent of *you* subjects occur with past tense predicates (48/335); this is in contrast to all other subjects whose past tense proportions range from 23 percent to 32 percent. Table 3.7. displays the distribution of second person singular subjects by verb type. With respect to the most frequent verb types in the corpus, 2s utterances make up 32 percent of all verbs of cognition and 17 percent of material verbs. Within the category of second person singular itself, 33 percent of the predicates are verbs of cognition (110/335), and 27 percent are material verbs (90/335). The combinations of *you* with these two most frequent verb types will be the topics of this section.

3.4.2 Verbs of cognition with 2s subjects

Table 3.8. displays the distribution of cognitive verbs with second person singular subjects. Quite clearly the frequency of this subject with this verb type is due to the conversational frequency of one item—the formulaic *you know*, which accounts for 98/110, or 89 percent of the total utterances in this group. As is true for *I don't know*, *I think*, *I mean*, and *I guess*, the frequency of *you*

Table 3.7. Second person singular subjects by tense and verb type (n=335).

	Present	Past	Modal	Total	% Verb Type
cognition	**100**	**6**	**4**	**110**	32.35%
	43.67%	**12.50%**	**6.90%**	**32.84%**	
corporeal	5	0	2	7	10.61%
	2.18%	0.00%	3.45%	2.09%	
existential	4	2	0	6	6.59%
	1.75%	4.17%	0.00%	1.79%	
feeling	7	1	1	9	20.00%
	3.06%	2.08%	1.72%	2.69%	
material	**44**	**19**	**27**	**90**	16.64%
	19.21%	**39.58%**	**46.55%**	**26.87%**	
perception	4	3	12	19	29.69%
	1.75%	6.25%	20.69%	5.67%	
possessive/relational	19	7	5	31	30.39%
	8.30%	14.58%	8.62%	9.25%	
relational	32	4	5	41	6.40%
	13.97%	8.33%	8.62%	12.24%	
verbal	14	6	2	22	9.05%
	6.11%	12.50%	3.45%	6.57%	
Total	229	48	58	335	15.42%
	100.00%	100.00%	100.00%	100.00%	

know is in large part responsible for the high frequency of grammatical and lexical categories in the database with which it is affiliated (e.g. 2s subjects, verbs of cognition, present tense).

Y*ou know* fulfills automated, interactive functions in conversation. Schiffrin (1987) discusses specific uses of *you know* in both conversation and narrative; one function of the expression in conversation is that it allows speakers to check in with other participants to make sure they share relevant knowledge as they are talking. Schiffrin also observes that in conversational narrative, *you know* occurs with evaluative utterances to encourage hearers to attend to speakers' own assessments. In (15), the speaker is justifying distancing herself from her brother during a period when he had a disruptive drug habit; *you know* in this episode appears to solicit support for her stance, illustrating the interactive evaluative function of this expression suggested by Schiffrin.

(15) but you have to --
you have to at one point let go.
you can't constantly be torn,

Table 3.8. Verbs of cognition with second person singular subjects (n=110).

	Present	Past	Modal	Total
find out		1		1
know	**98**	**2**		**100**
know better			1	1
learn		1		1
make up			1	1
remember			1	1
think	1	2	1	4
think about	1			1
Total	100	6	4	110

.. (H) just to=rn to pie=ces,
by=,
you know=,
somebody like tha=t. D23–1

However, as illustrated in (15), a functional distinction between *you know* serving as a metalinguistic check on the sharing of information by participants versus the expression's garnering support for the speaker's evaluative stance is not easy to determine, in part due to the more general problem of distinguishing informativeness from evaluativeness in linguistic utterances. Similarly, in (16), it is ambiguous whether the speaker's *you know* is guiding the participants to attend to the relevant facts of the story or to the speaker's own assessment of the facts.

(16) I started noticing,
about thirty-five I think,
that,
you know,
lines on the face, G12–7

With respect to the second person singular pronoun in the fixed expression *you know*, *you* is always slightly referential in that it tacitly refers to the addressee. But *you* is also generic in the sense that the pronoun in the collocation does not refer to an individual person. This weakness, or generality, of referentiality of the subject is not unlike the situation of *I* in the expressions *I mean* or *I think*, in which the speaker is not specifically referring to herself in these usages; rather, the entire collocation takes on pragmatic function with a concomitant reduction in specificity of the pronominal referent.

Table 3.9. Material verbs by 1s, 2s, 3s subjects and tense (n=409).

	Present	Past	Modal	Total
1s	46	61	**36**	143
	32.17%	42.66%	**25.17%**	100.00%
2s	44	19	**27**	90
	48.89%	21.11%	**30.00%**	100.00%
3s	83	70	**23**	176
	47.16%	39.77%	**13.07%**	100.00%
Total	172	150	**87**	409
	42.05%	36.68%	**21.27%**	100.00%

3.4.3 Material verbs with 2s subjects

Twenty-seven percent of clauses with second person singular subjects contain material verbs (90/335). As is true for first person singular utterances with material verbs, there are no highly frequent lexical verbs in this group; rather, the frequency of the type comes from the large number of lexical verbs in the category (51 types, 90 tokens in this case).

3.4.3.1 *Central modals with material verbs with 2s subjects*

Table 3.9. shows the distribution of tense and modal auxiliaries for material predicates with 1s, 2s, and 3s subjects. Recall that for utterances occurring with 1s subjects, the proportion of central modals is greater for material predicates than for the other frequent verb types in the corpus (relational, cognition, verbal). It was suggested in those discussions that speakers tend to personalize these more propositional predicates by using modal expressions.

Table 3.9. demonstrates that the proportion of material predicates occurring with modals is higher for first person and second person singular subjects (25 percent and 30 percent, respectively) than for third person singular subjects (13 percent). In the case of second person singular utterances, when speakers address another speech act participant directly, they often mediate their assertion towards the addressee by using a modal element, as in (17).

(17) **you can use** this for your muffins. A13–20

Because of interpersonal considerations such as attention to negative face needs, these second person singular utterances containing material verbs and modal elements reflect a kind of interactive subjectivity, in that speakers use

personalizing, or subjectifying, elements to construct polite or empathetic construals of another person's agentivity.

3.4.3.2 *Present tense material verbs with 2s subjects*

Table 3.9. indicates that almost half of the 2s+material verb clauses occur in the present tense (n=44), though there appears to be little evidence of an informative, or propositional, character to these utterances. Structurally and functionally, this group of 2s present tense material process clauses consists of: (a) predicates with intermediate function words, e.g. *ought to*, as in (20) (25 percent of this group, or 11/44 of the tokens), (b) interrogatives and other questions, (c) utterances with subjects that are best categorized as generic, often conveying first person singular reference (e.g. (18) in which the speaker is describing his own reaction), or (d) subordinate clauses following evaluative or evidential main clauses as in (19) and (20), respectively. What is common to all of these structures is that they restrict direct assertion of information.

(18) it smells like **you're walking** past a d=umpster. D1828b

(19) that's good **you're getting** r- good rest. A642b

(20) I think **you ought to change** doctors. D10–38b

3.4.3.3 *Past tense material verbs with 2s subjects*

There are only 19 utterances with second person singular subjects with material verbs in the past tense, and none of these predicates contains an intermediate function verb (a modal element). Another difference between this group and the more frequent present tense material clauses described immediately above is that in these past tense 2s material utterances (with the exception of one hypothetical clause) all the subject pronouns refer to specific discourse participants (i.e. they are not generic). Even so, there are two ways in which the informative, or propositional, nature of these past tense items are mediated structurally as discussed for the present tense: (a) almost half of the group are questions, and (b) several of the tokens are subordinate clauses preceded by epistemic and evaluative expressions (e.g. *I thought **you already brought** the cake in. well, glad **you stopped i=n,***). Unlike the present tense 2s+material verb group, however, there are a few clauses in this past tense set that do express direct assertions about the subject, illustrated in (21) and (22).

(21) but .... **you wrote**,
Z's aging again, G17–8

(22) **you met** Benno, D20–25

3.4.4 Summary of utterances with second person singular subjects

Second person singular subjects are the third most frequent group of utterances in the database, and they make up 15 percent of the corpus (n=335). The two most frequent verb types occurring with *you* subjects are verbs of cognition (33 percent) and material verbs (27 percent). Proportionally, there is less past tense and more frequent occurrence of modal auxiliaries in 2s predicates than for other subjects in the corpus.

Eighty-nine percent of 2s verbs of cognition (and 29 percent of all 2s utterances) are tokens of the fixed expression *you know*. These, of course, are all present tense. Forty-nine percent of 2s material verbs are also present tense; this group consists of questions, utterances with generic *you* reference (often interpreted as first person), and so-called subordinate clauses prefaced by epistemic and evaluative clauses. In all of these present tense forms, speakers personalize and situate their utterances in relation to another speech act participant; they do not assert information about the addressee. Additionally, the presence of generic *you* in many 2s utterances accompanies the making of generalizations, which are evaluative. Following the general trend noted above regarding the more frequent occurrence of central modals with 2s predicates, almost one third of second person subjects with material predicates also contain a modal auxiliary, and 25 percent of present tense material predicates with 2s subjects include an intermediate function word. These modal elements permit speakers to soften direct assertion about an addressee.

There are very few past tense predicates with second person singular subjects (6 tokens with verbs of cognition and 19 tokens with material verbs). Given that the past tense is typically an important vehicle for conveying propositional information (e.g. through narrated events), it is not surprising that English speakers use it infrequently with second person subjects due to face concerns. Though there are not many past tense tokens of material verbs in this group, a functional characteristic that distinguishes them from their present tense counterparts is that the past tense subjects are much more likely to be specifically referential than generic. The relationship between non-specific referentiality of the subject and subjective expression—in its broadest

sense, encompassing a variety of strategies and structures used by speakers to organize their expression—can be seen in fixed expressions such as *I mean* or *you know* in which the subjects exhibit low (general) referentiality. This issue will come up in discussion of third person plural utterances (section 3.7.) and also in the next chapter with regard to third person singular inanimate subjects of relational clauses.

3.5. Third person singular subjects

3.5.1 Introduction

Third person singular is the most frequent subject in the corpus; it accounts for 43 percent of all utterances (930/2172). Table 3.10 shows that 626/930 of 3s clauses (67 percent) are present tense, 240/930 (26 percent) are past tense, and 64/930 of these predicates (7 percent) contain modal auxiliaries. Third person singular is a more complex member of the English category of person than are first and second person singular subject pronouns discussed above (though the latter two are not without functional variation). This is because 3s includes several subject types that differ in animacy and referentiality, and these distinctions affect the distributions of 3s utterances in relation to verb type and tense. In its heterogeneity, third person singular subject is analogous to the material verb class, in that the frequency of the category as a whole is related to its being composed of many types. However, unlike the semantically broad class of material verbs for which there are no highly conventionalized combinations of subject + lexical verb in conversation, the class of third person singular subjects includes commonly occurring subject + verb types within it.

Table 3.10 indicates that the most frequent verb type to appear with third person singular subjects is relational, which accounts for 53 percent of 3s utterances (497/930). The second most frequent type with this subject—material verbs—occurs in 19 percent of 3s clauses (176/930). With respect to the database as a whole, the right-hand most column of Table 3.10. shows that 3s clauses account for more than half of several verb types. Notably, 78 percent of relational predicates (mostly copular constructions) in the corpus have 3s subjects. Third person singular utterances also constitute the majority of two lesser verb types: existential (e.g. *she wasn't there.*) and perception/relational (e.g. *that smells like ga=rbage,*).

Table 3.10. Third person singular subjects by tense and verb type (n=930).

	Present	Past	Modal	Total	% Verb Type
cognition	5	8	2	15	4.41%
	0.80%	3.33%	3.13%	1.61%	
corporeal	20	5	5	30	45.45%
	3.19%	2.08%	7.81%	3.23%	
existential	43	18	1	62	68.13%
	6.87%	7.50%	1.56%	6.67%	
feeling	7	3	0	10	22.22%
	1.12%	1.25%	0.00%	1.08%	
material	**83**	**70**	**23**	**176**	32.53%
	13.26%	**29.17%**	**35.94%**	**18.92%**	
perception	3	3	0	6	9.38%
	0.48%	1.25%	0.00%	0.65%	
perception/relational	31	4	0	35	89.74%
	4.95%	1.67%	0.00%	3.76%	
possessive/relational	14	10	4	28	27.45%
	2.24%	4.17%	6.25%	3.01%	
relational	**393**	**81**	**23**	**497**	77.54%
	62.78%	**33.75%**	**35.94%**	**53.44%**	
verbal	27	38	6	71	29.22%
	4.31%	15.83%	9.38%	7.63%	
Total	626	240	64	930	42.82%
	100.00%	100.00%	100.00%	100.00%	

Unlike previous sections in this chapter, the analysis of 3s utterances is not organized by verb type. Because the category English third person singular subject includes several members, it is necessary to examine distributions of usages of individual members of this larger group in combination with tense and verb type. This section, then, explores general interactions among the most frequent 3s subject types, tense (in particular, past tense), verb type (primarily material verbs), and the influence of animacy (humanness) of the subject referent on distributional patterns. The largest verb type appearing with 3s subjects—relational predicates—is the topic of chapter 4 and so will not be discussed in this section.

Table 3.11. presents counts of third person subject types by tense. Overall, the most frequent 3s subjects are *s/he* (27 percent of 3s subjects), *it* (25 percent), lexical noun phrases (19 percent), and *that* (16 percent). These four types comprise 86 percent of all utterances with 3s subjects in the database.

As a preface to discussion of members of this class, it is worth noting that

even though third person singular subject as a category is the most frequent subject in the corpus (43 percent of the total), none of the individual members of this group is as common in conversation as are first person singular subjects or even second person singular subjects. The most frequent 3s type in the database is *s/he* and occurs in 249 clauses; *I* subjects, on the other hand, appear in 617 utterances, and *you*, in 335. When 3s subject types are examined individually, then, global patterning of subjects in the corpus shifts.

There is an interesting commonality among the three most frequent subjects in the database (*I*, *you*, and *s/he*) with regard to animacy—they all have human referents. However, we will see below that this shared feature of human referentiality of subject does not produce identical distributions as regards tense and verb type. Speech act participant subjects occur with different predicates than does the 3s *s/he*. Moreover, as discussed above, the discourse participants *I* and *you* do not, with respect to each other, pattern in the same ways.

Table 3.11. Third person singular subject types by tense (n=930).

	Present	Past	Modal	Total	% 3 sg
s/he	**124**	**102**	**23**	**249**	**26.77%**
	49.80%	**40.96%**	**9.24%**	**100.00%**	
it	**164**	**51**	**14**	**229**	**24.62%**
	71.62%	**22.27%**	**6.11%**	**100.00%**	
lexical NP	**102**	**57**	**15**	**174**	**18.71%**
	58.62%	**32.76%**	**8.62%**	**100.00%**	
that	**117**	**16**	**12**	**145**	**15.59%**
	80.69%	**11.03%**	**8.28%**	**100.00%**	
there	40	10	0	50	5.38%
	80.00%	20.00%	0.00%	100.00%	
this	38	0	0	38	4.09%
	100.00%	0.00%	0.00%	100.00%	
what	28	3	0	31	3.33%
	90.32%	9.68%	0.00%	100.00%	
who	7	1	0	8	0.86%
	87.50%	12.50%	0.00%	100.00%	
they (sg)	4	0	0	4	0.43%
	100.00%	0.00%	0.00%	100.00%	
one	2	0	0	2	0.22%
	100.00%	0.00%	0.00%	100.00%	
Total	**626**	**240**	**64**	**930**	100.00%
	67.31%	**25.81%**	**6.88%**	**100.00%**	

3.5.2 Third person singular subject types and tense

As is true for all utterances in the corpus, for 3s predicates, present tense is more frequent than past tense—67 percent vs. 26 percent, respectively. However, as shown in Table 3.11., the proportion of present to past (to modal) varies by 3s subject type. For example, in the case of the human subject *s/he*, there is proportionally more past tense (present: 50 percent, past: 41 percent) than for the other subject types. This proportion of tense distribution for *s/he* clauses is in contrast to utterances with 3s subjects with inanimate referents, such as *it* (present: 72 percent, past: 22 percent), *that* (present: 81 percent, past: 11 percent), and *this* (present: 100 percent), for which there is proportionally less past tense than is found in utterances with the subject *s/he*. In other words, there appears to be an association in the database between third person singular subjects with human referents and a greater proportion of past tense predicates (or, conversely, proportionally fewer past tense predicates with 3s subjects with nonhuman referents). Examples of utterances with 3s subjects with human and nonhuman referents follow in (23) and (24), respectively.

(23) 3s subjects with human referents
  (a) **Mary**'s just lying. (present, verbal) A6–44
  (b) **he** just di- did a blood test, (past, material) D10–48

(24) 3s subjects with nonhuman referents
  (a) **everything** breaks on my kitchen floor. (present, material) A15–34
  (b) and **it** bothered me**,** (past, feeling) G10–40

Table 3.11. also shows that utterances with lexical noun phrase subjects are intermediate in their tense distribution, with respect to *s/he* utterances as one end of the continuum and *this* utterances as the other. Referents of lexical NP subjects may be either human or nonhuman (e.g. *some neighbor* vs. *the worst thing I ever had*, respectively). And what can be seen in the frequency counts is that the difference between present and past tense clauses for all 3s lexical NP subjects (present: 59 percent, past: 33 percent) falls between the present/past proportions for *s/he* utterances and those utterances with inanimate subjects (e.g. *it*, *that*). However, when utterances with lexical NP subjects are separated by animacy of referent, an interesting contrast emerges; thus this group provides an instructive example.

Table 3.12. indicates that clauses with lexical NP subjects with human referents have almost an equal number of present and past tense predicates (present: 45 percent, past: 43 percent), a pattern similar to the one seen for *s/he* utterances (present: 50 percent, past: 41 percent). On the other hand, lexical NP subjects with nonhuman referents pattern more like the inanimate 3s subjects *it, that,* and *this*. Seventy-three percent of the predicates with 3s lexical NPs with nonhuman referents are present tense, and only 22 percent are past tense. There is, therefore, a strong tendency for there to be a greater number of past tense predicates with all 3s subjects with human referents (not only for *s/he*) than for utterances with 3s inanimate referents. The consequences for subjectivity and propositionality of this association between animacy, or humanness, of subject and past tense usage will be discussed in section 3.5.4. following a review of frequency trends of third person singular subjects and verb types.

3.5.3 Third person singular subject types and verb type

Recall that the most frequent verb types occurring in third person singular utterances as a group are relational and material. Table 3.13. provides a breakdown of verb type by the most frequent 3s subjects. Consistent with the general trend for the group as a whole (see Table 3.10. above), the most frequent verb types cooccurring with *s/he, it,* lexical NP, and *that* are relational and material; however, the distributions of verb type vary by subject.

Notice in Table 3.13. that the most frequent verb types to occur with the human subject *s/he* are material verbs (35 percent), relational verbs (24 percent), and predicates of verbal process (17 percent). On the other hand, for the subjects *it* and *that*, the most frequent predicate is relational (68 percent and 84 percent, respectively).

Table 3.12. Third person singular lexical NP subjects by animacy and tense (n=174).

	Present	**Past**	**Modal**	**Total**
human referent	40	38	11	89
	44.94%	42.70%	12.36%	100.00%
nonhuman referent	62	19	4	85
	72.94%	22.35%	4.71%	100.00%
Total	102	57	15	174
	58.62%	32.76%	8.62%	100.00%

Table 3.13. Frequent third person singular subject types by verb type (n=797).

	s/he	it	lexical NP	that	Total
cognition	13	0	2	0	15
	5.22%	0.00%	1.15%	0.00%	1.89%
corporeal	19	5	4	2	30
	7.63%	2.18%	2.30%	1.38%	3.76%
existential	6	6	8	2	22
	2.41%	2.62%	4.60%	1.38%	2.76%
feeling	5	2	3	0	10
	2.01%	0.87%	1.72%	0.00%	1.25%
material	**88**	**31**	**42**	**10**	**171**
	35.34%	**13.54%**	**24.14%**	**6.90%**	**21.45%**
perception	1	1	3	0	5
	0.40%	0.44%	1.72%	0.00%	.63%
percept/rel	3	23	1	6	33
	1.20%	10.04%	0.57%	4.14%	4.14%
possess/rel	13	3	12	0	28
	5.22%	1.31%	6.90%	0.00%	3.51%
relational	**59**	**155**	**81**	**122**	**417**
	23.69%	**67.69%**	**46.55%**	**84.14%**	**52.32%**
verbal	42	3	18	3	66
	16.87%	1.31%	10.34%	2.07%	8.29%
Total	249	229	174	145	797
	100.00%	100.00%	100.00%	100.00%	100.00%

Table 3.14. provides a more succinct view of this information. In this table, frequent 3s subjects with material and relational predicates are broken down by animacy of the subject referent. Human subjects include *s/he* and lexical noun phrases with human referents. Nonhuman subjects are *it, that*, and lexical noun phrases with nonhuman referents.

Similar to tense distribution for 3s utterances discussed above, Table 3.14. indicates that animacy, or humanness, of subject referent is an important influence on cooccurring verb types (Dahl 2000). Thirty-three percent of 3s subjects with human referentiality appear with material predicates, and 25 percent of 3s with human subjects occur in relational clauses. On the other hand, 3s subjects with nonhuman referents overwhelmingly occur in relational utterances (copular clauses)—in 73 percent of the cases—while only 13 percent of these subjects appear with material verbs. In English interactive discourse, then, there is an association between 3s nonhuman subjects and relational verbs (see chapter 4), while 3s human subjects more frequently cooccur with both material and relational predicates.

Table 3.14. Frequent third person singular subjects with material and relational verbs by animacy (n=588).

	3s subjects with human referents	3s subjects with nonhuman referents
material	113	58
% subjects*	33.43%	12.64%
relational	83	334
% subjects*	24.56%	72.77%

* These percentages are based on the total number of subject tokens in each group. For example 3s subjects with human referents occurring with material predicates account for 33.43% of *all* 3s subjects with human referents (*s/he* tokens + human lexical NP tokens).

3.5.4 Humanness, verb type, and tense in 3s utterances

The preceding two sections delineated cooccurrence patterns among third person singular subject types, tense, and verb type found in the database. Frequency counts indicate that there are proportionally more past tense predicates with 3s subjects with human referents than with 3s subjects with nonhuman referents. With respect to verb type, we find that (primarily present tense) relational verbs account for the bulk of 3s subjects with nonhuman referents, whereas the majority of predicates with 3s subjects with human referents contain both relational and material verbs, and, to a lesser extent, verbal types.

Recall that the three most frequent subjects in the entire database are first person singular (n=617), second person singular (n=335), and third person singular *s/he* (n=229). All three of these items are singular pronominal subjects with human referents. In examining the effects of human referentiality on 3s predicates, it is informative to contrast these usages with the other frequent human subjects, 1s and 2s.

3.5.4.1 *Singular subjects with human referents and the distribution of verb type*

Table 3.15 and 3.16. show that distributions of frequent verb types overlap for 1s, 2s, and *s/he* in several ways.[9] First person singular and second person singular both have cognitive predicates as their largest group, whereas for *s/he* utterances, there are very few cognitive verbs. Human referentiality, then, is not a sufficient property of subjects in English to trigger frequent combination

Table 3.15. 1s, 2s, and *s/he* subjects by tense, breakdown by verb type (n=1,201).

	1s			1s Total	2s			2s Total	S/he			S/he Total
	Present	Past	Modal		Present	Past	Modal		Present	Past	Modal	
cognition	**131**	**50**	**14**	**195**	**100**	**6**	**4**	**110**	**4**	**7**	**2**	**13**
	37.64%	**26.18%**	**17.95%**	**31.60%**	**43.67%**	**12.50%**	**6.90%**	**32.84%**	**3.23%**	**6.86%**	**8.70%**	**5.22%**
corporeal	17	5	2	24	5		2	7	15	2	2	19
	4.89%	2.62%	2.56%	3.89%	2.18%	0.00%	3.45%	2.09%	12.10%	1.96%	8.70%	7.63%
existential	5	3	4	12	4	2		6	2	4	0	6
	1.44%	1.57%	5.13%	1.94%	1.75%	4.17%	0.00%	1.79%	1.61%	3.92%	0.00%	2.41%
feeling	10	6	3	19	7	1	1	9	3	2	0	5
	2.87%	3.14%	3.85%	3.08%	3.06%	2.08%	1.72%	2.69%	2.42%	1.96%	0.00%	2.01%
material	**46**	**61**	**36**	**143**	**44**	**19**	**27**	**90**	**39**	**38**	**11**	**88**
	13.22%	**31.94%**	**46.15%**	**23.18%**	**19.21%**	**39.58%**	**46.55%**	**26.87%**	**31.45%**	**37.25%**	**47.83%**	**35.34%**
perception	9	12	6	27	4	3	12	19	1	0	0	1
	2.59%	6.28%	7.69%	4.38%	1.75%	6.25%	20.69%	5.67%	0.81%	0.00%	0.00%	0.40%
perception/rel	0	0	0	0	0	0	0	0	1	2	0	3
	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.81%	1.96%	0.00%	1.20%
possessive/rel	12	7	2	21	19	7	5	31	5	6	2	13
	3.45%	3.66%	2.56%	3.40%	8.30%	14.58%	8.62%	9.25%	4.03%	5.88%	8.70%	5.22%
relational	23	22	5	50	32	4	5	41	45	11	3	59
	6.61%	11.52%	6.41%	8.10%	13.97%	8.33%	8.62%	12.24%	36.29%	10.78%	13.04%	23.69%
verbal	**95**	**25**	**6**	**126**	**14**	**6**	**2**	**22**	**9**	**30**	**3**	**42**
	27.30%	**13.09%**	**7.69%**	**20.42%**	**6.11%**	**12.50%**	**3.45%**	**6.57%**	**7.26%**	**29.41%**	**13.04**	**16.87%**
Total	348	191	78	617	229	48	58	335	124	102	23	249
	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Table 3.16. 1s, 2s, and *s/he* subjects by verb type, breakdown by tense (n=1,201).

	1s			1s Total	2s			2s Total	S/he			S/he Total
	Present	**Past**	**Modal**		**Present**	**Past**	**Modal**		**Present**	**Past**	**Modal**	
cognition	131	50	14	195	100	6	4	110	4	7	2	13
	67.18%	25.64%	7.18%	100.00%	90.91%	5.45%	3.64%	100.00%	30.77%	53.85%	15.38%	100.00%
corporeal	17	5	2	24	5		2	7	15	2	2	19
	70.83%	20.83%	8.33%	100.00%	71.43%	0.00%	28.57%	100.00%	78.95%	10.53%	10.53%	100.00%
existential	5	3	4	12	4	2	0	6	2	4	0	6
	41.67%	25.00%	33.33%	100.00%	66.67%	33.33%	0.00%	100.00%	33.33%	66.67%	0.00%	100.00%
feeling	10	6	3	19	7	1	1	9	3	2	0	5
	52.63%	31.58%	15.79%	100.00%	77.78%	11.11%	11.11%	100.00%	60.00%	40.00%	0.00%	100.00%
material	**46**	**61**	**36**	**143**	**44**	**19**	**27**	**90**	**39**	**38**	**11**	**88**
	32.17%	**42.66%**	**25.17%**	**100.00%**	**48.89%**	**21.11%**	**30.00%**	**100.00%**	**44.32%**	**43.18%**	**12.50%**	**100.00%**
perception	9	12	6	27	4	3	12	19	1	0	0	1
	33.33%	44.44%	22.22%	100.00%	21.05%	15.79%	63.16%	100.00%	100.00%	0.00%	0.00%	100.00%
perception/rel	0	0	0	0	0	0	0	0	1	2	0	3
	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	100.00%	33.33%	66.67%	0.00%	100.00%
possessive/rel	12	7	2	21	19	7	5	31	5	6	2	13
	57.14%	33.33%	9.52%	100.00%	61.29%	22.58%	16.13%	100.00%	38.46%	46.15%	15.38%	100.00%
relational	23	22	5	50	32	4	5	41	45	11	3	59
	46.00%	44.00%	10.00%	100.00%	78.05%	9.76%	12.20%	100.00%	76.27%	18.64%	5.08%	100.00%
verbal	95	25	6	126	14	6	2	22	9	30	3	42
	75.40%	19.84%	4.76%	100.00%	63.64%	27.27%	9.09%	100.00%	21.43%	71.43%	7.14%	100.00%
Total	348	191	78	617	229	48	58	335	124	102	23	249
	56.40%	30.96%	12.64%	100.00%	68.36%	14.33%	17.31%	100.00%	49.80%	40.96%	9.24%	100.00%

with verbs of cognition. Lexical items in the cognitive group (e.g. *think, know*) with 1s and 2s subjects have subjective function in discourse, and these usages do not extend to a subject that is not a speech act participant.

We also see in Table 3.15. that all three of these human subjects have large percentages of material verbs which are the most frequent predicates to occur with *s/he* subjects and the second most frequent type for both 1s and 2s. It is not surprising that material predicates—processes of 'doing' and 'happening'—should mainly occur with subjects with human referents, given that activities designated by these verbs are amenable to the presence of a human agent.

There appears, then, to be a strong association between material verbs and all subjects with human referents in English conversation, and, as discussed above, these propositional predicates tend to be fairly general in their informational value (recall that the only frequent lexical items in this group are the basic verbs *do, go,* and *get*.). The utterance in (25) below provides an example of a common *s/he*+material verb combination in the database.

(25) and **he did** a blood test, D10–37

There is, however, a difference in material predicates that occur with 1s and 2s subjects and those that occur with *s/he* subjects. For the speech act participants—1s and 2s subjects—we find proportionally more central modals (25 percent and 30 percent, respectively) than for *s/he* utterances (13 percent) (see Table 3.16.). This indicates that for third person human subjects, there tends to be less personalization of presentation than occurs in 1s and 2s material clauses. That is, *s/he*+material verb groupings may be said to be less mediated by speaker point of view than 1s+material verbs and 2s+material verbs based on there being fewer modal elements in these *s/he* predicates.

With respect to frequent verb types cooccurring with these three human subjects, both 1s and *s/he* utterances frequently contain verbal process predicates (see Table 3.15.), but the functions of the two types differ. As discussed in section 3.3.4., the ubiquitous *I mean* performs epistemic functions in discourse, whereas the most frequent usages of *s/he*+verbal type combinations are past tense quotatives. Again, distributionally, we see that even though humanness of subject triggers a higher frequency of verbal predicates (Table 3.13. shows there are very few of this type with 3s inanimate subjects), the formal characteristics of these combinations (e.g. verb lexeme and tense) and their functions vary by subject. Moreover, not all subjects with human refer-

ents necessarily cooccur with verbal process predicates; notice that there are relatively few of these types with 2s subjects (Table 3.15.) Notably, 1s+verbal type collocations carry subjective import, while the *s/he*+verbal predicate combinations seem to have a more propositional function—that of reporting the speech of another person. However, many of these *s/he*+verbal type utterances also personalize the speaker's contribution. In (26), *he just said*, introduces a direct quote, and this use fulfills a relatively objective, or reporting, function.

(26) **he just said** <Q bring it i=n, F8–5a

But in (27), the same quotative (*said*) preceding an indirect quotation has a slightly more interpreted reading; it is reminiscent of example (5) (*he thought I think…*) above in which assertion of the state or emotions of a third person are often mediated by the speaker. In (27), characterization of the subject's *not minding* is anchored by the speaker to the subject's having said it himself, an evidential function. Indirect quotations, then, are "an act of translation; the quoter translates the utterance of the original speaker into his or her own frame of reference" (Haiman 1998: 116). In (27), then, the actual 'saying' is not as concrete, or, perhaps, literal, as it is in direct quotations, such as (26). Such utterances are more imbued with the speaker's stance.

(27) and **he said** he didn't mind it. F8–34a

3.5.4.2. *Singular subjects with human referents and the distribution of tense*

Differences in animacy among third person singular subject types highlights the association of past tense with human subjects. Recall that frequency counts indicate that there are proportionally more past tense predicates with 3s subjects with human referents (*s/he* and some lexical NPs) than with 3s subjects with nonhuman referents (*it, that,* and some lexical NPs).

Table 3.17. outlines tense distribution for all 3s subjects by humanness of the subject referent. There is proportionally more past tense in predicates that have 3s subjects with human referents than in predicates with 3s subjects with nonhuman referents. For human referents, 50 percent of cooccurring finite verbs are present tense and 41 percent are past tense. In contrast, 78 percent of clauses with nonhuman subjects are present tense, and only 16 percent of these clauses are past tense.

As was the case in examining relationships between animacy of subject

and verb type above, it is worthwhile to contrast findings regarding tense and animacy in 3s human utterances with 1s and 2s utterances to try to tease apart the influence of humanness of subject from the function of person. Table 3.18. is parallel to Table 3.17., except the subjects with human referents in Table 3.18. include *all* singular subject utterances (1s, 2s, 3s), not just third person singular tokens.

Like Table 3.17., Table 3.18. reveals that there is proportionally more past tense in predicates with singular human subjects than in those with nonhuman subjects—58 percent present tense to 29 percent past tense for human types vs. 78 percent present tense to 16 percent past tense for nonhuman types. However, in comparing the distribution of tense in utterances with human subjects in Tables 3.17. and 3.18., we see that there is proportionally less past tense in the group of all singular subjects (29 percent) than there is in the group of third person singular subjects (41 percent). One suspects that the smaller proportion of past tense within the group that includes all singular subjects is due to the very low frequency of past tense

Table 3.17. Third person singular subjects by animacy, breakdown by tense (n=930).

3s subject	Present	Past	Modal	Total
human referent	**177**	**148**	32	357
	49.58%	**41.46%**	8.96%	100.00%
nonhuman referent	449	92	32	573
	78.36%	16.06%	5.58%	100.00%
Total	626	240	64	930
	67.31%	25.81%	6.88%	100.00%

Table 3.18. Singular subjects by animacy, breakdown by tense (n=1882).*

Singular subject	Present	Past	Modal	Total
human referent	741	379	170	1290
	57.44%	**29.38%**	13.18%	100.00%
nonhuman referent	462	100	30	592
	78.04%	16.89%	5.07%	100.00%
Total	1203	479	200	1882
	63.92%	25.45%	10.63%	100.00%

*subjects with human referents include *I, you, s/he*, and 3s lexical NPs with human referents.

predicates for 2s utterances (only 14 percent; see Table 3.16.).[10] We see, then, that while there is a general tendency for all subjects with human referents to cooccur with a greater proportion of past tense clauses than do subjects with nonhuman referents, there are differences due to person that affect the more general pattern.

Notice in Table 3.16. that both 1s and *s/he* utterances have a relatively large proportion of past tense predicates (31 percent and 41 percent, respectively). Table 3.15. shows that the largest proportion of past tense tokens for 1s, 2s, and *s/he* tokens occurs with material verbs. Material predicates account for 32 percent of 1s past tense clauses (n=61), 40 percent of 2s past tense utterances (n=19), and 37 percent of past tense *s/he* items (n=38). Moreover, for 1s+material verb combinations, there are more past tense predicates than present tense ones (Table 3.16)—an unusual switch in tense patterning in the database. For *s/he* clauses, there is an equal number of past and present tense tokens (n=38 and n=39, respectively). With respect to 1s+past tense material verb groupings, it was suggested above that these expressions are used by speakers to report on past personal experiences, and *s/he*+past tense material verb combinations appear to function similarly. In contrast, previous discussion of 2s+past tense material verbs suggests that items in this group assert little propositional information; most of these utterances are questions or subordinate clauses preceded by epistemic and evaluative expressions.

It appears, then, that past tense material predicates are used by speakers to report on past events for 1s and *s/he* utterances—a propositional function—even though the level of informativeness is often quite general, given that the most frequent verb lexemes in the material class are basic English verbs, such as *do* and *go*. However, this reporting does not occur in 2s utterances. In this case, there is similar patterning of first person singular and third person singular subjects (with human referents) as distinct from 2s utterances. The influence of humanness of subject on tense distribution, then, is mediated by the unique, interactive usages of second person singular utterances in conversation.

3.5.5 Summary of utterances with third person singular subjects

Third person singular subject provides an excellent example of how criteria for category membership can affect linguistic analyses. A global count of utterances in the database suggests that 3s is the most frequent subject in

English conversation—a pattern that is compatible with *referentially descriptive* modes of grammatical analysis that view language as primarily propositional in function, with third person singular utterances serving as the representative site for reporting. But 3s subject is a class that is composed of several subtypes, and these subtypes pattern differently with respect to animacy and cooccurring grammatical and lexical material. When English conversational usages are examined, one finds that even the most frequent 3s subtype, *s/he*, is not as frequent as are utterances occurring with the speech act participant subjects, *I* and *you*. Using this perspective (i.e. viewing 3s subtypes as subjects in their own right), first person singular, then, is the most frequent subject in the database. And because this is the prototypical site for expression of speaker point of view, its high frequency is compatible with the hypothesis that linguistic expressions most commonly found in English conversation are those that are subjective in nature.

In this section, associations were found between humanness of subject and occurrence of material verbs and past tense. The group of *s/he*+past tense predicates functions propositionally for speakers—to report on real-world events. However, in this corpus, there are only a total of 104 third person singular human subjects+past tense material verb tokens—11 percent of all 3s utterances. English speakers do in fact report and narrate in conversation, but, in interactive discourse, this is not as frequent a function as is, for example, the evaluative function of relational clauses, which account for 53 percent of 3s utterances (see chapter 4). In other words, 3s human subjects+past tense material predicates appear to be the prototypical site for reporting about the world, and they are less mediated by structural expression of speaker point of view than other subject-tense-verb type combinations, such as 1s, 2s, or 3s inanimate utterances. Given the relatively low frequency of these expressions in conversation, however, one suspects that they may not be as *prototypical* as they have been perceived to be in linguistic analysis.

And, finally, the issue of humanness of 3s subject referent invoked comparisons of the 3s subject *s/he* with the speech act participant subjects, *I* and *you*. In general, humanness of subject referent in English correlates with more past tense and with material and verbal predicates. However, these trends are influenced by whether or not the subject is a speech act participant and which speech act participant it is. For example, though both 1s and *s/he* commonly cooccur with verbal process predicates, the most common lexical items and tense uses vary for each subject. Similarly, though there is an association

between humanness of referent and past tense—and, indeed, there is a tendency for there to be a higher percentage of past tense with 1s and *s/he* subjects—2s utterances have the fewest tokens of past tense in the database.[11] In this case, the unique characteristics of 2s expressions in English conversation (e.g. numerous interrogatives and other questions, frequency of *you know*, high percentage of modal elements) preclude this subject's functioning very often as an agent, or even as much of a referential entity at all. Additionally, we find a very low frequency of verbs of cognition cooccurring with *s/he* subjects; in contrast, this is the most frequent verb type to appear with both 1s and 2s subjects. Here, again, humanness of subject does not on its own condition consistent distributions of verb type. As discussed in section 3.3.2., combinations of cognitive predicates with first person singular subjects function epistemically—a subjective function that rarely, if ever, is used with a 3s human subject.

These seemingly skewed distributions, then, in many cases can be explained by looking at the local combinatory patterns, and, what has been noted so far in this chapter is that the most frequent cooccurrences of subject, tense, and verb type tend to be those that allow speakers to personalize their contributions. The analyses in this section also identified a structurally more informative, or propositional, use of language in American English conversation, but these usages are not frequently occurring in English interactive discourse.

3.6. First person plural subjects

3.6.1 Introduction

Utterances with first person plural subjects are the least frequent type in the database (see Table 3.1., section 3.2.); in total, they account for only 3 percent of items in the corpus (n=66).[12] With respect to tense breakdown for 1p utterances, 53 percent of these clauses are present tense, 32 percent are past tense, and 15 percent of these predicates contain modal auxiliaries. Because this is such a small group of tokens, analyses presented in discussions in this section should be viewed as suggestive.

Table 3.19. displays the breakdown of 1p utterances by tense and verb type. Within the category 1p itself, 45 percent of utterances contain material verbs (30/66), and 15 percent (10/66) of *we* predicates are verbs of perception.

Table 3.19. First person plural subjects by tense and verb type (n=66).

	Present	Past	Modal	Total	% verb type
cognition	1	4	1	6	1.76%
	2.86%	19.05%	10.00%	9.09%	
corporeal	0	1	0	1	1.52%
	0.00%	4.76%	0.00%	1.52%	
existential	1	0	2	3	3.30%
	2.86%	0.00%	20.00%	4.55%	
feeling	1	1	0	2	4.44%
	2.86%	4.76%	0.00%	3.03%	
material	**15**	**9**	**6**	**30**	5.55%
	42.86%	**42.86%**	**60.00%**	**45.45%**	
perception	**6**	**4**	**0**	**10**	15.63%
	17.14%	**19.05%**	**0.00%**	**15.15%**	
possessive/rel	4	0	1	5	4.90%
	11.43%	0.00%	10.00%	7.58%	
relational	5	1	0	6	0.94%
	14.29%	4.76%	0.00%	9.09%	
verbal	2	1	0	3	1.23%
	5.71%	4.76%	0.00%	4.55%	
Total	35	21	10	66	3.04%
	100.00%	100.00%	100.00%	100.00%	

Perhaps because the class of 1p clauses is so small, they do not make up a significant percentage of any major verb type; they do, however, account for 16 percent of the small group of perception verbs (10/64).

3.6.2. Verbs of perception with 1p subjects

Verbs of perception are the second most frequent verb type occurring with first person plural subjects, and 1p subjects appear in 16 percent of these types of predicates in the corpus. Nine of the 10 items in the 1p+verb of perception group are usages of the verb *see* (see Table 3.20.), and there is one token of *hear*. In the majority of these cases, the subject pronoun, *we*, refers to individuals who are speech act participants (inclusive usage), as opposed to its referring to the speaker plus an outside person or group (exclusive usage), or the pronoun's functioning generically (e.g. *we're talking .. ba=king monster.*). Furthermore, all of these 1p+*see* predicates are semantically concrete in the sense that they express the more referential meaning of *see*—perceiving some-

Table 3.20. Verbs of perception by subject and tense (n=64).

		1s	2s	3s	1p	3p	Total
find	Past	1	1				2
	Modal	1	1				2
find Total		2	2				4
hear	Present	2	1				3
	Past	2		2	**1**		5
hear Total		4	1	2	**1**		8
look	Present			2			2
	Past	1		1			2
	Modal	2	1			1	4
look Total		3	1	3		1	8
notice	Past	3					3
notice Total		3					3
pick up	Present	1					1
pick up Total		1					1
see	**Present**	**6**	**3**		**6**		**15**
	Past	**4**	**2**		**3**	**1**	**10**
	Modal	**2**	**8**				**10**
***see* Total**		**12**	**13**		**9**	**1**	**35**
tell	Modal	1	2				3
tell Total		1	2				3
watch	Present			1			1
	Past	1					1
watch Total		1		1			2
Total		27	19	6	**10**	2	64
		42.19%	29.69%	9.38%	**15.63%**	3.13%	100.00%

thing visually—and not the more cognitive senses of 'understanding' or 'knowing'. For example, in (28), *see* (past tense) is used quite lexically.

(**28**) **we saw** her at ... Diamond Jim's. A8–50

Similarly, in (29), *see* (present tense, habitual meaning) also conveys the concrete sense of the word.

(**29**) **we don't see** him all winter. G6–13

Table 3.20 shows the distribution of verbs of perception in the entire corpus. Because 1p accounts for 16 percent of this group, it is of interest to compare these usages to the two subjects that occur even more frequently with this small class: 1s and 2s. Moreover, because in most cases (with the possible exception of some *generic* uses of the pronoun) at least one referent of *we* is

included in the discourse (the *I* portion), then it is useful to examine the ways in which utterances with *we* are similar to or different from those occurring with the other speech act participant subjects, *I* and *you*. This discussion will focus on the lexical item, *see*, since it accounts for over half (55 percent, or 35/64) of all uses of verbs of perception.

Table 3.20. indicates that *see* occurs with a 1s subject 12 times and with a 2s subject 13 times. For both subjects, these *see* predicates in English conversation can be propositional in character as are the *we+see* expressions illustrated above. This propositionality may be attributed to: (1) the pronoun's referring to an experiencing subject versus a speech act participant, or to (2) the expression of the verb's concrete meaning. For example, in (30), *I* refers to an entity/experiencer in the world (as opposed to its referring to the speaker as speaker in this utterance), and the verb is used lexically.

(**30**) **I saw** her at Scott's, A8–48

Similarly, in (31), *you* refers to a specific person in the discourse context (as opposed to making generic reference), and the verb is most focally used in the sense of visually perceiving something, even though *saw* here also allows the meaning 'know about'. Note, too, in (31), the presence of *I mean* in the same intonation unit as the *see* clause. As discussed for 2s+material verbs, this is a common way English speakers temper 2s assertions with *you* subjects that are used referentially.

(31) I mean **you saw** his other book <X Henria X>. C4–39b

However, not all 1s and 2s *+see* utterances are as propositional, or referential, with respect to the meaning of the verb and the specificity of the pronominal referent. For example, there are several *I+see* tokens for which the meaning of the verb is something on the order of 'understand'. In these utterances, as illustrated by (32), the referent of the subject is most readily interpreted as the speaker (i.e. a participant in the conversation), and not an acting subject separate from the discourse context. The presence of the modal, *can*, contributes to the expression of this more interactive usage.

(32) .. **I can see** how,
.. that would vary,
.. fr- from region to region, C9–48

In (33) below, the use of *I see* is even more pragmatically extended than the

usage in (32). Similar to example (32), *I* in (33) refers to the speaker, and, in both utterances, the verb conveys a more extended meaning of understanding, something on the order of 'acknowledging'. However, in (32), what the subject 'sees' or 'understands' is encoded in the complement clause *how that would vary from region to region* which follows *I see*. In contrast, in (33), there is no expressed object or complement of *I see*; in this usage, the *substance* that the speaker acknowledges precedes her utterance in the discourse and is, in fact, part of the discourse context. In this case, the complement of *I see* is the acknowledgment that Mar (and Stephan), who are visiting Mabel, will have to leave soon, even though this piece of information is not explicitly stated.

(33) MAR: we haven't gone to visit my grandmother so,
MABEL: .. oh,
MAR: (0) we need to go spend .. time with he=r,
… and then back to my parents,
MABEL: … (2.0) **I see**=. M8–3

In some usages, this type of *I see* may be so bleached of propositional substance than it can function as a backchannel, or minimal response—a cooperative device used in conversation to indicate participation, interest, or support. This situation is parallel to the case of *I don't know* in which the more pragmatic uses of the expression do not make reference to propositional material in the discourse (e.g. in the case of *I don't know* that the subject/speaker has insufficient knowledge of something) but rather to interaction among the participants (e.g. hedging assertions, mitigating disagreements), and these usages have grammaticized as markers of speaker change (Scheibman 2000). Moreover, in the case of *I don't know*, these pragmatic uses are more consistently articulated in a phonetically reduced form than are the more concrete, propositional ones.

For 2s+*see* combinations—as discussed for 2s+material verbs above—there is a tendency for these expressions to be referentially mediated by phenomena such as the presence of modal expressions and/or the use of *you* as generic, often with first person singular construal. [13] For example, in (34), the speaker is talking about her impressions while she was visiting Normandy; the *see* here is a concrete usage ('to perceive visually'); the subject pronoun, however, has generic expression with 1s meaning (i.e. the speaker/subject is projecting, or universalizing, her experiences).

(34) but **you still see** the marks in the la=nd. C10–40

In (35) below, the meaning of *see* is 'understand' and the referentiality of *you* is ambiguous between pointing to the addressee and being referentially underspecified in indexing a particular person, because it is part of the pragmatic collocation *you see*.

(35) STEPHAN: … my grandpa my gra%- --
Winston.
.. he couldn't remember.
MABEL: … (1.1) he couldn't.
well **you see**,
.. (h) I= … don't know .. how come, M10–6

In summary, then, first person plural+*see* utterances are descriptively informative in the sense that the subject tends to be referential and inclusive, and the verb is used in its most concrete sense ('to visually perceive something or someone'). This situation is in contrast to the two more frequent subjects that occur with this lexical item—*I* and *you*—which, along with expressing this more propositional meaning, also convey pragmatic usages that are consistent with the character of other 1s and 2s utterances. For first person singular clauses, there is a tendency for certain combinations of *I*+verb (e.g. verbs of cognition and perception) to cohere as a unit and fulfill subjective and interactive functions. These more pragmatic utterances express the viewpoint of *I* as *speaker*, not *I* as *subject*. For second person singular clauses, these collocations—as is true for 2s+material verb utterances—tend to be propositionally mediated by the presence of modals and the use of generic/subjective *you*. Based on this analysis, it appears that *we* utterances in conversation are more similar to 3s propositional clauses (e.g. those with human subjects and material predicates) than to utterances with speech act participant subjects.

3.6.3. Material verbs with 1p subjects

The largest verb class to cooccur with 1p subjects is the material group (n=30). Table 3.21. indicates that half of these clauses are present tense (n=15), 30 percent are past tense (n=9), and 20 percent contain central modals (n=6). With respect to the proportion between present and past tense for this subject with this verb type, *we*+material verb combinations pattern more like 2s+material verb utterances than like 1s+material verbs or *s/he*+material verbs. For

Table 3.21. First person plural subjects with material verbs by tense and referentiality (n=30).

	Present	Past	Modal	Total
SAP (inclusive)	**10**	3	**6**	19
	52.63%	15.79%	**31.58%**	100.00%
SAP/Entity (exclusive)	3	**6**	0	9
	33.33%	**66.67%**	0.00%	100.00%
Generic	2	0	0	2
	100.00%	0.00%	0.00%	100.00%
Total	15	9	6	30
	50.00%	30.00%	20.00%	100.00%

we+material verbs (and for 2s to an even greater extent), there is a higher proportion of present tense than past tense. Conversely, there is less past tense for both 2s and 1p subjects with material predicates than for 1s and *s/he* subjects with the same verb class. Recall that for 1s+material verbs there is more past than present tense, and for *s/he*+ material verbs, present and past tense tokens are essentially the same. Even though the number of *we*+material predicate tokens is small in this corpus (n=30), the distribution with respect to tense and referentiality is suggestive.

Because all human subjects have a frequent association with material predicates, it is useful to review some of the details of subject referentiality coding used in this study in order to include in the examination of *we* its similarities to and differences from other speech act participant subjects. The four referentiality codes for *we*, then, are summarized below. Examples may be found in chapter 2, section 2.4.16.

1. SAP (Speech Act Participant): referents of *we* who are all participants in the discourse, the *inclusive* usage.
2. SAP/Entity: *we* includes both a speech act participant and another person or several who are not present during the conversation, the *exclusive* usage.
3. SAP/Generic: *we* is construed as simultaneously referring to speech act participants and also to a larger, more universal, or generic, group. There are only 4 tokens of *we* with this code, and only 2 with material predicates.
4. Generic: *we* designates a class of referents as opposed to a specific referent. There are only 4 tokens of *we* with this code, and none appear with material verbs.

Note in Table 3.21. that two-thirds of the present tense *we*+material verb tokens (10/15) have subjects whose referents are all speech act participants (SAPs). Additionally, modals only occur in clauses with these inclusive subjects. This is consistent with the finding that there is a greater percentage of central modals in utterances with 1s and 2s subjects (SAP subjects) than those with 3s subjects (non-SAP subjects). Though not shown in Table 3.21., six of the present tense utterances with SAP referents contain intermediate function verbs (see example (36)); this means that 12 out of the 19 utterances in the group with SAP referents (=the inclusive usages), or 63 percent, are mediated by some kind of modal expression.

(36) **we gotta get** a picture. A18–38

On the other hand, two-thirds of *we*+material verb past tense tokens contain subjects whose referents include an SAP and someone who is not a discourse participant (the exclusive usage). In (37), the speaker is discussing her medical school coursework; in this case, *we* refers to the speaker and her classmates.

(37) **we did** a little bit in foundations, G15–41

Though the data set is quite small, there is a suggestive pattern that the subject *we* patterns more similarly to other speech act participants with material verbs—in particular, second person singular—when it refers to SAPs (inclusive usage) than when it refers to the speaker and to nonpresent others (SAP/Entity, the exclusive usage). There is a greater proportion of present to past tense with SAP *we* subjects with material verbs and also a higher percentage of modal expressions.

3.6.4. Summary of utterances with first person plural subjects

The most general pattern found for first person plural utterances in the corpus is that there are so few of them. Since in traditional grammatical descriptions, *we* is assumed to be the plural version of *I*, it is of interest to investigate the ways in which this subject patterns structurally like the other speech act participant subjects, *I* and *you*. On the surface, the two analyses in this section present conflicting findings; therefore, the results from this section are summarized below to facilitate discussion.

1. In combinations of *we+see* in conversation, the subject in the majority of cases (8 out of 10) refers to the speaker and another speech act participant

(inclusive usage), and the verb used in these constructions conveys its most concrete meaning of 'visually perceiving' and does not extend to a more cognitive sense of 'understanding' or 'acknowledging'. In comparison, combinations of *I+see* may express the extended, cognitive sense of 'understanding'. The 2s+*see* collocations, on the other hand, tend to be mediated by modal expressions, and in some cases, the subject pronoun expresses generic referentiality.

2. In combinations of *we*+material verbs—as with 2s utterances containing the same predicates—there is proportionally more present tense (or less past tense) than in the case of material clauses with 1s and *s/he* subjects. In particular, there is more present tense when *we* functions inclusively, whereas there is more past tense in this group when *we* is used exclusively.

How, then, can it be explained that *we* patterns propositionally when it occurs with *see* (e.g. no semantic extension to include the verb's cognitive meaning) whereas *we* patterns like 2s utterances when it occurs with material verbs (e.g. the presence of modal expressions, more present/less past tense [when the referents are all SAPs])? In all cases, *we* implicitly includes the point of view of the speaker. And when the other referents are discourse participants, those participants become open to interactive focus, as when the speaker negotiates assertions using the second person singular subject, *you*. In other words, when a speaker uses *we* to refer to herself and another participant, it becomes necessary to tacitly negotiate cooperation or agreement from that person, in similar ways that English speakers mediate 2s utterances in conversation. On the other hand, when *we* refers to the speaker and to people who are *not* present in the discourse, speakers do not have to negotiate consensus with another person. To varying degrees, speakers may assume, or impose, referential inclusion and consensus from the nonpresent others. In some sense, then, when *we* refers to SAPs, it functions as an interactive subject, like *you*, and when *we* includes the speaker and outside entities, it patterns like a third person subject.

Though speculative, this idea of *we* functioning both interactively and referentially helps to explain the disparate findings of the analyses of this subject pronoun with material verbs and verbs of perception (*see*). As with all the analyses presented in this study, the assumption is that general structural trends are built on the details of local activities in context. For material predicates, *we* utterances pattern like 2s utterances when the pronoun refers to speech act participants (inclusive usage). In this case, the speaker is medi-

ating her assertions to negotiate agreement for her point of view with the other SAP referent(s), just as we see for 2s clauses of this type. There is little unmediated reporting of events and activities within this group, and, concomitantly, less past tense usage. In contrast, there is the suggestion of more past tense and no modal mediation in 1p utterances when the nonspeaker referents of *we* are not in the discourse (exclusive usage); these utterances are structurally more like *s/he* utterances and the *1s*+past tense material verb group.

But when *we* cooccurs with *see*, recall that the subject refers to speech act participants and the verb is not used in its more extended cognitive sense (e.g. 'understand'). There are very few tokens of *we* occurring with verbs of cognition in the corpus (n= 6), and though there are three items with subjects that are referentially coded SAP, they do not function epistemically as do 1s+verb of cognition utterances (e.g. *we couldn't remember,*). However, it is not surprising that *we* would not occur with a cognitive process predicate (i.e. when *see* means 'understand'). It may be the case that in spontaneous discourse, it is interactively safer to say *we saw this* than to say *we understand this*, in that the former requires less presumption of consensus than the latter.

Though first person plural is an infrequent subject in the database, its usage appears to be complex. The data presented in this section suggest that its interactive nature influences its structure in discourse and that this quality is derived from the pragmatic consequences of its referentiality.

3.7. Third person plural subjects

3.7.1 Introduction

Utterances with third person plural subjects account for 10 percent of all items in the database (219/2172). Table 3.22. shows that 147 out of the 219 third person plural clauses (67 percent) are present tense, 50 verbs are past tense (23 percent), and 22 predicates with third person plural subjects contain modal auxiliaries (10 percent). With respect to the proportion of present to past tense found for other subjects, 3p utterances, though a much smaller group than 3s, pattern most similarly to 3s utterances (67 percent present tense and 26 percent past tense for the 3s group). Recall that 2s clauses also show a high proportion of present tense verbs (68 percent). However *you* utterances differ from 3s and 3p tokens in that 2s clauses show very little past tense (only 14

Table 3.22. Third person plural subjects by tense and verb type (n=219).

	Present	Past	Modal	Total	% verb type
cognition	8	2	4	14	4.12%
	5.44%	4.00%	18.18%	6.39%	
corporeal	1	2	0	3	4.55%
	0.68%	4.00%	0.00%	1.37%	
existential	7	0	1	8	8.79%
	4.76%	0.00%	4.55%	3.65%	
feeling	4	0	1	5	11.11%
	2.72%	0.00%	4.55%	2.28%	
material	**63**	**30**	**7**	**100**	18.48%
	42.86%	**60.00%**	**31.82%**	**45.66%**	
perception	0	1	1	2	3.13%
	0.00%	2.00%	4.55%	0.91%	
perception/rel	4	0	0	4	10.26%
	2.72%	0.00%	0.00%	1.83%	
possessive/rel	15	2	0	17	16.67%
	10.20%	4.00%	0.00%	7.76%	
relational	**33**	**9**	**3**	**45**	7.02%
	22.45%	**18.00%**	**13.64%**	**20.55%**	
verbal	12	4	5	21	8.64%
	8.16%	8.00%	22.73%	9.59%	
Total	147	50	22	219	10.08%
	100.00%	100.00%	100.00%	100.00%	

percent of the utterances) and contain many more predicates with central modals (17 percent) than either 3s (7 percent of predicates have modal auxiliaries) or 3p utterances (10 percent of predicates have modal auxiliaries).

Table 3.22. also indicates that the most frequently occurring verb types with 3p subjects are material (46 percent) and relational (21 percent)—the same verb types most commonly found with 3s subjects. Additionally, in material predicates with 3p subjects, there is the greatest proportion of present tense (63 percent) and the fewest modal auxiliaries (7 percent) than in utterances with material predicates occurring with any other subject in the corpus. Discussion of 3p+material verbs will be presented in section 3.7.3.

3.7.2 Third person plural subject types

Like third person singular, third person plural is composed of several subject types. Table 3.23. provides a breakdown of individual 3p subjects by verb type.

Table 3.23. Third person plural subject types by verb type (n=219).

	they	lexical NP	these	those	Total
cognition	14	0	0	0	14
	8.64%	0.00%	0.00%	0.00%	6.39%
corporeal	2	0	1	0	3
	1.23%	0.00%	33.33%	0.00%	1.37%
existential	4	4	0	0	8
	2.47%	7.69%	0.00%	0.00%	3.65%
feeling	3	2	0	0	5
	1.85%	3.85%	0.00%	0.00%	2.28%
material	86	14	0	0	100
	53.09%	26.92%	0.00%	0.00%	45.66%
perception	2	0	0	0	2
	1.23%	0.00%	0.00%	0.00%	0.91%
perception/rel	1	1	2	0	4
	0.62%	1.92%	66.67%	0.00%	1.83%
possessive/rel	17	0	0	0	17
	10.49%	0.00%	0.00%	0.00%	7.76%
relational	17	26	0	2	45
	10.49%	50.00%	0.00%	100.00%	20.55%
verbal	16	5	0	0	21
	9.88%	9.62%	0.00%	0.00%	9.59%
Total	162	52	3	2	219
	100.00%	100.00%	100.00%	100.00%	100.00%

The two most frequent subject types are *they* (162/219, or 74 percent) and lexical noun phrases (52/219, or 24 percent). *They* occurs in the corpus much more frequently than full NP subjects in this 3p group.[14] Fifty-three percent of *they* utterances contain material predicates, and the most frequent verb type to cooccur with lexical NP subjects is relational. Only 10 percent of *they* subjects occur with relational predicates, and, though 27 percent of utterances with lexical NP subjects cooccur with material verbs, the absolute number of lexical NP+material verbs is quite small—only 14 tokens.

3.7.3 Material verbs with 3p subjects

One hundred out of the total 219 third person plural utterances contain material predicates, and 86 percent of these have *they* subjects. Recall from analyses in previous sections that material verbs are a frequently occurring

Table 3.24. Material verbs by subject and tense (n=539).*

	Present	Past	Modal	Total
1s	46	61	36	143
	32.17%	42.66%	25.17%	100.00%
2s	44	19	27	90
	48.89%	21.11%	30.00%	100.00%
3s (all)	83	70	23	176
	47.16%	39.77%	13.07%	100.00%
s/he	39	38	11	88
	44.32%	43.18%	12.50%	100.00%
1p	15	9	6	30
	50.00%	30.00%	20.00%	100.00%
3p (all)	63	30	7	100
	63.00%	30.00%	7.00%	100.00%
they	53	27	6	86
	61.63%	31.40%	6.98%	100.00%

* Two tokens of 2p+material verbs in the corpus are not included in this count.

verb type with all subjects. Table 3.24. summarizes the occurrence of material predicates in the corpus by subject and tense. Discussion of distributional trends of subjects with material verbs and tense in this section is based on this table and findings from preceding portions of this chapter.

Note in Table 3.24. that for verbs of 'doing' and 'happening' (material verbs), utterances with speech act participant subjects (1s, 2s, 1p) show a greater proportion of modal auxiliaries than utterances with non-speech act participant subjects. These clauses are hedged when the speaker is the subject and mediated in assertion both when directed towards the addressee (2s) and when the subject includes both the speaker and another discourse participant (1p inclusive). In comparison, utterances with 3s subjects tend to be intermediate in the proportion of central modals found in material predicates (13 percent). And for third person plural utterances, there is the smallest proportion of modals in material predicates (7 percent) compared to other subjects.

Regarding the ratio of present to past tense for material predicates, Table 3.24. indicates that 3p tokens show the greatest proportion of present tense verbs of all subjects for this verb type (63 percent for all 3p utterances and 62 percent for *they* subjects). In contrast, for 1s+material verbs, there is more past tense than present, and for 3s human subjects (*s/he*), the difference in frequency between present and past tense clauses is negligible. It was sug-

Table 3.25. Third person plural material predicates by animacy, breakdown by tense (n=100).

	Present	Past	Modal	Total
human	53	29	5	87
	60.92%	33.33%	5.75%	100.00%
nonhuman	10	1	2	13
	76.92%	7.69%	15.38%	100.00%
Total	63	30	7	100
	63.00%	30.00%	7.00%	100.00%

gested in section 3.5.—with the exception of 2s utterances, which because of interactive concerns show little evidence of unmediated assertion of material processes in the past—that there is an association between human subjects and a greater proportion of past tense material predicates, or predicates of 'doing' in conversational English.

Of the 100 3p+material predicates in the corpus, 87 percent occur with subjects with human referents (see Table 3.25.). Consistent with the pattern found for 3s+material verb constructions discussed in section 3.5. and summarized in Table 3.26., there is a greater proportion of past tense predicates occurring with subjects with human referents—in particular for material verb types—than in clauses with subjects with nonhuman referents.[15] But in comparing Table 3.25. to 3.26., one can see that there are distributional differences in how animacy of subject and tense pattern for the two third person subject groups. For both third person singular and plural subjects, there is proportionally more past tense in clauses with subjects with human referents than in those with nonhuman ones. For 3p+material utterances with subjects with human referents, 61 percent of the predicates are present tense and 33 percent are past. On the other hand, for 3s+material verb combinations with human subjects, there are more past tense predicates than present tense (present: 39 percent, past: 50 percent). For both 3s and 3p material utterances with nonhuman subjects, there is proportionally less past tense in these clauses than in those with human subjects.

There appear, then, to be factors beyond humanness of the subject referent that may influence the distributional differences found in these two third person subject groups. That is, though the majority of 3p subjects with material verbs have human referents (87 percent), proportionally fewer of these predicates are past tense than was seen for the human subjects 1s and 3s.

Table 3.26. Third person singular material predicates by animacy, breakdown by tense (n=176).

	Present	Past	Modal	Total
human	43	54	12	109
	39.45%	49.54%	11.01%	100.00%
nonhuman	40	16	11	67
	59.70%	23.88%	16.42%	100.00%
Total	83	70	23	176
	47.16%	39.77%	13.07%	100.00%

Section 3.7.4. below explores the role of referentiality of third person plural subjects (*they* and lexical NP) with material predicates in discourse in an attempt to home in on these different patterns of distribution.

3.7.4 Referentiality of third person plural subjects

As described in chapter 2, subjects of utterances in the database were coded for referentiality. The three codes relevant to third person plural subjects with material verbs are *speech act participant (SAP)*, *entity*, and *generic*.

SAP labels subjects whose referents are participants in the discourse. *They* in (38) refers to the speaker's mother and father in-law who are themselves participants in the conversation. There are only four such tokens in the database, all uttered by the same speaker in the same stretch of discourse. These four items—because they make reference to specific individuals—are grouped with the *entity* class for the present analysis.

(38) **they** picked it up at a gar- -- .. trash dump, A18–24

Entity is the value assigned to subjects whose referents are specific people, animals, or things. Because the analysis in this section pertains to human subjects, *entity* labels a specific person in these cases. In the database, these subjects—which account for 25 percent of 3p+material verbs —are all tokens of *they* (i.e. there are no human lexical NP subjects coded as *entity*), and they make reference to people either known to the participants and/or previously talked about in the discourse. In (39), *they* refers to a couple all of the participants know, Trish and Scott, and the conversation concerns whether or not Trish has to work for a living given Scott's income.

(39) KENDRA: I guess Scott's making some good bucks.
MARCI: .. yeah but **they** bought like three cars in a row. A9–11

However, the most frequent referential code assigned to 3p subjects is *generic* (63 percent of all 3p subjects or 137/219). Sixty-seven percent of all 3p subjects with material verbs have this code (67/100), and 70 percent of human third person plural subjects with material verbs are *generic* (61/87). *Generic* characterizes subjects that designate a class of referents as opposed to specific referents. These subjects exhibit a range, or continuum, of generality of reference—from designating a particular group (e.g. all teachers in an elementary school or students in a medical school class) to referring to a more general (usually institutional) authority, such as the medical establishment (e.g. **they** *were going to have to take his tonsils out*). Both lexical NP (example (40)) and *they* subjects (example (41)) are coded *generic* in the database. For the most general referential usages, typically *they* is used.

(40) I'd never had,
.. known that **the French** ... compromised so terribly, C4–32b

(41) and in high school,
you know,
they teach about it, L16–40

3.7.5 Third person plural generic subjects

To explain there being proportionally less past tense in third person plural material utterances than was found in similar predicates with other (non–2s) human subjects, it is useful to examine the influence of subject referentiality on the distribution of tense and animacy (specifically, humanness) in these clauses. Table 3.27. displays tense and referentiality for third person plural human subjects with material verbs. Utterances with human subjects account for 87 percent of all 3p+material verb clauses (87/100). Note that 61 percent of these tokens are present tense and only 33 percent are past tense. In contrast, in Table 3.28.—which displays for third person singular subjects the same information presented in Table 3.27. for third person plural subjects—we find that 50 percent of these 3s utterances are past tense and 40 percent are present tense. In comparing these two tables, we also note that only 6 percent of the 3p group contain modal auxiliaries, while 11 percent of the 3s utterances do. While there are too few tokens to analyze this aspect of the data with much

Table 3.27. Third person plural human subjects with material verbs by referentiality, breakdown by tense (n=87).

	Present	Past	Modal	Total
SAP or entity	10	16	0	26
	38.46%	61.54%	0.00%	100.00%
generic	43	13	5	61
	70.49%	21.31%	8.20%	100.00%
Total	53	29	5	87
	60.92%	33.33%	5.74%	100.00%

Table 3.28. Third person singular human subjects with material verbs by referentiality, breakdown by tense (n=109).

	Present	Past	Modal	Total
SAP or entity	40	54	11	105
	38.09%	51.43%	10.48%	100.00%
generic	3	0	1	4
	75.00%	0.00%	25.00%	100.00%
Total	43	54	12	109
	39.45%	49.54%	11.01%	100.00%

certainty, one wonders whether the paucity of modals with 3p+material verbs is due to the frequency of generic subjects in this group. There may be less social need to mediate an assertion when the subject does not refer to a specific person.

With respect to referentiality of subject, Table 3.27. indicates that 61 of the 87 third person plural items (70 percent) are generic, while Table 3.28 shows that only 4 out of the 109 third person singular tokens (4 percent) are generic. Clearly, third person singular human subjects in English conversation are overwhelmingly used to refer to specific people. However, it is not true that third person plural subjects primarily refer to a collection of distinct individuals. Instead, material assertions about third person plural subjects are most frequently construed by speakers as being about an un(der)specified human group. In some sense, then, third person plural does not for the most part function as a plural subject in conversation, but as a generic one. Moreover, there are proportionally fewer past tense predicates with these generic subjects.

Table 3.27. also shows that 70 percent of 3p generic subjects are present tense while only 21 percent are past tense. However, in looking at the 3p human subjects with specific referentiality (i.e. those coded SAP or entity), one sees that 62 percent of these are past tense. This proportion of past to present is parallel to what is shown in Table 3.28. for 3s items: 51 percent of 3s SAP or entity subjects are past tense. This general pattern—the appearance of a relatively greater number of past tense predicates with human subjects—not only characterizes third person singular human subjects with material predicates, but also first person singular subjects with these same verbs. And this pattern exists as well for 3p human subjects which are *referentially specific*, but not for the majority of 3p human subjects whose referents are generically construed.

It was suggested above that the structural convergence of humanness of subject, past tense, and material predicates for 3s and 1s, and, now, for the minority of 3p subjects with specific referentiality, represents a more propositional function of discourse than is found elsewhere in conversation—one of reporting, in which speakers discuss past 'doings' and 'happenings' of particular individuals. For the majority of 3p utterances, however, the generic nature of these subjects does not lend itself to this kind of unmediated conveying of information on the part of speakers. Instead, in most 3p clauses, speaker point of view is integrated into the conveying of information; by using subjects with generic referentiality, speakers are generalizing to varying degrees (an evaluative function), not describing events in the world.

This lack of separation between the message and the speaker's take on it, seen in numerous contexts throughout this chapter, in the case of generic third person plural utterances, may be attributable to the invariably subjective character of the linguistic act of discursively delineating, or labeling, a group of people for purposes of predication. The subject in these clauses is in a sense itself a predication of the speaker's attitude, and the material predicate serves as an evaluation, not a description of an activity. For example, in (42), the material process designated by the verb *shave* is not what is in focus; that is, this is not an utterance reporting on an act of *shaving*. Rather, the subject, *most women*, coupled with the shared knowledge of the cultural import of women shaving themselves (and the presence of the modality adverb, *probably*) renders this utterance an expression of the speaker's point of view and not a description of an activity in the world.

(42) even in Berkeley most women probably don't,
… **most p- women probably shave their legs**,
but even so it's not,
.. I mean,
f- .. one is more tolerant of feminist vie=ws. C12–51

Recall that the most frequent 3p subject is *they*, and 67 percent of these tokens have generic reference. The use of generic *they*, then, is well grammaticized in American English interactive discourse. Unlike generic lexical NP subjects that have linguistic support to aid interpretation of their referential sense (e.g. the presence of lexical items in the subject NP), *they* is always phonologically *they*. Its referentiality, or indexicality, is typically expressed by both nonlinguistic and linguistic factors. In the conversational corpus, the referentiality of generic *they* hinges on the speaker's attitudes and her assumption that these attitudes are shared by the other participants (it is typically the case, that whoever *they* is, is presupposed). And, as is true for all linguistic expressions, both the preceding discourse context and the local predicate contribute to the interpretation of *they*. This easy and frequent use of generic *they*, however, does not depend on shared *information* among the contributors, but rather on shared (cultural) *attitudes* about the situating of authority and groups; and this understanding is tacit because it is so conventionalized. For example, in the discourse preceding (43), the speaker, Miles, was talking about his experience going to a club where people who were lambada dancing were engaging in sexual play with each other even though they hadn't known each other before dancing. Harold's last line in this exchange is the one in focus: *no wonder they forbid this dance*. The referent of *they* in this case is not retrievable from earlier discourse, nor is it readily apparent what group or institution *they* refers to. What is clear, however, is that the participants have no trouble construing and accepting the discursive creation of an authority that may or may not exist. It is also the case that in these most generic (the least discursively delineated) of *they* uses, the entire clause tends to be evaluative, in that it expresses the speaker's opinion. The predicate of such a generic subject, then, tends to function as a generalization, not as a description.

(43) MILES: it looks like,
.. these people aren't going home alo=ne tonight.
@@
HAROLD: … go [sh,
MILES: [but I just thought that] —

HAROLD: **no wonder] they forbid this dance.**
((J, M, P LAUGHING 7.8 SEC)) L25–6

Another distinction between generic 3p subjects and other human subjects in English (both speech act participants and third person singular) manifests in the distribution of verbs of cognition with third person plural subjects. Table 3.22. and 3.23. show that there are 14 tokens of 3p+verb of cognition combinations in the database. For all of these, the subject is *they*, and they are all instances of the verbs *think* and *know*. Recall that these verbs most frequently occur with 1s and 2s subjects. As pointed out earlier in this chapter, cognitive verbs with non-speech act participants (e.g. 3s subjects) tend to be hedged in some way by the speaker. The distribution of this verb type, then, is constrained to first person singular and, as a formulaic expression, to second person singular subjects. Cognitive predicates occurring with a third person subject tend to be mediated, as is the case for two of the 3p items as well. What is of interest here with respect to the combination of *they* with verbs of cognition is that in 12 of the 14 tokens, the subject is generic.[16] Examples include several variants of the conventionalized ***they don't know*** *what they're doing* and also utterances such as *what are* ***they thinking?*** or ***they think*** *of William the Conqueror.* These generic subjects allow speakers to make assertions about others in a way that specifically referential third person subjects do not. In usage, the meanings, or referents, of these generic *theys* are dependent upon the speaker's attitudes and the cultural/discursive sharing of these attitudes.

3.7.6 Summary of utterances with third person plural subjects

The third person plural subject in English, like third person singular, is composed of subtypes; the most frequent of these is *they* followed by lexical noun phrases. The most frequent verb type to occur with 3p subjects is the material class, but in contrast to other subjects with material predicates, 3p utterances most often occur in the present tense and have the fewest predicates with modal auxiliaries. Though 87 percent of 3p subjects with material verbs have human referents, these combinations are less frequently marked past tense than is the case for other human subjects with this verb class. It was suggested that the high frequency of generic 3p human subjects with material verbs (70 percent) contributes to the unique distributional pattern found for this particular group. That is, these 3p human generic subjects with material predicates are primarily used by speakers to generalize about, or evaluate, a

group of people who, to varying degrees of conventionality, are recognizable but viewed as *other* by the speaker and the discourse participants. These utterances are not typically used to report on past events and activities. Unlike third person singular utterances whose distributions of tense pattern according to animacy (humanness) of the subject, for third person plural clauses, the important factor seems to be referentiality. On the other hand, third person plural human subjects whose referentiality is specific (making reference to more than one specific individual) do pattern in tense like third person singular utterances.

A question that remains is what exactly is a (third person) plural subject in English conversation? The prototypical third person plural subject in interactive discourse is not a collection of specific entities; rather, it is (to varying degrees) a generally conceived group of people most frequently labeled with the pronoun, *they*. And it is not clear in some contexts whether the referent of *they* is human exactly, because the usage is metonymic in that it may refer to some accepted function or value of a particular institution or social construct and not necessarily to the specific people involved. This high frequency of third person plural generic subjects in the corpus has consequences for subjectivity vs. propositionality in American English conversation. Because these generic subjects are themselves grammaticized predications of speaker point of view (i.e. generalizations), the utterances that include them tend to be more evaluative than descriptive. In other words, in the most frequent third person plural clauses, speaker attitude is well integrated into the full grammatical structure of the utterance.

3.8. Summary

In examining the local details of frequent combinations of subject/person, verb type, and tense in a corpus of American English conversation, what one finds is that subjective phenomena—structural indications of speaker point of view—are integrated into the conveying of information throughout interactive discourse in a variety of ways. For first person and second person singular subjects we find highly frequent lexical collocations with subjective and pragmatic import, in particular with verbs of cognition and, for first person, with verbal process predicates as well (e.g. *I guess, I don't know, I mean, you know*). Moreover, text counts indicate that the high use of these conventionalized 1s

and 2s expressions are in large part responsible for the frequency of these categories as a whole (e.g. 1s, 2s, verbs of cognition), as well as for the individual verb lexemes (e.g. *know, think,* and *mean* are the second, fourth, and fifth most frequent lexical verbs, respectively, in the database). First person singular, second person singular, and first person plural inclusive utterances more frequently contain modal elements than do predicates with other subjects—a situation that formally displays the mediated nature of assertion characteristic of clauses (especially material predicates) that occur with speech act participant subjects.

For 1s utterances, subjectivity often manifests epistemically; 2s utterances, on the other hand, reflect an interactive, or empathetic, subjectivity on the part of the speaker towards the addressee in cooperative talk, in the sense that personalizing properties are employed to fulfill face-maintaining functions in 2s expressions. The relationship between subjective stance and interaction is complex; however, the high frequency of first person singular epistemic expressions coupled with the paucity of structures promoting direct assertion in second person singular utterances in the corpus suggests that in conversation, details of interaction are recruited for the negotiation of individual subjectivities, or the sharing of points of view.

This delicate play between expression of point of view and interactive focus—interaction being the hallmark of a conversational genre—is also seen in the patterning of first person plural utterances. Though English is not considered a language that marks an inclusive/exclusive distinction—and certainly, it doesn't morphologically—one sees a difference in lexical and grammatical patterning of *we* utterances. Notably, first person plural inclusive subjects have an interactive, or negotiated, function like 2s clauses; the subject in these utterances includes both the speaker's point of view and a tacitly (and likely, highly conventionalized) negotiated exchange with another speech act participant in order to subsume the other's point of view into the speaker's as *we*.

For utterances with third person subjects examined in this chapter, the findings are somewhat different than those described above for speech act participant subjects. Third person singular is the most frequent subject in the corpus because it is composed of several subtypes, and these subtypes pattern differently as regards their animacy and referentiality and also with respect to surrounding grammatical and lexical material. The distinctive distributions of utterances with 3s subject subtypes provide evidence that third person singu-

lar subject may not be a coherent formal category in English interactive discourse. With respect to third person plural subjects, however, there is really only one frequent subtype in this group—the pronoun *they*—which functions as a generic subject in discourse, and not a plural.

The analysis of third person nonrelational utterances (both singular and plural) pointed out an association in the corpus among third person human subjects, past tense, and material verbs. This convergence is especially robust for *s/he* clauses, but is weak for material predicates with third person human plural subjects. It was suggested that the convergence of humanness of subject, past tense, and verbs designating material processes represents the most propositional function of language in English conversation (i.e. the least subjectively mediated, the one that most closely approximates reporting) and only occurs in 3p utterances when the subject has specific, human referentiality. The high frequency of 3p generic subjects, then, separates these clauses functionally from their more referentially specific counterparts, in that the generic subjects mediate propositionality by anchoring the predicate to the generalizing subject, thus producing an evaluation and not a description. The structural contexts, then, that most frequently exhibit this concurrence of human animacy of subject, specific referentiality, past tense, and material predicates in the corpus are third person singular and first person singular. These are the forms that speakers use when they are engaging in more propositional modes of language use (e.g. narrative), and these utterances do not by any means constitute the majority of what occurs in interactive discourse in American English.

The dual goals of this research—the investigation of subjective structure in interactive discourse in English and the tacit methodological one of looking at local patterning of structure—have produced useful and complex results in this portion of the study. Structural indications of speaker point of view in interactive discourse, as they are delineated in the individual analyses presented in this chapter, manifest in diverse ways. Even within these complex and contextually variable linguistic details, what one sees overwhelmingly is that in English conversation, the expression of speaker stance and its impact on the structure of interaction is woven into—not separate from—the relating of descriptions of events and actors in the world.

Notes

1. Though second person plural utterances appear in several tables in chapters 3 and 4, the group is not discussed since there are very few of these subjects (n=5). They are *you guys, you both,* and one case of *you* where it was clear that the speaker was referring to a married couple.

2. See Kärkkäinen (1998) for an indepth analysis of *I think* in American English conversation.

3. Tokens of *I'm thinking* described above, though certainly subjective, are not included in this count since their use in the corpus is limited to one speaker.

4. There are also a greater proportion of modals for two less frequent verb types: existential and perception.

5. First person singular perception verbs also have more past tense tokens than present tense, and 1s relational predicates have an equal number of present and past tense usages.

6. Separate discussion of predicates with intermediate function verbs and those that express habitual meaning is for purposes of organization only—to provide as clear as possible a presentation of this complex verb class and its functions in discourse. In fact or in theory, there is nothing that precludes a predicate both from containing an IFV and expressing habitual meaning (e.g. in the utterance *I have to buy bags?*, the speaker is complaining that she will always have to have bags on hand to be able to use her new vacuum cleaner).

7. Note, too, that even though this utterance contains the specific lexical verb *play*, it is part of the lexicalized predicate *play…game*.

8. Additionally, 29 percent of verbal process types (79/245) have third person singular subjects, and a little over half of these are past tense quotatives (e.g. *she said, he was just talking,*). This group will be discussed in section 3.5.4.1.

9. This and subsequent discussions are based on presentation of data in both Table 3.15. and Table 3.16. The layout and frequency counts are the same in both tables. The difference is that in Table 3.15., percentages within person are given by verb type; in Table 3.16., percentages within person are given by tense.

10. Note, too, that there is a higher proportion of modals in predicates with all human subjects than in predicates with 3s human subjects (13 percent and 9 percent, respectively). This is likely due to the greater number of modal auxiliaries cooccurring with 1s and 2s than with 3s.

11. In fact, 2s has less past tense (14 percent) than 3s inanimate subjects (17 percent). This distributional patterning correlates with there being very few descriptive, or propositionally assertive, utterances in both groups.

12. Recall that five items are coded second person plural subjects, but, because there are so few tokens, this group is not discussed in this study.

13. Note in Table 3.20. that 8 of the 13 *you+see* items contain central modals.

14. The infrequently occurring 3p subject types *these* and *those* are excluded from subsequent discussions.

15. With the notable exception, of course, of 2s utterances.

16. The two utterances in which *they* refers to specific people are *they must know each other* and *surely they must know each other* (both from the same conversation and speaker).

Chapter 4

The evaluative character of relational clauses

4.1. Introduction

4.1.1 Description of relational utterances

The focus of this chapter is the group of clauses containing the most frequent verb type in the database—utterances with *relational* predicates. *Relational* processes are processes of being. Halliday (1994: 119) explains that "[i]n relational clauses, there are two parts to the 'being': *something* is being said to 'be' *something else*. In other words, a relation is being set up between two separate entities. [italics added]" In English, relational processes may be schematically represented as: '*x* is *a*', '*x* is at *a*' (where 'is at' represents 'is at, in, on, about', etc.), or '*x* has *a*' (Halliday 1994: 119).[1]

Though 96 percent of the relational predicates in the database contain the copula, *be*, the relational group is not completely coextensive with the category copular clause. For example, predicates other than *be* are regarded as relational (e.g. an*d you* ***get*** *really good, this tractor's* ***going*** *sour on him*.), and some clauses with *be* predicates may not be classified as relational (e.g. the existential: *there's the mail*.).

4.1.2 Relational utterances and linguistic subjectivity

Relational clauses are the most frequently occurring predicate types in the corpus, and they account for 30 percent of the data tokens in this study (641 of 2,172). Thompson and Hopper (2001) also report that 37 percent of one participant clauses in their database of American English conversational utterances are copular.[2] Given the role of discourse frequency in the conventionalization of linguistic expression, examination of the function and form of these ubiquitous expressions is relevant to the elaboration of theories and practices that view structure as rooted in usage. Moreover, descriptions of how these relational clauses express speaker point of view in interactive dis-

course in English (i.e. what kinds of relations do speakers typically express?) are important to the general hypotheses underlying this study concerning the expectation that the most frequent structures found in conversation should be those that participate in conveying the speaker's stance.

A general point regarding the subjectivity of relational utterances proposed in this chapter is that the prototypical function of these clauses in conversation is evaluative and not descriptive. This characterization is based on analysis of local distributional details and classifications of the subjects of these clauses (subject type, referentiality) and the elements found in the predicate complements (adjective and semantic adjective type, nominal and semantic noun class, adverbials and adverbial type). To illustrate, in isolation (1) below might be interpreted as a descriptive utterance since *metal* describes a physical property of the item designated by *it* (=a vacuum cleaner). However, the utterance's context and structure demonstrate its evaluative quality. The speaker of (1) is a mother who was trying to convince her less than enthusiastic daughter that the vacuum cleaner she bought for her (the daughter) for her birthday was high quality. In other words, it is an evaluation. That the vacuum cleaner is presented as *all metal* is not primarily a statement about its physical characteristics; rather, it is a positive assessment on the part of the speaker.

(1) it's all metal, A19–5

Distinctions (drawn) between description and evaluation can sometimes be obscure; however, in discussing performative and constative (statement) speech acts, Austin (1997: 134) remarks "Surely to state is every bit as much to perform an illocutionary act as, say, to warn or to pronounce".

Section 4.1.4. below presents general distributions of relational clauses in the corpus, but by way of preview, it is interesting to note that these predicate types do not equally occur with all subjects; rather, the great majority of relational utterances have third person singular subjects (78 percent), and 79 percent of these 3s relational predicates are present tense. Furthermore, over half of all verbs occurring with 3s subjects are relational (53 percent), and the most frequent 3s subjects that combine with these predicates are those with nonhuman referents (e.g. *it, that*). In other words, the bulk of relational clauses do not occur with first person singular or second person singular subjects which are the more transparent sites of expression of subjectivity (see chapter 3 and Benveniste 1971). Instead, most relational predicates appear

with third person singular inanimate subjects. This finding leads to consideration of the functional and referential aspects of the English subject in interactive discourse, in particular, how third person singular inanimate subjects with relational predicates contribute to the expression of speakers' discursively-changing points of view. These analyses touch on claims put forth by Kuno and Kaburaki (1977) that the subject is the locus of the speaker's empathy, and Stein (1995: 130) who suggests "that the subject position is the bleached, grammaticalised . . . position of the emanational locus of subjective viewing".

4.1.3 Analysis of relational clauses

The analysis of relational utterances in interactive discourse is a linguistically eclectic task. Consistent with the examination of subject and lexical verb type in chapter 3, the organization of discussion in this chapter proceeds by linguistic form, and analyses focus on semantic and pragmatic aspects of utterances and their elements.

English relational verbs overwhelmingly appear in predicate adjective, predicate nominal, and predicate oblique clauses, illustrated in (2), (3), and (4), respectively. Relational predicates appearing without complements (without predicate adjectives, predicate nominals, or predicate obliques) are tagged *other*; examples of these appear in (5).

(2) Predicate adjective clause
that's terrible. D13–25

(3) Predicate nominal clause
this is .. a raging bureaucracy, R13–6

(4) Predicate oblique clause
he's still out there. M7–29

(5) Other
but their hips are like, L19–43
it must be. C5–31

The most important coding categories for the analyses in this chapter are: (1) subject and referentiality of subject; (2) adjectives and adjective type for predicate adjectives; (3) noun and noun type for predicate nominals; and (4) adverbial expressions and adverbial type for oblique utterances.

4.1.4 Global distributional patterns of relational clauses

There are 641 relational tokens in the database which comprise 30 percent of the total number of utterances. Table 4.1. summarizes their distribution by subject and tense. Recall from chapter 3 that 64 percent of all utterances in the corpus are present tense, 25 percent are past tense, and 11 percent contain central modals. In contrast, for relational clauses alone, there is a higher percentage of present tense (76 percent) and proportionally fewer past tense predicates (18 percent) and clauses with modal auxiliaries (6 percent) than for the total group of utterances.

Table 4.1. indicates that 3s subjects are the most frequent subjects to occur with relational predicates. Seventy-eight percent of relational utterances have third person singular subjects. Note, too, in Table 4.1., that the majority of

Table 4.1. Relational utterances by subject and tense (n=641).

	Present	**Past**	**Modal**	**Total**	**% Relational***
1s	23	22	5	50	7.80%
	46.00%	44.00%	10.00%	100.00%	
2s	32	4	5	41	6.39%
	78.05%	9.76%	12.20%	100.00%	
3s (total)	393	81	23	497	77.54%
	79.07%	16.30%	4.63%	100.00%	
3s (human)	72	26	6	104	16.22%
	69.23%	25.00%	5.77%	100.00%	
3s (nonhuman)	321	55	17	393	61.31%
	81.68%	13.99%	4.33%	100.00%	
1p	5	1	0	6	0.94%
	83.33%	16.67%	0.00%	100.00%	
2p	1	0	1	2	0.31%
	50.00%	0.00%	50.00%	100.00%	
3p (total)	33	9	3	45	7.02%
	73.33%	20.00%	6.67%	100.00%	
3p (human)	16	8	2	26	4.06%
	61.54%	30.77%	7.69%	100.00%	
3p (nonhuman)	17	1	1	19	2.96%
	89.48%	5.26%	5.26%	100.00%	
Total	487	117	37	641	100.00%
Percent	75.98%	18.25%	5.77%	100.00%	

*The *% relational* column shows the distribution of relational clauses by subject (e.g. 6.39% of relational clauses have 2s subjects; 77.54% of relational clauses have 3s subjects).

Table 4.2. Percent of subjects that occur with relational clauses.*

	% subject
1s	8.10%
2s	12.24%
3s (total)	53.00%
3s (human)	*30.77%*
3s (nonhuman)	***66.39%***
1p	9.10%
2p	40.00%
3p (total)	20.55%
3p (human)	*14.94%*
3p (nonhuman)	*42.22%*
Total	29.51%

*This table displays the percentage of each subject that occurs with relational clauses (e.g. 8.10 % of all 1s utterances have relational predicates; the total percentage of subjects in the database that appear with relational clauses is 29.51%).

these 3s subjects of relational clauses have nonhuman referents (393/497 or 79 percent of 3s relational, and 393/641 or 61 percent of all relational clauses).

Recall from chapter 3 that in the entire database the most frequently occurring subjects are, in order, 3s, 1s, and 2s. However, the right-hand most column of Table 4.1. indicates that for relational predicates alone, 1s and 2s subjects only appear with 8 percent and 6 percent of this verb type, respectively. Table 4.2 shows that relational verbs do not frequently appear with speech act participant subjects. Only 8 percent of all 1s subjects cooccur with relational clauses, 12 percent of 2s subjects combine with this verb type, and 9 percent of 1p clauses are relational. This is in contrast to 3s nonhuman subjects for which relational predicates occur with 66 percent of subjects in this group.

Table 4.3. summarizes the distribution of relational utterances by subject and complement type. Relational predicates include a complement (one of the two *entities* involved in the relation, the other being the subject) with the exception of those coded *other* illustrated in (5) above. The distribution of relational clauses by subject and complement type displayed in Table 4.3. reveals a few general frequency patterns. For the entire relational group, there is almost an equal number of predicate adjectives (41 percent) and predicate nominals (43 percent); on the other hand, only 12 percent of relational clauses are predicate oblique.

Table 4.3. Relational utterances by subject and predicate type (n=641).

	Adjective	Nominal	Oblique	Other	Total
1s	25	11	11	3	50
	50.00%	22.00%	22.00%	6.00%	100.00%
2s	27	10	2	2	41
	65.85%	24.39%	4.88%	4.88%	100.00%
3s (total)	182	247	53	15	497
	36.62%	49.70%	10.66%	3.02%	100.00%
3s (human)	44	40	18	2	104
	42.31%	38.46%	17.31%	1.92%	100.00%
3s (nonhuman)	138	207	35	13	393
	35.11%	52.67%	8.91%	3.31%	100.00%
1p	2	0	4	0	6
	33.33%	0.00%	66.67%	0.00%	100.00%
2p	1	1	0	0	2
	50.00%	50.00%	0.00%	0.00%	100.00%
3p (total)	26	8	8	3	45
	57.78%	17.78%	17.78%	6.67%	100.00%
3p (human)	14	5	6	1	26
	53.85%	19.23%	23.08%	3.85%	100.00%
3p (nonhuman)	12	3	2	2	19
	63.16%	15.79%	10.53%	10.53%	100.00%
Total	263	277	78	23	641
Percent	41.03%	43.21%	12.17%	3.59%	100.00%

There is, however, a general trend in these data for there to be proportionally more predicate adjective complements than the other complement types for all subjects with the exception of third person singular subjects with nonhuman referents. For this latter group there are proportionally more predicate nominals.[3]

Eighty-five percent of the predicate oblique class contain complements of space or time (e.g. *Harold's in the room, it's been a year,)*, and there is the greatest proportion of these types occurring with subjects with human referents with the exception of second person singular, for which there are only two tokens. These two 2s utterances appear in (6) below; one of these is a question, and the other is an assertion. The scarcity of clauses that situate an addressee in place or time is not surprising given the general finding discussed in chapter 3 that speakers tend not to assert propositional information about a 2s participant. In (6)(b), note the presence of a pause between *you='ve* and *been* which makes this assertion less direct.

(6) (a) you were the=re before? L25–19
(b) you='ve .. been in Paris. C9–39

4.2. Referentiality

4.2.1 Referentiality, coding, and indexicality

Analysis of referentiality of linguistic elements takes for granted the idea that linguistic material is mapped onto tangible and discursive *entities*. And, indeed, it is the case that for many subject noun phrases in English—in particular the human third person pronouns, *she* and *he*—assignment of reference is relatively straightforward. This is because individuals are easier to construe, track, and delineate than are events or states. Related to the more general issue concerning expression of subjectivity vs. unmediated information in (analysis of) discourse, the concept of referentiality itself is theoretically embedded in a propositional view of language in the sense that linguistic form (referring expressions) is presumed to encode and transfer ideas and events (referents) from one person to another without much focus on the activities of speakers and the influence of context (Reddy 1993). The data examined here force a reconsideration of these assumptions, particularly with regard to referentiality in conversation.

Identification of entities, events, and relations in the world or in the discourse (context) that are designated by linguistic expressions requires categorization, or reification, on the part of the analyst. Particularly in situations where the subject noun phrase does not point to an individual entity, it is often difficult to characterize the referent of the subject. One area of indeterminacy occurs in coding a referent as an *event-state*; in these situations it is not clear how much or what aspects of the event or state should be linguistically present in the conversation to warrant this referential label (see also chapter 2, section 2.4.16.). In the following exchange, the participants are discussing the social unacceptability of being a feminist in France; according to the speaker, it is better in that country to proclaim oneself a communist than a feminist. The referent of the subject *it* in the utterance *it's rather sad* in (7) is what is in focus here.

(7) yeah.
it's just that,
when you had this immediate image of <X bra-burning X>,
and that's it,
<X and isn't X> going any further.
it's rather sad. C12–30

It in this example is coded *event-state* in the database. The state designated by *it* is construed (summarized, reified) as a belief on the part of the speaker that feminism has not advanced much beyond stereotyped media images of bra-burning. Additionally, this state/referent has no existence apart from its having been uttered by the speaker. What *it* refers to in this utterance is not "something" (except as analytically construed and coded). It is an on-the-fly, constructed designation and evaluation on the part of the speaker. *It* indexes a "state", or better, an evaluation, that is coextensive with its uttering; it is neither making reference to a reality outside the discourse context, nor explicitly to aspects of the discourse itself.

In a discussion of discourse deixis in a cross linguistic study of demonstratives, Himmelmann (1996: 224) comments on the difficulties inherent in determining a referent for an event, and he suggests that the problem "is due to the fact that no referent exists in advance to which one may point. Instead, the referent is first created at the very moment when this use occurs".[4] Discussing the importance of context to the identification of referents, Hanks (1990: 34) writes: "[I]n general, the descriptive information encoded in a linguistic expression is insufficient in itself to uniquely identify a referent". One suspects that the frequent lack of fitness between referential categories (e.g. *event-state, discourse*) and linguistic elements in interactive discourse may in large part be due to the propositional, or referential, bias in our conceptions of language and its analysis (Hopper 1997; Silverstein 1976). In other words, our analytical expectations contribute to a characterization of referents as discrete, and in English conversation, this does not fit well with the majority of referents of 3s inanimate subjects.

In most cases, referents of subject noun phrases in English conversation are not naturally substantive nor in any transparent sense existing prior to their being pointed to by live participants within a rich discourse context. Relevant to the idea that referring expressions are embedded in a process of contextualized and interactive meaning-construction—and are not themselves containers for meaning—is the notion of indexicality (Silverstein 1976).

Most often indexical expressions are equated with deixis (Levinson 1983), but Tao (1996: 488) provides an example of a more expansive characterization by noting that indexicality "refers to the phenomenon that the interpretation of an expression (a word or a phrase) is not based on some pre-fixed meaning but is relative to or shaped by the local discourse context". To return to (7) above, then, *it* can be interpreted as indexing (cooccurring with) a real-time evaluation on the part of the speaker and not signifying a preexisting meaning package. In this sense, then, the subject functions as the anchor for the speaker's empathetic construal, or subjective focus (Kuno & Kaburaki 1977; Stein 1995).

4.2.2 Relationships between referentiality and subjectivity in nonrelational verb types

The impact of referentiality of subject—both on distributional patterns and on construction of subjective expression in English conversation—surfaced in a variety of ways in chapter 3. For example, we saw a reduction in referential specificity in first person singular subjects occurring in routinized expressions such as *I think*, *I don't know*, and *I mean*, as well as in the second person singular subject in the ubiquitous collocation, *you know*. In these frequently occurring expressions, *I* and *you* do not refer to acting subjects in the world. As part of very common constructions with interactive import, speakers are not specifically referring to themselves (with *I*) nor to addressees (with *you*) in these usages. Instead, the entire routinized expression has pragmatic function with attendant reduction in specificity of the pronominal referent (see also Bybee & Scheibman 1999; Scheibman 2000).

It was suggested in chapter 3 that the split distributional patterning of verb types and tense with *we* subjects is due to the dual referentiality of the first person plural pronoun. Recall that when *we* is referentially inclusive, it patterns similarly to *you* utterances, and when it is referentially exclusive, these subject-predicate combinations function more propositionally. It was also discussed in chapter 3 that the most frequent third person plural subjects in the corpus are those that are referentially generic (primarily *they*), and that speakers use these utterances to evaluate and generalize in conversation. These findings demonstrate that referentiality of subject is an important element in the distributional and expressive patterns found in English interactive discourse.

4.2.3 Referentiality of third person singular inanimate subjects of relational clauses

The focus of referentiality in this chapter is on third person singular subjects with nonhuman referents. Recall that 79 percent of 3s relational utterances occur with subjects with nonhuman referents. The high frequency of this subject-predicate combination makes this an important group to investigate as part of this study on associations between expressive trends and structural details in English conversation.

For 3s relational clauses, the most commonly occurring subject types are *it* (155/497 or 32 percent) and *that* (122/497 or 26 percent). It is well known that the majority of subjects in English conversation are pronouns; 90 percent of subjects in utterances in the full database used for the present study are pronominal. Chafe (1994) reports that 81 percent of the subjects in his corpus of English conversational utterances are *given*, and 98 percent of these *given* subjects are pronouns. He suggests that subjects in English function as starting points for additional information (MacWhinney 1977) and are governed by the *light subject constraint*—the notion that English subjects tend to be given or accessible or, less frequently, of trivial importance (Chafe 1994: 82–92). Chafe (1994: 91) argues that in English "heavy referents do not occur in conversational language as starting points". However, these analyses of the formal and cognitive properties of subjects in English conversational utterances do not extend to a discussion of what a conversational referent might actually look like (i.e. Chafe's discussion assumes a traditionally discrete character of the referent). Therefore, the present study investigates the functional correlates of some of these pronominal elements that serve as starting points in interactive discourse.[5]

The claim that English subjects are starting points is a productive notion, especially for the description of the majority of relational clauses found in this conversational corpus. These constructions do not have subjects that are semantic agents. Given that the typical subject in these utterances (e.g. *it, that*) is not a sentient actor that can serve as a structural anchor for elaboration of an activity by a propositional predicate, then what motivates these starting points for discourse participants? And with respect to the present study, how do speakers use these frequent 3s inanimate subjects in the expression of point of view? In general, then, a central question to ask is what

is the relationship between subject and subjectivity in English as revealed by these relational utterances?

These inanimate starting points (grammatical subjects)—whether referentially concrete or analytically diffuse—are constructed from speakers' evaluations (see (7) in section 4.2.1.). What is common to these utterances is that they are linguistically structured around a third person point of view, not the speaker's. Analyses in subsequent sections of this chapter demonstrate that lexical material in these relational predicates (e.g. elements of adjectival and nominal complements) is frequently ambiguous in characterizing properties of a non-speech act participant entity and evaluating or discursively delineating that entity. This study proposes that what speakers are doing with these relational utterances is projecting their evaluations and points of view onto external entities and events in the world using third person inanimate subjects, and endowing them with properties and characteristics (expressed in the predicate complements), the substance of which is subjective, or evaluative. In other words, these inanimate subjects of relational clauses function in conversation to locate viewpoint away from the speaker; in essence they elevate evaluation and attitude to the level of description. Functionally, they reflect a kind of *covert subjectivity* in that the speaker is projecting her subjective stance outward. And because of the absence of a first person subject in these utterances, these forms are not overtly marked as speaker point of view.

4.3. Relational predicates with adjectival complements

4.3.1 Distribution of predicate adjective constructions

There are 263 relational clauses with adjectival predicates in the database. Sixty-nine percent of these have third person singular subjects (182/263), and 52 percent occur with 3s subjects with inanimate referents (138/263). Table 4.4. summarizes the distribution of these clauses by subject and semantic adjective type. Recall from chapter 2 that lexical adjectives in these complements are classified into semantic classes (e.g. *age, difficulty, dimension*). These adjective type codes are from Dixon (1977; 1991). Refer to Table 2.6. and section 2.4.30. in chapter 2 for examples of these classes.

With respect to the distribution of predicate adjectives by subject, Table 4.4. reiterates the general pattern found for all relational predicates intro-

Table 4.4. Relational utterances with adjectival predicates by subject (n=263).

	1s	2s	3s	1p	2p	3p	Total
age	0	2	2	0	0	0	4
	0.00%	7.41%	1.10%	0.00%	0.00%	0.00%	1.52%
difficulty	0	0	7	0	0	0	7
	0.00%	0.00%	3.85%	0.00%	0.00%	0.00%	2.66%
dimension	0	1	9	0	1	2	13
	0.00%	3.70%	4.95%	0.00%	100.00%	7.69%	4.94%
function	0	0	2	0	0	0	2
	0.00%	0.00%	1.10%	0.00%	0.00%	0.00%	0.76%
human propensity	**13**	**9**	**29**	**1**	**0**	**13**	**65**
	52.00%	**33.33%**	**15.93%**	**50.00%**	**0.00%**	**50.00%**	**24.71%**
physical property	**7**	**6**	**45**	**1**	**0**	**6**	**65**
	28.00%	**22.22%**	**24.73%**	**50.00%**	**0.00%**	**23.08%**	**24.71%**
qualification	3	3	17	0	0	0	23
	12.00%	11.11%	9.34%	0.00%	0.00%	0.00%	8.75%
similarity	1		5	0	0	2	8
	4.00%	0.00%	2.75%	0.00%	0.00%	7.69%	3.04%
value	**0**	**3**	**64**	**0**	**0**	**3**	**70**
	0.00%	**11.11%**	**35.16%**	**0.00%**	**0.00%**	**11.54%**	**26.62%**
other	1	1	1	0	0	0	3
	4.00%	3.70%	0.55%	0.00%	0.00%	0.00%	1.14%
indecipherable	0	1	0	0	0	0	1
	0.00%	3.70%	0.00%	0.00%	0.00%	0.00%	0.38%
none	0	1	1	0	0	0	2
	0.00%	3.70%	0.55%	0.00%	0.00%	0.00%	0.76%
Total	25	27	182	2	1	26	263
	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

duced in Tables 4.1. and 4.3.: compared to 3s subjects, there are few predicate adjective clauses with the speech act participant subjects, 1s and 2s. The most frequent adjective types found in this group of predicates as a whole are *value* (e.g. *good, horrible, great*), *human propensity* (e.g. *happy, famous, interested*), and *physical property* (e.g. *dirty, dry, blue, sick*). Notably, 91 percent of the predicates containing *value* adjectives (64/70) are third person singular; there are none of these adjective types with 1s subjects and 3 tokens each for 2s and 3p subjects.

In order to refine the presentation of the distribution of 3s predicate adjective clauses, Table 4.5. presents frequencies of adjectival complements for 3s subjects by humanness of the subject referent; included as a point of comparison are the two singular speech act participant subjects, 1s and 2s. There are more tokens of 3s human forms than there are 1s or 2s in this group,

Table 4.5. Relational utterances with adjectival predicates and singular subjects by adjective type and animacy of subject (n=138).

	Human			**Nonhuman**
	1s	2s	3s	3s
age	0	2	1	1
	0.00%	7.41%	2.27%	0.72%
difficulty	0	0	0	7
	0.00%	0.00%	0.00%	5.07%
dimension	0	1	0	9
	0.00%	3.70%	0.00%	6.52%
function	0	0	0	2
	0.00%	0.00%	0.00%	1.45%
human propensity	**13**	**9**	**21**	**8**
	52.00%	**33.33%**	**47.73%**	**5.80%**
physical property	7	6	16	29
	28.00%	22.22%	36.36%	21.01%
qualification	3	3	0	17
	12.00%	11.11%	0.00%	12.32%
similarity	1	0	1	4
	4.00%	0.00%	2.27%	2.90%
value	**0**	**3**	**4**	**60**
	0.00%	**11.11%**	**9.09%**	**43.48%**
other	1	1	0	1
	4.00%	3.70%	0.00%	0.72%
indecipherable	0	1	0	0
	0.00%	3.70%	0.00%	0.00%
none	0	1	1	0
	0.00%	3.70%	2.27%	0.00%
Total	25	27	44	138
	100.00%	100.00%	100.00%	100.00%

and these 3s utterances pattern similarly to the other two human subjects. That is, the bulk of the adjectives in these predicates with singular human subjects are, not unexpectedly, of the *human propensity* type, and there is a paucity of *value* adjectives in this group. On the other hand, 43 percent of 3s nonhuman utterances contain *value* adjectives and very few *human propensity* tokens (again, not surprisingly). These and other details will be discussed below.

4.3.2 Predicate adjectives with human subjects

Seventy-six percent (86/113) of the adjectival predicates occurring with subjects with human referents contain *human propensity* and *physical property* adjective types. Table 4.6. summarizes this distribution.

Table 4.6. *Human propensity* and *physical property* adjective types by human subjects (n=86).

	1s	2s	3s	1p	2p	3p	Total
human propensity	13	9	21	1	0	10	54
	24.07%	16.67%	38.89%	1.85%	0.00%	18.52%	100.00%
physical property	7	6	16	1	0	2	32
	21.88%	18.75%	50.00%	3.13%	0.00%	6.25%	100.00%

4.3.2.1 *Human propensity adjectives in clauses with human subjects*

Naturally, 54 of the 65 predicates with *human propensity* adjectives occur with human subjects. Dixon (1991) includes in this semantic class of adjectives items that describe an emotional response (e.g. *angry, proud*) and those that refer to ability or attitude (e.g. *clever, lucky, generous*). The majority of the clauses with 1s subjects—11 out of 13—contain adjectives that express an emotional reaction, as illustrated in (8).

(8) (a) actually I was **amazed**, G15–22
(b) I'm r=eal **proud** of him. D22–6
(c) I'm sure **glad** (it's only thirty-five a=cres= @.) F10–45a

For 2s utterances, on the other hand, the *human propensity* adjectives in 8 out of these 9 tokens do not express an emotional response; instead they tend to describe human characteristics and attitudes (see (9)). Notice, too, that both utterances in (9) are mediated; the assertion in (9)(a) is hedged with the use of *in case*, and the subject *you* in (9)(b) is used generically.

(9) (a) in case you're **interested**. H7–11
(b) you have to be **Catholic** first. L16–12

The one example of an occurrence of a *human propensity* adjective that conveys the subject experiencer's emotional reaction in a 2s clause is given in (10). Note the presence of the mediating modal elements *'ll* and *just*.

(10) you'll just be **disgusted**, R13–10

For *human propensity* predicates occurring with non-speech act participants (i.e. 3s and 3p), the distinction between expressing an emotional reaction (of a third person subject) and characterizing that subject becomes blurred. For example, there are four predicates that contain the adjective *happy* with a third

person singular subject (all occur in the same conversational episode). This lexical item (*happy*) appearing with a first person singular subject might easily be interpreted as conveying an emotional response, but with a third person subject, the emotional response of the subject can also be interpreted as being a characteristic of the subject, and not the subject's emotional reaction. This characteristic, however, is one that is based on the speaker's judgment. The four tokens containing the adjective *happy* follow in (11). The first three intonation units occur consecutively in the conversation; the last one appears subsequently. In these examples, the mother's *happiness* is an evaluation of her character by the speaker (the daughter) and not an expression of the subject's *own* emotional response.

(11) (a) so your mother's **happy** now. D23–4
(b) my mother's <MRC never **happy** MRC>. D23–5
(c) my mother wouldn't be **happy** D23–6
(if everything was -- .. everything was great,)
(d) and she's totally **ha=ppy** now. D23–20

In a sense there is a performative aspect to the use of *human propensity* adjectives, in that only in association with a speaking subject can these lexical items unambiguously express an emotional reaction or state. Otherwise, they are interpreted as *descriptions* of a third person subject. The presence of the verb *be* also contributes to this descriptive or authoritative rendition of the attributions. Johnstone (2002:47–48) observes that "[u]sing *be* in the simple present tense is a way of presenting a claim as universally and hence incontrovertibly true". That is, these utterances are understood as statements of fact about the world, not as expressions of opinion.

Human propensity adjectives occurring with third person subjects appear to characterize the subject separate from the speaker's point of view in part because there is no overt marking of first person reference in these usages. But these adjective types (and others; see below) typically convey speaker judgment as part of their meanings. For example, in (12), *opinionated* is conventionally accepted and interpreted as a personality attribute, but its usage is relative to the judgment of the speaker. Implicit in the semantics of this particular lexical item is a shifting, or indexical, point of view that is dependent on the speaker and context. Similarly, in (13), the meaning of *lazy* is simultaneously a description of a third person subject and an evaluative comment on the part of the speaker.

(12) she's really, definitely **opinionated**, C14–22

(13) because their parents, were too **lazy** to come,
... and, and fill out the stupid form, R7–45

4.3.2.2 *Physical property adjectives in clauses with human subjects*

The class of adjectives Dixon (1991) calls *physical property* adjectives are those items that describe a physical state (of humans, of chairs, of soup, etc.), such as *hard, strong, clean, sick, absent*. Of the 32 tokens of *physical property* adjectives occurring in utterances with human subjects, over half of these make reference to the subject's health. These usages are illustrated in (14).

(14) (a) yet he's still **heal=thy**, D19–54
(b) I think he was pretty **sick**, F3–48b
(c) all of a sudden, I became **i=ll**. M6–35

Because measures of physical health are in large part cultural constructions, assessments of symptoms and manifestations of illness may be more conventionalized in interpretation (also more reified) than are expressions of personal characteristics. When *physical property* adjectives are predicated of subjects with human referents, there is less ambiguity of the type discussed above for *human propensity* adjectives occurring with human subjects. That is, when speakers discuss physical characteristics of people, the property expressed by the adjective is interpreted as an attribute of the subject. However, we will see in the following section that this is not the case for *physical property* predicates occurring with nonhuman subjects; in these utterances the boundary between the semantic and pragmatic meanings of the adjective is analytically less distinct.

4.3.3 Distribution of predicate adjectives with nonhuman 3s subjects

Table 4.7. presents the distribution of relational utterances with adjectival predicates with 3s nonhuman (inanimate) subject types by semantic adjective class. The most frequently occurring subjects with predicate adjective constructions are *it* (41 percent, or 57/138), *that* (38 percent, or 52/138), and lexical noun phrases (16 percent, or 22/138) The most common predicate adjective types are *value* adjectives (43 percent) (e.g. *good, bad, important*), *physical property* adjectives (21 percent), and *qualification* (12 percent) (e.g. *sure, true, possible*). For *it* and *that* utterances, *value* adjectives are the most

Table 4.7. Relational utterances with adjectival predicates with third person singular nonhuman subject types by adjective type (n=138).

	it	that	lexical NP	this	what	Total
age	0	1	0	0	0	1
	0.00%	1.92%	0.00%	0.00%	0.00%	0.72%
difficulty	6	1	0	0	0	7
	10.53%	1.92%	0.00%	0.00%	0.00%	5.07%
dimension	4	0	4	0	1	9
	7.02%	0.00%	18.18%	0.00%	33.33%	6.52%
function	0	2	0	0	0	2
	0.00%	3.85%	0.00%	0.00%	0.00%	1.45%
human propensity	5	2	1	0	0	8
	8.77%	3.85%	4.55%	0.00%	0.00%	5.80%
physical property	**13**	**4**	**10**	**1**	**1**	**29**
	22.81%	**7.69%**	**45.45%**	**25.00%**	**33.33%**	**21.01%**
qualification	**1**	**13**	**2**	**0**	**1**	**17**
	1.75%	**25.00%**	**9.09%**	**0.00%**	**33.33%**	**12.32%**
similarity	2	2	0	0	0	4
	3.51%	3.85%	0.00%	0.00%	0.00%	2.90%
value	**26**	**26**	**5**	**3**	**0**	**60**
	45.61%	**50.00%**	**22.73%**	**75.00%**	**0.00%**	**43.48%**
other	0	1	0	0	0	1
	0.00%	1.92%	0.00%	0.00%	0.00%	0.72%
Total	57	52	22	4	3	138
	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

frequent type, and these adjectives less commonly occur in utterances with lexical subjects. And finally, 76 percent (13/17) of the *qualification* adjectives appear in clauses with *that* subjects, and are tokens of frequently occurring collocations in discourse (e.g. *that's right*). Only the three most frequent subject types (*it, that,* lexical NP) and the two commonly occurring adjective types (*value, physical property*) will be discussed in this section.

Table 4.8. presents the distribution of adjectival predicates by referentiality of (inanimate) subject. See chapter 2, section 2.4.16., especially Table 2.2., for a full discussion of referentiality coding, but a summary of these codes follows. *Entity* refers to those subjects whose referents are people, animals, or things; *generic* categorizes those subjects that designate a class of referents as opposed to individuated referents. *Event-state* codes subjects that refer to a wide range of both general and specific beliefs, evaluations, descriptions, and responses to external and discursive situations. The referential code, *discourse,* marks those subject noun phrases that explicitly refer to linguistic

Table 4.8. Relational utterances with adjectival predicates with third person singular nonhuman subject types by referentiality (n=138).

	it	lexical NP	that	this	what	Total
entity	16	13	7	3	0	39
	28.07%	59.09%	13.46%	75.00%	0.00%	28.26%
generic	3	3	0	0	0	6
	5.26%	13.64%	0.00%	0.00%	0.00%	4.35%
event-state	38	6	28	1	2	75
	66.67%	27.27%	53.85%	25.00%	66.67%	54.35%
discourse	0	0	3	0	1	4
	0.00%	0.00%	5.77%	0.00%	33.33%	2.90%
event-state/discourse	0	0	14	0	0	14
	0.00%	0.00%	26.92%	0.00%	0.00%	10.14%
Total	57	22	52	4	3	138
	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

elements (e.g. words, names) or to bigger stretches of talk (e.g. renditions of what a speaker says). And finally, the double code *event-state/discourse* classifies subject noun phrase referents in the database that are interpreted as both *event-state* and *discourse* referential.

The most general pattern revealed in Table 4.8. is that over half of all subjects in this set of utterances are coded *event-state* (54 percent). However, given that the coding values ev*ent-state, discourse, and event-state/discourse* are continuous in their designation rather than categorical (as opposed to, say, *entity*, which is much more discrete), a decision was made to combine these three groups here for purposes of discussion. The aggregate of these three categories, then, accounts for 67 percent of the total (93/138). This is in contrast to the number of utterances with subjects coded as *entity,* which equal 39, or 28 percent of the total. The trend in these data, then, is for less than one-third of the subjects in these clauses to make reference to specific entities. Conversely, two-thirds of inanimate subjects occurring with predicate adjective clauses refer to conceptually larger, less stable phenomena (i.e. to referents coded *event-state*, *discourse, event-state/discourse*).[6]

Note, too, in Table 4.8. that *entity* referents are mainly expressed by *it* and lexical noun phrase subjects (16/39 and 13/39, respectively), while *that* is the least likely of these frequent subjects to index an entity (7/39 tokens). Additionally, *that* subjects are the only ones that are coded *event-state/discourse* and *discourse*. Local interactions between subject, referent, and the most

frequently occurring adjective types (*value, physical property*) in relational utterances with adjectival predicates will be discussed below.

4.3.3.1 *Value adjectives in clauses with nonhuman subjects*

Though Dixon does not explicitly define *value* adjectives, this class includes items that in English conversation derive their primary meanings from the speaker's evaluation of what is in discursive focus (the subject). Basic English *value* adjectives are *good* and *bad*, but many of the items in this semantic class (e.g. *funny, annoying, interesting*) express more specific meanings. In comparison to other adjective types, it seems to be the case that *value* adjectives are more dependent on the speaker's evaluation as part of their meanings than are, for example, *physical property* adjectives (e.g. *cheap, dangerous, ugly*), which may be said to have fuller propositional meaning separate from the speaker's point of view. However, in usage, it is not clear on what grounds one can argue that words such as *funny* and *annoying* (*value* adjectives) fulfill a greater evaluative function than do *cheap* and *ugly* (*physical property* adjectives). This kind of distinction may be based on people's tacit expectations that opinions about some kinds of activities and behaviors are more widely shared (e.g. that what is *cheap* and *ugly* in English-speaking cultures may be understood as achieving broader consensus than, say, what is *annoying* or *interesting*, which are often understood to be dependent on individual tastes). In other words, what we conceive of as ph*y*sical properties or human propensities are those *opinions* assumed to be commonly held, and it is also those evaluations that speakers tend to interpret as representing objective reality.

Recall from section 4.3.1. that there are very few predicates containing *value* adjectives cooccurring with subjects with human referents; however this semantic class of adjectives is the most frequent type found with subjects with nonhuman referents. Table 4.9. summarizes the distribution of *value* adjectives by 3s nonhuman subject and referent type. Of these 60 tokens, 48 (80 percent) have subjects that refer to non-entities (event-states, etc.), and only 9 subjects with these predicate types (15 percent) have starting points that are referentially discrete (i.e. they refer to entities). [7]

Utterances with *entity* subjects are illustrated below. In (15), the speaker is talking about something she is cooking, so *it* refers to the food (chicken). Similarly, in (16), *that* refers to a vacuum cleaner. In all of the nine *entity* tokens, the subject refers to a specific thing.

Table 4.9. Adjectival predicates with value adjectives by nonhuman third person singular subject type and referent (n=60).

	it	lexical NP	that	this	Total
entity	4	0	3	2	9
	15.38%	0.00%	11.54%	66.67%	15.00%
generic	1	2	0	0	3
	3.85%	40.00%	0.00%	0.00%	5.00%
event-state	21	3	20	1	45
	80.77%	60.00%	76.92%	33.33%	75.00%
discourse	0	0	2	0	2
	0.00%	0.00%	7.69%	0.00%	3.33%
event-state/discourse	0	0	1	0	1
	0.00%	0.00%	3.85%	0.00%	1.67%
Total	26	5	26	3	60
	100.00%	100.00%	100.00%	100.00%	100.00%

(15) it's **good**. G16–41

(16) that's **ni=ce**. A18–54

The subjects *it* and *that* account for the great majority of the 48 predicates with *value* adjectives appearing with subjects with nonentity referents (i.e. subjects with referents coded as *event-state, discourse, event-state/discourse*). When looking at these *it* and *that* utterances in isolation, they appear to be performing identical functions in discourse: that of evaluating an abstractly construed event or state. Moreover, the *value* adjectives that appear in these predicates are similar; we find lexical items such as *funny* or *good* in predicates with both subject types, though there are more frequently occurring collocations with *that* than with *it* subjects (e.g. *that's nice, that's good, that's interesting*).

In their local conversational contexts, however, *it* and *that* predicates with *value* adjectives can be distinguished. Specifically, *it* evaluative clauses are typically uttered by the speaker as part of her own contribution to the discourse. *That* utterances, on the other hand, come from speaker's evaluating (ratifying, sanctioning) the previous participant's contribution. In (17), Z is talking about how she has been pulling out her gray hairs whenever she sees them appear. Her evaluation *it's overwhelming* (*it*=the situation of continually finding herself with gray hair) is part of Z's construction of the "event" in discourse, and it becomes reified both by giving it subject status and in the choice of *it*.

(17) Z: I pull it out,
<CR and I quit CR>
because [it's]--
O: [you know what]?
Z: **it's overwhelming.** C10–11

Example (17) is in contrast to *that* clauses appearing with *value* adjectives, which are typically used by speakers to comment on, or evaluate, a previous speaker's contribution. In (18), Ken is discussing his experiences of being an American in Mexico, eating food from street carts, and becoming devastatingly ill. In this excerpt he is saying that he has decided not to go back to Mexico because it is not worth it to him to risk getting sick again. Lenore's *that's terrible* (*that*=Ken's situation of getting so ill while traveling in Mexico that he has decided not to return and the distress that this state of affairs has caused him) is uttered in response to Ken's story.

(18) KEN: now I'll never go back,
I won't have to worry about it,
I won't be put in that position @anymore.
LENORE: [**that's terrible**]. D13–25

All but 2 of the 23 *that* utterances with *value* predicates are similar to (18); they function as empathetic evaluations of a *previous* speaker's contribution by a *subsequent* speaker. Interestingly, the two *that* clauses in this group that don't quite fit with this pattern (i.e. they are utterances used by speakers to evaluate their own contributions) are both tokens of *isn't that great?*. What is notable about these two utterances is that in requesting evaluation from other participants, the speaker is using *that*—the form most expected were the addressee to have initiated the exchange. The pattern for *it* utterances in the database is also noteworthy; all of these 21 *value* clauses with *event-state* subject referents are uttered by speakers to comment on their *own* discursive constructions, as illustrated in (17) above. Though the counts of these *it* and *that* clauses are relatively small (21 and 23, respectively), the pattern is striking. Both utterance types are evaluative; however, speakers tend to use *it* forms in their own commentaries, while *that* variants are used by participants to respond to a previous speaker.

What theoretical insights, then, might be useful in understanding the different functions of *it* and *that* evaluative utterances in English conversation? Greenberg (1985) suggests that, typologically, distance demonstratives

are generally used to refer to the past, including metaphorical reference to past discourse. Results from the present study are amenable to this claim in that speakers use the distal form, *that*, to make reference to another participant's previous contribution. However, in a study of Finnish demonstratives in conversation, Laury (1997:94) finds that distality and proximality (and their metaphorical extensions, e.g. time, discourse) do not provide an ideal framework in which to investigate the social nature of demonstratives in interactive discourse.

In English evaluative *that* utterances, participants acknowledge and ratify a previous speaker's discursive construction. However, when speakers provide commentary on their own contributions using these relational expressions, they do not use the proximal *this* as subject, but rather, the impersonal pronoun *it*. In these data, *it* functions in these forms to signal that the speaker's point of view is anchored in her own sphere, not the addressee's (Laury 1997). Given the paradigmatic asymmetry of usage in this situation (i.e. that English speakers are using *that* to refer to another's previous discourse, but not using *this* to refer to their own conversational contribution), an explanation based solely on the metaphorization of spatial deixis may not be the best choice in this interactive context. In his discussion of the social and relational nature of deictic reference, Hanks (1992:52) remarks, "The standard assumption that space is always foundational in deixis is an inconvenient fiction not borne out comparatively".

Bolinger (1977) discusses cases of *it* that are not anaphoric, which he calls *ambient it*. He suggests that these uses make reference to the immediate situation—in the case in question, to the speaker's construction of her current discourse, or reality. He states that "*it* is more general than *things*. It embraces weather, time, circumstance, whatever is obvious by the nature of reality or the implications of context" (Bolinger 1977:85). The uses of *it* subjects with *value* adjective clauses in the present study are compatible with Bolinger's analysis. In these evaluative utterances, speakers are evaluating their own contributions; they are using *it* as a subject to construct their discursive reality for themselves and others. What may follow from this usage in natural discourse, then, is the subsequent speaker's indexing these constructions with the demonstrative, *that*, to demonstrate her own evaluation of what has preceded. In a sense, then, in these interactive contexts, *it* constructs a discursive fact and *that* is used to make reference to it by another participant. On the other hand, *this* subjects of relational clauses in the conversational corpus perform a

Table 4.10. Adjectival predicates with physical property adjectives by nonhuman third person singular subject type and referent (n=29).

	it	lexical NP	that	this	what	Total
entity	8	8	2	1	0	19
	61.54%	80.00%	50.00%	100.00%	0.00%	65.52%
generic	1	0	0	0	0	1
	7.69%	0.00%	0.00%	0.00%	0.00%	3.45%
event-state	4	2	2	0	1	9
	30.77%	20.00%	50.00%	0.00%	100.00%	31.03%
Total	13	10	4	1	1	29
	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

more specialized function; they are most frequently used to refer to concrete objects in the physical context (see section 4.4.2.1. below).

4.3.3.2. *Physical property adjectives in clauses with nonhuman subjects*
Physical property adjectives are the second most frequent type of adjectival predicates occurring with 3s inanimate subjects. Table 4.10. summarizes the distribution of these utterances by subject and referentiality.

Recall from the previous section that the most frequent subject referent to occur with predicate complements containing *value* adjectives is a nonentity (i.e. *event-state*, etc.). In that group, only 15 percent of the tokens have subjects that refer to specific entities, while 80 percent of the subjects refer to events and states as discursive constructions. However, the *physical property* group of utterances with 3s nonhuman subjects examined in this section shows a reversal of this pattern. Table 4.10. indicates that the majority of these subject referents are coded *entity* (66 percent)—twice as many as the tokens with subjects marked *event-state*. Also note in Table 4.10. that 16 out of the 19 utterances with *entity* referents (84 percent) are expressed with either *it* or a lexical noun phrase; only two of these clauses have *that* subjects.

Recall that the great majority of *physical property* predicates that occur with human subjects are from the semantic domain of health. In contrast, the physical properties that are predicated of nonhuman *entity* subjects examined in this section are colors and physical states (e.g. *green, toxic, dirty, absorbable, wrapped*). The subjects of these utterances are overwhelmingly *it* and lexical noun phrases (e.g. ***it**'s like, ... distorted, **your blood**'s all shot,*). While there is an impression with these adjectival complements that what is being linguistically

expressed are characteristics inherent to the cooccurring inanimate subject entities, the majority of meanings of these forms are, in fact, determined by speaker point of view. For example, there is no nonsubjective measurement of what is *dirty* or *ugly*. The most propositional, or objective, uses in this group of *physical property* adjectives are those predicates with color adjectives, illustrated in (19).

(19) it's not **green**, A13–7

However, even a relatively straightforward expression of a relation between a thing and a color is less objective when the utterance contains a linguistic expression of speaker stance such as an adverb of intensity or modality. Examples follow in (20) and (21).

(20) if her skin weren't like .. **really dark brown**, R9–48

(21) well it's **just as yellow as ... (1.3) ... can be**. F11–15a

In (22) and (23) below, the adjectives *dirty* and *toxic* are linguistically marked as expressive by the amplifier *all* and the modality adverb *almost*, respectively.

(22) and his face was **all dirty**, F6–28

(23) so it's **almost impervious** to destruction. A19–6

Of the nine tokens of *physical property* predicates whose subjects refer to *event-states*, three of these have *it* subjects, two have *that* subjects, three of these subjects are lexical noun phrases, and there is one *what* subject. In this database, there is no lexical overlap between the adjectives in this *event-state* group with those that appear with nonhuman *entity* subjects discussed above. The *physical property* adjectives occurring with *event-state* subject referents are words such as *strong, ugly, distorted, dangerous*, and *cheap*—all lexical items whose meanings are in large part determined by the speaker's evaluation. In other words, when these so-called *physical property* adjectives attribute characteristics to discursive events or states (e.g. opinions, attitudes, perceptions)—as opposed to entities—their meanings have a tendency to be figurative and relative to speaker stance. For example, in (24), the speaker is an elementary school teacher who is complaining about what she sees as the general behavior of substitute teachers in her school. In this utterance it is obvious that the adjective *distorted* does not refer to a physical situation; rather it is an evaluation on the part of the speaker.

(24) it's like, ... **distorted**.
… they don't know what they're doing. R12–15

Similarly, in (25), Speaker A is commenting on an episode in a movie in which a French character hid a Jewish friend in France during World War II but was simultaneously working with the Nazis so that she wouldn't arouse suspicion. Her utterance, *that's just really ugly*,—while containing the *physical property* word *ugly*—is an evaluation of the situation and not an ascribing of physical characteristics. This usage is not qualitatively different from the *value* predicates discussed immediately above.

(25) B: he- here she was helping … somebody but,
.. she tried to look like she wasn't,
she --
A: **that's just really ugly**. C6–36

And finally, it is worth noting that eight out of the nine *physical property* predicates with nonhuman subjects referring to event-states contain an intensifying or modal expression (e.g. *really, just, pretty*). For the same utterances with *entity* subjects, only 9 out of 19 clauses (47 percent) have a similar adverb. This suggests that speakers recognize the evaluative character of the former group and often enhance this use lexically.

4.3.4 Summary of predicate adjective constructions

Using Dixon's (1991) semantic taxonomy of English adjectives to code and analyze the 263 predicate adjective constructions in this database provides a useful organization for examining the meanings and functions of these predicates in interactive discourse.[8] While the theoretical goal in this portion of the study is not to evaluate the validity of these semantic types as analytical tools in their own right, the categorical indeterminacy among semantic classes (e.g. *physical property* vs. *value* adjectives) that surfaced during the coding of lexical items (e.g. *ugly, strong*) points to a more general analytical problem—that of viewing linguistic expressions as referring to preexisting semantic substance recruited to characterize the world separate from these elements' more pragmatic meanings, specifically, those based on the speaker's point of view.

The analyses in this section suggest that to varying degrees English adjectives have a subjective component to their semantics. In a study of styles of

speaker stance in English, Biber and Finegan (1989: 118) similarly note that "attributive adjectives sometimes seem to mark stance in addition to marking descriptive elaboration or referential identification".

Human propensity adjectives (e.g. *lazy, happy*), in particular those occurring with 3s human subjects, are ambiguous in expressing a characteristic of the subject and the speaker's evaluation of that subject referent. For 3s nonhuman subjects, the frequently occurring *value* adjectives (e.g. *good, funny*) are indistinguishable from many *physical property* adjectives (e.g. *ugly, distorted*) as regards their expression of speaker evaluation. I have suggested that the cognitive and analytical distinction between adjectives with more specific lexical content and less semantic room for speaker opinion (e.g. *physical property* adjectives, *human propensity* adjectives) and adjectives with more general lexical meaning and a robust expressive component (e.g. *value* adjectives) may in many cases be a distinction between characteristics that are construed as more conventional vs. those considered to be more idiosyncratic.

Examination of the data in this section also highlights some interesting usage patterns of 3s nonhuman subjects in predicate adjective clauses. In general, the frequent subjects, *it* and *that*, both make reference to events and states. *It*, however, also commonly indexes entities, though *that* rarely fulfills that function. With *value* adjective predicates, two-thirds of the nonhuman 3s subjects in this set refer to discursive events and states, and both *it* and *that* occur in this context. But these two subjects are not used in the same interactive contexts. Tokens of *it* with *value* predicates are used by speakers as commentaries on their own conversational contributions; *that* + *value* predicates, on the other hand, are used by a subsequent speaker to evaluate a previous participant's discourse. It was suggested that the occurrence of *it* in these utterances is compatible with Bolinger's (1977) characterization of ambient *it* and is used by speakers in English conversation to construct a discursive reality. These two frequent inanimate pronominal subjects of adjectival clauses, then, are interactively and referentially distinct. In terms of general function, however, they both index discursive, often abstract starting points for the speaker's evaluations (Kuno & Kaburaki 1977; Stein 1995).

4.4. Relational predicates with nominal complements

4.4.1 Distribution of predicate nominal clauses

The 277 predicate nominals in the database are a heterogeneous group as regards subject referentiality, type of noun phrase and semantic class of the lexical noun in the complement, and linguistic context. Unlike the distinctive distributional patterns found for predicate adjective clauses in the previous section (e.g. the high frequency of *value, physical property*, and *human propensity* adjective types with particular subjects and referents, the differential uses of *that* and *it* utterances in interaction), the distribution of categories and associations among linguistic elements in predicate nominal utterances is less striking, in the sense that there are no highly frequent correlations between noun type (e.g. *human, artifact*, *time-place*) and subject. The differences in patterning of these two relational clause types (predicate adjective and predicate nominal clauses) are accompanied by differences in the relations expressed by these two types of utterances.

Table 4.11. presents frequencies for nominal predicates by subject. As is true for all relational clauses, the bulk of these utterances have 3s subjects (89 percent), and 75 percent of all predicate nominal clauses occur with 3s subjects with nonhuman, or inanimate, referents. In contrast, only 52 percent of predicate adjective clauses occur with nonhuman subjects. That there is a greater percentage of inanimate subjects (in comparison with human subjects) for the nominals than for the adjectival group may be due to the different functions of these two utterance types. When conversational participants introduce or construct a discourse phenomenon, they tend to do so with 3s inanimate relational clauses with nominal complements. Using these utterances, speakers situate discursive starting points (subjects) and *identify* them in a variety of ways with a nominal predicate. In contrast, the relation between subject and adjectival complement typically evaluates an extant referent of some sort; it does not construct a new one. There are not as many subjects with human referents of nominal clauses (compared to adjectival clauses) because people are not typically constructed in interactive discourse. Human beings concretely exist; therefore, their materiality is not reliant on discursive construal or negotiation. The *evaluation* of human beings, on the other hand, is commonly constructed and negotiated through the discourse. This difference may contribute to there being more human subjects with predicate

Table 4.11. Relational utterances with nominal predicates by subject (n=277).

	% nominal predicate group
1s	11
	3.97%
2s	10
	3.61%
3s (total)	247
	89.17%
3s (human)	40
	14.44%
3s (nonhuman)	**207**
	74.73%
1p	0
	0.00%
2p	1
	0.36%
3p (total)	8
	2.89%
3p (human)	5
	19.23%
3p (nonhuman)	3
	1.08%
Total	277
Percent	100.00%

adjective clauses (43 percent) than with nominal clauses (24 percent) (see Table 4.3.).

As is the case for relational clauses with adjectival complements, there are few nominal clauses occurring with the speech act participant subjects 1s and 2s (approximately 4 percent for each), and these are smaller percentages than those found for adjectival clauses (see Table 4.3.). Again, this may have to do with the distinct functions of these clause types discussed above. Subjects with human referents are semantically and pragmatically constrained with respect to cooccurring predications (e.g. because of the social appropriateness of how people may be labeled or situated in cooperative conversation). For example, the great majority of nominal complements that occur with human subjects are, not surprisingly, of the *human* type (e.g. *teacher, women, very very good writer, an old lady*), and there are very few nouns of other types appearing with human subjects.

With respect to subject referentiality for nominal predicates—similar to relational clauses with adjectival predicates—Table 4.12. indicates that there

Table 4.12. Relational utterances with nominal predicates with third person singular nonhuman subject types by referentiality (n=207).

	it	lexical NP	that	there	this	what	Total
entity	18	5	10	0	19	17	69
	27.27%	21.74%	17.54%	0.00%	79.17%	68.00%	33.33%
generic	0	2	1	0	0	0	3
	0.00%	8.70%	1.75%	0.00%	0.00%	0.00%	1.45%
event-state	36	8	21	1	3	6	75
	54.55%	34.78%	36.84%	8.33%	12.50%	24.00%	36.23%
discourse	7	7	18	0	1	2	35
	10.61%	30.43%	31.58%	0.00%	4.17%	8.00%	16.91%
event-state/discourse	1	1	7	0	1	0	10
	1.52%	4.35%	12.28%	0.00%	4.17%	0.00%	4.83%
nonreferential	4	0	0	11	0	0	15
	6.06%	0.00%	0.00%	91.67%	0.00%	0.00%	7.25%
Total	66	23	57	12	24	25	207
Percent	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Percent subject	31.88%	11.11%	27.54%	5.80%	11.59%	12.08%	100.00%

are more subjects in this group with *event-state, discourse,* and *event-state/discourse* referents (58 percent or 120/207) than there are with *entity* referents (33 percent or 69/207). Additionally, in line with the distribution of all subjects of relational clauses, the most frequent 3s inanimate subject types are *it* (31 percent) and *that* (28 percent). Both of these subjects refer to event-states, though *it* also more frequently indexes an entity than does *that*.

4.4.2 The subject-nominal relation in predicate nominal clauses

As discussed in section 4.3., relational clauses with adjectival complements in American English conversation function primarily as evaluations. In general, the relation expressed by these structures is one between a subjective focus, or starting point, (typically expressed by an inanimate 3s subject) and an evaluation of that focus by the speaker (expressed by the lexical adjective in the predicate).[9] To varying degrees, these lexical adjectives include as an important part of their meanings the speaker's opinions and attitudes. This act of predicating properties, or characteristics, of a discursively construed entity or event is a subjective act, even when those properties are considered to be an inherent part of the referent, thus unacknowledged as contingent on the speaker's stance.

In contrast to these adjectival utterances, the relation expressed by clauses with nominal predicates does not appear to be overtly subjective. Unlike the adjectival predicates, these nominal expressions are formally heterogeneous.[10] For example, these noun phrases can be as compositionally simple as a pronoun (e.g. *and that's **it,***) or a determiner and a noun (e.g. *it's **a kni=fe=,***); they can include adverbs and modifying adjectives (e.g. *it is **a different sensation entirely.***), or the nominal predicate can consist of an entire clause (e.g. *that's **what they smell like.***).[11]

There is a correlation between the function and referentiality of the subject and that of the nominal predicate. For example, only one of the 23 predicate nominals coded *clause/event* has an *entity* subject (*that's **what you got–***);[12] instead, the majority of these *clause/event* noun types occur with subjects that refer to events, states, and discourse material (e.g. *it was **what was convenient,***). On the other hand, *entity* subjects overwhelmingly occur with noun phrases containing nouns semantically coded as *artifact* (e.g. *this is **beta food,***), *body part* (e.g. *it's just dry **skin.***), and *proform* (*what is **it,***).[13] These examples illustrate a functional relationship between the subject of the utterance and the predicate nominal in that the two elements share referentiality. That is, concrete nouns such as artifacts and body parts have concrete, individuated subject referents (i.e. referents coded in this study as entities); similarly, nominal predicates expressing events or abstract phenomena (often these are full clauses) typically occur with subjects with *event-state* referents.

This matching between subject referentiality and semantic and formal characteristics of noun phrases in the complement suggests that the type of relation between subject and predicate expressed by these nominal utterances is an *identifying* one. Halliday (1994: 122) defines the *identifying* mode of relational clauses as one in which "one entity is being used to identify another: '*x* is identified by *a*', or '*a* serves to define the identity of *x*". Halliday also states that the *identifier* (in this case, the nominal in the predicate) is typically definite. However this is not always the case for nominal relational clauses found in conversation—a striking illustration are those with head nouns coded as *empty* (e.g. *that's just total quack **stuff,***). Halliday's examples of identifying clauses tend to have lexically substantive subjects and noun complements, both with *entity* reference (e.g. ***The deadliest spiders in Australia** are **the funnelwebs.***). However, this is not the typical kind of relational clause found in English interactive discourse. Because relational utterances with nominal predicates in discourse exhibit variation in both complement

type and subject referentiality, it is useful to examine how this *identifying* relation manifests in these clauses in conversation. The next two sections discuss this identifying function in utterances cooccurring with subjects referring to entities (section 4.4.2.1.) and in those with event-state-discourse referents (section 4.4.2.2.).

4.4.2.1 *Entity subjects with nominal predicates*

Recall from Table 4.12. that 33 percent of nominal predicates with 3s nonhuman subjects occur with subjects referring to entities, and 78 percent of these subjects are *it, this,* and *what*. The most common noun types occurring in predicates with 3s nonhuman *entity* subjects are *artifacts* (e.g. *spoon, chart, bag*) and *proforms* (e.g. *it, that*). However, these particular utterances cluster in discourse in very specific situations; they occur in conversation when participants are interacting with objects in the physical context. For example, in the conversation, *Appease the Monster*, many of these *entity* utterances appear during a segment when one of the participants is opening her birthday presents and the focus of the conversation is on objects in the immediate environment. In (26), the speaker is unwrapping a present (a cookie baking set) and she sees a knife in the package. The relation between *it* and *knife* in this utterance is a straightforward identification of the item (i.e. 'the entity referred to by *it* is a knife').

(26) it's a kni=fe=, A13–47

Another conversation where these clauses cluster is *Deadly Diseases*, a discourse that contains a fairly long segment in which two of the participants are irreverently inquiring about a third participant's cache of vitamin and mineral supplements. In this episode, there are several tokens of inquiry/response pairs of the type *What is it? This is vitamin C.*[14] In (27), the speaker is replying to one of these requests for information.

(27) and then this is just basic, uh=, .. very absorbable iron, D18–36

Notice, too, in (27) that the speaker hedges her identification of the supplement; this utterance follows several turns of the other participants' joking about the smell of the pills.

Both *this* and *it* subjects in this *entity* group participate in an identifying relation with their cooccurring nominal predicates. *This* utterances identify objects in the physical environment and only appear in the present tense (e.g.

this is the plastic bag., this is fake fried chicken.). *It* subjects with *entity* referents, on the other hand, are less functionally constrained. These *it* utterances occur in the past tense as well as the present, and, unlike *this* clauses, only four of the 18 *it* tokens identify objects in the physical context (e.g. *it's a used vacuu=m, oh it's a piece of Wendy's hair.*). In the other 14 items, *it* has slightly more abstract reference. Seven of these tokens are past tense, which typically indicates that the referent is not present in the physical context (e.g. *well, .. it was a test plot,*).

There is another subset of *entity it* clauses for which the predicate nominal element is less concrete than are those designating objects in the immediate context (e.g. *and it's* ***a story*** *of a .. Jewish woman, it's* ***a good bra=nd,***). In the majority of these predicate nominal clauses with *it entity* subjects, the identifying relation between subject and predicate is more interpretive, or subjective, than are those expressed by *this entity* utterances. The latter relation is one of identity (e.g. ***this*** *is* ***the spoo=n,***). For the majority of *it* clauses, however, the predication evaluates or describes the subject referent—a function beyond simple identity.

4.4.2.2 *Event-state-discourse subjects with nominal predicates*

One hundred and twenty of the 207 relational utterances with nominal predicates are coded as having *event-state, discourse,* and *event-state/discourse* subject referents (58 percent),[15] and the most frequent subject types are *that* (46/120 or 38 percent) and *it* (44/120 or 37 percent). The most common predicate nominal type to occur with these event-state-discourse subjects is the *clause/event* category (21/120 or 18 percent) followed by the *empty noun* category (13/120 or 11 percent). Unlike the utterances with *entity* subjects discussed in the previous section, these nominal forms with event-state-discourse subject referents are distributed more evenly throughout the nine conversations coded in the database; they do not cluster during specific activities as do the *entity* utterances.

Sixty-seven percent (14/21) of the predicates coded *clause/event* occur with *that* subjects. This common association between *that* (as opposed to *it* or *this*) and a noun phrase complement may be due to the frequency of the English construction *that's what* + CLAUSE (e.g. *that's what they smell like, that's what I think too*).

Of the total 21 tokens in this *clause/event* group, seven of the nominal predicates (nominal relative clauses) contain first person singular expressions. Examples are provided in (28).

(28) (a) but that's **what I want to learn**. L20–26a
(b) now that's **what I just go by memories**. M9–53
(c) that's **what I assumed**. L23–33

In these utterances, the identifying relation is between the starting point (the subject) and something the speaker associates with it in feeling, interest, or action. In (28)(a) and (28)(c), *that* indexes a previous piece of discourse. In (28)(b), however, *that* is cataphoric since the referent is only made explicit at the end of the utterance—the referent being the speaker's *go[ing] by memories*. In general, then, the identifying relation expressed in these utterances is more abstract than the relation of *identity* noted above for nominal utterances with *entity* subjects. In the present group, the relation of *identifying*, or, perhaps, *identifying with*, is illustrated by the speaker's linguistically linking ideas and constructions in the discourse with her own experiences.

Another small set of utterances containing *clause/event* noun phrases and event-state-discourse subjects are nominal predicates composed of predicate adjective constructions. In (29), the speaker's utterance, *that's what's so ama=zing*, links her telling of her husband's "miraculous" recovery from a litany of illnesses with her evaluation of the situation.

(29) .. he's only had,
.. <@ since .. since we've been @> married,
… ca=ncer,
@@@@
.. @ leukemia=,
… bronchitis=,
… uh=,
.. tuberculosis,
.. @@@@
.. and he's recovered from all of them.
that's what's so ama=zing. H4–18

The second most frequent nominal type occurring with event-state-discourse subjects is the *empty noun* group.[16] Seven of these thirteen nominals contain the lexical item *thing*, and all of these occur with attributive adjectives (e.g. *it was the same thing, it's a very complex thing.*). Thompson (1988:175) discusses cases in English conversation "where an attributive adjective functions predicatively because its predicate nominal head is (relatively) *empty*". These examples in the database are amenable to Thompson's analysis in that an

utterance such as *it's a very complex thing* is similar in function to an analogous predicate adjective construction (e.g. *it's complex*). Both of these utterance types (the predicate nominal and the predicate adjective) are evaluations on the part of the speaker. In the predicate nominal expression, however, there is a sense in which the relation between the subject referent (*it* in this case) and the predicate (*a very complex thing*) is constructive as well as evaluative. In the nominal form, *it* both creates (=recognizes, or sanctions) a particular reality *and* links it to an evaluation, whereas the adjectival evaluation presupposes the existence of the point of focus represented by the subject and the *empty noun* in the predicate (e.g. *thing*).

4.4.3 Summary of predicate nominal constructions

Relational clauses with nominal complements are formally and functionally a heterogeneous group, and analyses of utterance types in context reveal local patterns of usage. Third person singular nonhuman subjects of these clauses may refer to individuated, often concrete, entities or, more commonly, to abstract construals and constructions of events and states. Based on the coding of these elements (the subject and the head noun in the predicate), it was suggested that functionally there is an *identifying* relation between these two nominal constituents (Halliday 1994), but that the character of the expressed relation varies depending on the referentiality of the subject and the semantic and structural composition of the predicate.

For utterances with *entity* subjects—especially the proximal demonstrative *this*—the relation expressed is one of identity in which the speaker verifies the identity of something which is often present in the physical context (e.g. *this is salsa, right?*). In utterances with *it* entity subjects whose predicates are past tense or whose head nouns in the complement are not concrete (e.g. *but it was a … sce=ne,*), the expressed relation is no longer one of simple identity but one of *associating* one "thing" (indexed by the subject) with another (expressed by the predicate). Identification in such a situation in which one is gauging whether something is the same as, or similar to, something else is a subjective act. This is because anything beyond a relation of strict identity (e.g. as we saw for *this* utterances) is necessarily based on the speaker's judgment and perception of what constitutes the similarity or association between discursive entities in a given situation.

For utterances with event-state-discourse subjects, there is also a func-

tional continuum with regard to the type of relation that exists between subject and predicate nominal. For culturally reified events (e.g. *Christmas*), the relation between ambient *it* (the most frequent subject in these utterances) and the predicate typically situates (identifies) the event in time (e.g. *it's midterm*). The most subjective identifying relation in the nominal group, however, is expressed by utterances whose subjects refer to discursively construed events and states and whose predicates are clauses (e.g. *that's what's so peculiar, that's what I don't like about him.*).[17] In these constructions, evaluative adjectives within the nominal clause and clauses with first person singular expressions—overtly subjective linguistic material—function as the identifying constituent in the relation. That is, the speaker associates her focus (the subject of the clause) with an explicit evaluation. In doing so, speaker evaluation in these clauses (the clausal complement) is syntactically, and perhaps conceptually, nominalized. The identifying relation in these expressions, then, is one between the discursive starting point and the speaker's reified (nominalized) point of view.

4.5 Relational predicates with oblique complements

In this study, the label *oblique* classifies those relational predicates in the database that are neither adjectival nor nominal. There is, however, structural and functional coherence to this group. Oblique complements as defined here primarily designate relations of space (example (30)) and time (example (31)), and are often expressed as prepositional phrases.

(30) (I think) she's still **at home**. A9–3b

(31) that's **right now**=. L17–16

Table 4.13. summarizes the distribution of oblique clauses by humanness of subject referent in comparison to predicate adjective and predicate nominal clauses. Most striking about this distribution is the low frequency of oblique predicates in the corpus compared to the other two relational predicate types. Only 12 percent (78/641) of relational clauses have oblique complements, while predicate adjective and predicate nominal tokens account for 41 percent and 43 percent of the relational group, respectively. Additionally, unlike relational utterances with adjectival or nominal complements, oblique predicates occur with proportionally more subjects with human referents than

Table 4.13. Relational utterances by predicate type and animacy of subject referent (n=641).

	Adjective	Nominal	Oblique	Other	Total
human	113	67	**41**	8	229
	42.97%	24.19%	**52.56%**	34.78%	35.73%
nonhuman	150	210	**37**	15	412
	57.03%	75.81%	**47.44%**	65.22%	64.27%
Total	263	277	**78**	23	641
Percent	100.00%	100.00%	**100.00%**	100.00%	100.00%
Percent Relational	41.03%	43.21%	**12.17%**	3.59%	100.00%

Table 4.14. Relational utterances by tense (n=641).

	Adjective	Nominal	Oblique	Other	Total
Present	212	216	**47**	12	487
	80.61%	77.98%	**60.26%**	52.17%	75.98%
Past	33	51	**27**	6	117
	12.55%	18.41%	**34.62%**	26.09%	18.25%
Modal	18	10	**4**	5	37
	6.84%	3.61%	**5.13%**	21.74%	5.77%
Total	263	277	**78**	23	641
Percent	100.00%	100.00%	**100.00%**	100.00%	100.00%

those with nonhuman referents (53 percent vs. 47 percent).

Table 4.14. presents the distribution of relational clauses by tense and by predicate type. As indicated in this table and in Table 4.1., 76 percent of all relational clauses are present tense.

Table 4.14. indicates that there is proportionally more present tense than past tense for all relational predicate types; however, notice that there is a greater percentage of past tense for oblique predicates than there is for predicate adjectives or predicate nominals (35 percent for obliques vs. 13 percent and 18 percent, for adjectives and nominals, respectively). As was suggested in chapter 3 for nonrelational utterances, past tense is often found in predicates that function propositionally in discourse to narrate past events. This issue will be discussed below.

Aside from there being proportionally more human subjects and more past tense in predicate oblique utterances than in predicate adjectives and predicate nominals, there is another distributional trend that distinguishes the obliques from the two other relational types. Table 4.15. presents frequencies

Table 4.15. Relational utterances by referentiality of subject (n=641).

	Adjective	Nominal	Oblique	Other	Total
SAP	50	20	17	6	93
	19.01%	7.22%	21.79%	26.09%	14.51%
entity	**88**	**103**	**41**	**6**	**238**
	33.46%	**37.18%**	**52.56%**	**26.09%**	**37.13%**
SAP/entity	0	0	2	0	2
	0.00%	0.00%	2.56%	0.00%	0.31%
generic	26	14	4	2	46
	9.89%	5.05%	5.13%	8.70%	7.18%
SAP/generic	3	1	1	0	5
	1.14%	0.36%	1.28%	0.00%	0.78%
event-state	78	76	9	6	169
	29.66%	27.44%	11.54%	26.09%	26.37%
discourse	4	38	1	1	44
	1.52%	13.72%	1.28%	4.35%	6.86%
event-state/discourse	14	10	1	1	26
	5.32%	3.61%	1.28%	4.35%	4.06%
event-state-discourse	**96**	**124**	**11**	**8**	**239**
	36.50%	**44.77%**	**14.10%**	**34.78%**	**37.29%**
nonreferential	0	15	2	1	18
	0.00%	5.42%	2.56%	4.35%	2.81%
Total	263	277	78	23	641
Percent	100.00%	100.00%	100.00%	100.00%	100.00%

of relational clauses by referentiality of subject. This display shows that there are proportionally fewer event-state-discourse subject referents (the aggregate group) with oblique clauses than with the adjectival and nominal complements (14 percent for obliques vs. 37 percent and 45 percent for adjectives and nominals, respectively). For the oblique group, then, with respect to subject referentiality, there is less subjective construal of a starting point on the part of participants, in the sense that these utterances tend to be used to situate individuated, extant entities in interactive discourse.

In contrast to adjectival and nominal clauses, then, relational utterances with oblique predicates are not very frequent in conversation; they tend to appear more often in the past tense; and their subject referents are more likely to be human, or, more generally, entities, rather than events or states. A similar convergence of distributional characteristics (human subjects, past tense) was noted in chapter 3 for 3s utterances with material verbs. It was suggested then that those clauses are less personalized, structurally and functionally, than the majority of utterances found in English conversation and are

used by participants to report on events in the world.

Though there are only 78 oblique tokens in the database, the distributional patterning of structural and functional elements of these items points to a consistent function for this relational group. For oblique utterances, the relation that is most often expressed is one in which speakers index entities (human and nonhuman) in terms of their (past and present) locations, as illustrated in (32) (a) and (b). Sixty-two percent, or 48/78, of oblique utterances have spatial predicates.

(32) (a) we were out, M6–29
(b) peas are in em so, F5–12

This relation between subject and predicate which situates entities in space is especially strong for oblique utterances with human subjects (31/41 or 76 percent of this group), and for this set there is an equal number of present and past tense tokens.

For oblique clauses with nonhuman subjects, on the other hand, the distribution is suggestively different. Only 8 of these 37 tokens (22 percent) have past tense predicates, and the relations expressed in these 8 tokens are equally divided between those that are spatial (e.g. (32)(b), a*nd the Mexican governor= — .. general's house=, .. was right there=*,) and temporal (e.g. *my test was Friday,*). The majority of oblique predicates with nonhuman subjects (28/37 or 76 percent) are present tense. Of these present tense tokens, 13/28, or 46 percent, express spatial relations between subject and predicate, and nine utterances (32 percent of present tense items with nonhuman subjects) express other case relations (e.g. benefactive: *it's for people who can't move their hips.*, ablative: *it's from her mother's lover.*) The trend, then, in these oblique clauses with subjects with nonhuman referents is for speakers to use less past tense and a wider range of functional predicate types than for similar clauses with human subjects.

Even with this small sample, one finds that it is in this particular structural domain that relatively unpersonalized, descriptive utterances occur. The analysis of oblique clauses offered in this section provides support for the argument proposed in chapter 3—that in interactive discourse in English, it is only in specific structural and pragmatic contexts that more propositional language occurs. It seems to be the case that when English speakers are focusing on human beings as subjects, they use utterances that are less mediated by point of view than is the case for most of their other talk. Though this

kind of language occurs in conversation—in particular to situate people with respect to space and time—it is relatively infrequent compared to other relational utterance types.

4.6. Summary

This chapter has examined the forms and uses of the most frequently occurring utterances in the conversational corpus—relational clauses. Unlike the lexical verb types analyzed in chapter 3, relational predicates do not commonly occur with subjects with human referents. The most typical subject appearing in the relational group is third person singular with a nonhuman referent, usually *it* or *that*. Clauses with nominal and adjectival predicates are the most frequent of the relational utterances, while oblique predicates account for only 12 percent of this group, and they are more functionally constrained than the other two types. With respect to distribution of tense, relational clauses as a group have proportionally more present tense than past compared to all other utterance types in the database.

Because utterances in this verb group express a relation between subject and predicate (see Halliday 1994), it has been useful to examine combinations of subject referents and lexical items in predicates in order to investigate the kinds of relations that these structures express in interactive discourse. Recall that predicate adjectives and predicate nominals more frequently occur with event-state-discourse subjects than with entity subjects. Notably, in these adjectival and nominal clauses in the corpus, the relation between subject and predicate is not one between an acting subject and an activity. For predicate adjective clauses, the relation is between a discursively conceived, often co-constructed, focus or starting point (the subject) and an evaluation of that focus, linguistically expressed by the adjective in the predicate. For predicate nominals, the relation between the subject and the predicate is an identifying one, and in most situations—those beyond the simplest cases of identity of objects present in the physical context—speakers "identify" the subject referent based on subjective criteria (e.g. experience, expectation) as to what constitutes similarity.

Both adjectival and nominal constructions routinely function subjectively in English conversation in the sense that the relations conveyed by these utterances are contingent upon speaker point of view. It is only the compara-

tively infrequent predicate obliques which usually express a more objective relation—one between discrete entities in the world (e.g. humans, objects) and, typically, their locations (expressed in the oblique complement). Fox (1995:158) also notes a distinction in copular clause constructions in English conversation between those that are evaluative and 'non-evaluative utterances, which can serve to locate, describe, identify or name an entity'. As we saw in chapter 3 during discussion of third person singular human subjects with material verbs, there are structures in English conversation that speakers use to report on events, but in interactive discourse, these are the least frequent utterance types. The bulk of third person singular utterances are not used by speakers to *describe* properties of people and processes in the external world, but, rather, to discursively endow events, ideas, and entities with characteristics that are based on participants' *evaluations*. Moreover, because these utterances are not marked as first person, these subjective construals are covert—conventionally understood as characterizations separate from the speaker's point of view.

Notes

1. In this study, '*x* has *a*' clauses are grouped separately in a verb class called possessive/relational and are not included in the larger relational group discussed here.

2. In the Thompson and Hopper corpus, one participant clauses themselves account for 73 percent of the utterances.

3. Due to the frequency of predicate nominals with 3s nonhuman subjects, one also finds proportionally more predicate nominal than predicate adjective complements with all 3s subjects.

4. In Himmelmann's (1996) coding system, reference to events is understood as referential to the discourse. In the current study, the tag *discourse* is defined in a more narrow sense, grouping those subject noun phrases that explicitly refer to linguistic material. The codes *event-state* and *event-state/discourse* together are analogous to Himmelmann's *discourse* label.

5. There are, however, numerous studies on the functional aspects of referentiality. Most relevant to the present investigation are examinations of deixis (e.g. Hanks 1992; Hanks 1995; Himmelmann 1996; Laury 1997).

6. In an unrelated discussion, Halliday (1994:197) observes that 'things are more highly organized than events', though he does not elaborate on the meaning of *organized* in this context (e.g. that these discursive phenomena—entities and events—are conceptually organized by speakers? intrinsically organized as individuated phenomena?).

7. Not included here are the three tokens with generic subjects.

8. See Thompson 1988 for a concise synopsis of Dixon's major semantic types.

9. The focus, or starting point, of these utterances is itself reified by being represented as a grammatical subject.

10. An example of this heterogeneity arose during coding of semantic noun type in these predicates. During the coding process it was necessary to modify Dixon's semantic classification system to include formal categories (e.g. *clause/event, empty noun, proform*); it was not possible to classify this group by meaning alone because of structural variation.

11. Most of the nominal complements coded *clause/event* are nominal relative clauses. They resemble *wh*-interrogatives.

12. *That* refers to a vacuum cleaner that uses disposable bags

13. Many of these *proform* tokens are utterances such as *what is this,* and *what's tha=t*.

14. The parts of these pairs are not typically adjacent in the conversation; there is usually a flutter of teasing banter that occurs between the question and the answer.

15. This group will be referred to as event-state-discourse.

16. Thirteen out of the total 16 predicates containing empty nouns occur with event-state-discourse subjects; only 3 items appear with *entity* referents.

17. Related to discussion in section 4.3.3.1. of the interactive uses of *that* with *value* adjectives, the subject *that* in these nominal relational utterances with clausal complements often refers to a previous participant's conversational contribution.

Chapter 5

Summaries and conclusions

5.1. Introduction

Like all studies in linguistics and in other disciplines as well, discussions and analyses in this work necessarily reflect the author's subjectivity. From the outset of investigating the structural expression of speaker stance in American English conversation, I have been keenly aware that neither the construction of a coding system and the coding of data nor the interpretation of distributional patterns found in the corpus are separate from influences from my own points of view, theoretical and otherwise. Just as in spontaneous conversation there is little passing of neutral, neatly specified information among participants, it is also the case that academic discussion of data and theory is itself situated (indexed or constructed) by speakers and writers in their presentations. Both spoken and written scholarly discourse are considerably less interactive than casual conversation. Nonetheless, even given the strong expectation for the communication of information in these genres, *transfer* of information does not occur independent of the participants. Even at a distance, writer and reader work from their points of view to negotiate context and sense of the language. In recognition of this process, I have attempted throughout the study to leave interactive *space* between my presentations and the reader's own experiences and insights by being mindful that the organization and substance of these analyses are best viewed—not as discrete products or packages—but as contributions to an ongoing dialog.

This final chapter provides general discussion and synthesis of the major ideas that have surfaced during the analyses. Section 5.2. surveys issues concerning the employment of usage-based methods in this work, including the importance of local structural patterns, and it focuses on the impact of the findings of this study on traditional perspectives on linguistic categories. Section 5.3. explores the central themes of subjectivity and objectivity in relationship to linguistic expression, including what I have called *covert subjectivity*. And finally, section 5.4. offers some preliminary ideas on the links between speaker point of view and interaction in English conversation.

5.2. Usage, linguistic patterning, and categories

In view of the importance of usage in the conventionalization and change of linguistic form, directions for analyses in this investigation were determined by frequency of grammatical and lexical expressions in the corpus.[1] This study was intentionally designed to be broad in scope in an effort to allow, as much as possible, unanticipated structural and functional patterns to emerge during analyses. Having access to a variety of formal and semantic properties of utterances in the form of coding categories facilitates more sensitive local analyses. For example, in examining the influence of humanness of subject on third person plural utterances compared to third person singular clauses in chapter 3, it was possible, given the flexibility of the coding system, to augment the analysis by including referentiality of the subject as a factor in the distribution of these expressions. This allowed for a more informative characterization of 3p utterances by illustrating that in conversation these subjects are more frequently generic than referentially specific.

Throughout this study, global distributional frequencies—those delineated by general, or high-level, linguistic parameters (e.g. tense, verb type, subject)—are not typically attributable to any one major structural trend. Rather, these larger patterns are usually found to be the result of several distinct local patterns. One of the most consistent findings in this corpus is the greater use of present tense than past tense predicates by speakers. The actual expressions contributing to this strong trend, however, are varied. For example, in the present tense we find first and second person singular subjects with verbs of cognition (fixed expressions with pragmatic function), 2s utterances as a group (mediated assertions), 3s nonhuman subjects with relational clauses (evaluations), and 3p clauses with generic subjects (generalizations). On the other hand, past tense predicates are much more constrained in their distribution in conversation. With the exception of 1s+verbs of cognition (e.g. *I thought*), past tense verbs are mainly found with 1s and 3s human subjects with material verbs.

By looking at local contexts, then, we see a functional generalization in tense distribution in English conversation: there is an association between the majority of present tense usages and various expressions of speaker point of view (e.g. epistemic expression, evaluation, generalization), and an association between past tense predicates and speakers' reporting on actors and events in the world. But it is also the case that in this corpus, English Present

and Past are neither distributionally symmetrical nor contrastive members of the category tense. These results are compatible with Bybee et al.'s (1994) cross-linguistic research on tense, aspect, and modality. Bybee et al. suggest that it is difficult to view present tense as a "tense" since it is not primarily used to express time reference; rather, present tense is general and is used in several types of expressions, such as habitual utterances and clauses with generic subjects (Bybee et al. 1994: 126). In a discussion of the English Progressive and Simple Present, Bybee et al. (1994: 148) also note "that grammatically expressed meaning categories do not necessarily have to form maximal contrasts". Given this kind of asymmetry of membership in grammatical categories that is revealed when looking at language in context, it may be useful for analysts to reconsider what it is exactly that we expect from our formal categories.

With the introduction of conversational data into the analytic process, linguistic distributions and categories can be investigated by studying contextual and local properties of utterances. For example, in the present work I have suggested that third person singular subject is not a coherent category in English conversation because utterances with these subjects exhibit a variety of formal and functional properties depending on the subject type (e.g. humanness), regardless of the fact that all 3s subjects trigger the same verb agreement.

Another example in which the use of conversational data elaborates characterizations of conventional linguistic categories is the relationship between subject and predicate. While *subject* and *predicate* are viewed as independent constituents in most grammatical treatments, we find that in English interactive discourse, they are not as paradigmatically autonomous as is often assumed. Distributional patterns of conversational usage indicate that subjects and verbs are not arbitrarily combined (see also Bybee & Scheibman 1999). Pragmatic demands, verb type, subject type, referentiality and animacy of subject, tense, and other linguistic features all affect the subject-verb combinations seen most frequently in conversation. Examples of subject-predicate cooccurrences commonly used by English speakers in interactive talk are: cognitive verbs with speech act participant subjects, value adjectives in predicates with nonentity 3s subjects, and referentially corresponding subject NPs and predicate nominals in relational clauses.

Another issue related to the coherence of grammatical categories raised in this study has to do with the role of pragmatics in ascertaining the meanings of

linguistic expressions in context. The traditional distinction between semantics (more or less synchronically invariant meanings of linguistic expressions) and pragmatics (meanings given by context and speakers' intentions) is in a sense parallel to the distinction made between referential language and expressive or indexical uses. Though the latter set of terms is typically used to describe more global aspects of language than the former set, both of these pairs presume separations between linguistic material (e.g. messages, autonomous meanings) and participants in context (e.g. speaker point of view, interactional demands). For example, in the analysis of semantic adjective types in predicate adjective constructions in chapter 4, I suggested that the meanings of value adjectives and physical property adjectives with nonhuman subjects are determined by the speaker's judgments. That is, the *semantic* meaning of lexical items is not separate from pragmatic or expressive use. What we see, then, when studying interactive discourse, is that maintaining a distinction between meanings of linguistic expressions as independent units and meanings of linguistic expressions in context can be restrictive, if not impossible.

There is an analytical tradition in linguistics of treating meanings of grammatical and lexical elements as referential, in the sense that they designate decontextualized events and entities in the world. For example, though verbal aspect is defined as "different ways of **viewing** the internal temporal constituency of a situation [bolding added]" (Comrie 1976: 3)—a characterization that only implicitly includes a speaker doing the viewing—most work on aspectual categories proceeds by assigning meanings that take the form of *descriptions of events* (e.g. progressive, iterative) with little or no elaboration of the speaker's *perspective* on events in context.

This situation regarding the convention of referentially describing verbal aspect contrasts with studies on modality, a phenomenon that is explicitly described as "the grammaticization of speakers' (subjective) attitudes and opinions" (Bybee et al. 1994: 176). However, despite the differential treatment of aspect and modality in the literature, it is not clear how speaker point of view is more influential when speakers use modal elements (e.g. to express degree of commitment to a proposition) than when they use aspectual categories (e.g. to indicate a temporal perspective on the viewing of events). Perhaps because of the referential bias in linguistic analysis, when defining grammatical classes the default labeling of meaning and meaning categories tends towards describing events separate from—rather than in terms of—the

speaker's point of view.[2] In contrast to this convention, results from the present study suggest that the most frequently found lexical and grammatical categories and elements in English conversation do not function referentially, but rather, subjectively (to varying degrees) to convey speaker stance.

We see, then, that analyses of structural patterns using conversational data highlight some discontinuities in our traditional understanding of categories. Because there is variation in pragmatic and formal patterns of conversational utterances in different local contexts, conventional grammatical categories (e.g. subject, tense, verb type) sometimes show a lack of functional coherence at a more global level (e.g. present tense). Given this variability in the distribution of linguistic categories in local contexts, one wonders how similar work using interactive data may influence broader typological studies which rely on the integrity of categories to make cross-linguistic comparisons of structural and semantic classes.

The inclusion of details of local patterning of structural and functional properties in conversation can potentially deepen our understanding of the structure of a given language. However, the introduction of interactive data puts added requirements on analysts, in particular that they be familiar enough with the language and culture they are working with to understand the pragmatics of interaction. This may be why for several years much work in conversation and grammar had been carried out by English-speaking analysts on English. Fortunately, more studies are being conducted on languages other than English (e.g. Iwasaki 1993; Laury 1997; Ono, Thompson, & Suzuki 2000).

Ochs (1996: 409) makes the point that many universal human behaviors (including language) may be due to commonalties in the ways that people structure and accommodate to social life, and are not necessarily the result of innate mechanisms. This may mean that the use of conversation in typological work would not only provide this field with data based on usage, but it would also broaden cross-linguistic study by introducing interactive and subjective motivations into these investigations.

An example of such an analytical correspondence between local description of conversational data and a typological phenomenon surfaced in the present study. We saw in chapter 3 that first person plural utterances in the database pattern along the lines of an inclusive/exclusive distinction, and that this difference is related to conventional ways English-speaking participants interact with one another in conversation. Therefore, even at this early stage,

one suspects that attention to local linguistic patterning will both elaborate and enhance cross-linguistic research.

5.3. Subjectivity and objectivity in English expression

As discussed in chapter 1, *linguistic subjectivity* has been very broadly defined in the literature to include speakers' perceptions and opinions in discourse as expressed by a myriad of linguistic forms (e.g. pronouns, lexical items, modal and deictic elements, grammatical morphemes, syntactic constructions, etc.). One of the goals at the outset of this project was to investigate the structure of utterances in conversation; the expectation was that the most frequent formal patterns would be those that convey speaker point of view, and I hope that in the previous chapters I have demonstrated that this is the case. However, as useful as this study may be in having highlighted the forms English speakers most commonly use in interactive discourse, it has also raised more questions. An important question is how these conversational data can contribute to a broader understanding of subjective expression in English.

Throughout the analyses in this work I have made reference to a variety of terms, such as *subjective, evaluative, personalizing, referential, descriptive*, and *conveying information* in characterizing and comparing utterances as regards their expression of speaker point of view. In surveying the utterance types in the data and their analyses, it becomes clear that these characterizations are relative ones and are best understood in terms of a continuum between *subjective* expression (i.e. structures and functions which to varying degrees are grammatically, lexically, or pragmatically anchored to the speaker) and *objective* expression (i.e. relatively unpersonalized utterances which typically report on activities of an entity subject).

In the analyses in chapters 3 and 4, several structural and functional elements were identified as contributing to the expression of speaker point of view. As a reminder to the reader, these general features are summarized below.

- *I*
- present tense (i.e. absence of past tense)
- modal elements (especially with material verbs)
- verbs of cognition

- referentially nonspecific, or nonentity, subjects (e.g. *event-state-discourse, generic*)[3]
- lexical adjectives in predicate adjective clauses
- in predicate nominal clauses, the identifying relation between NPs which is based on the speaker's conception of similarity
- intensifiers and modal adverbs

Note that these properties do not have equal status with respect to their contribution to subjective expression. For example, the presence of *I* does not necessarily subjectify an utterance on its own (e.g. as in past tense *I*+material verb utterances), nor does the presence of a present tense predicate. In other words, these features are listed separately for purposes of presentation only and should not be viewed as functioning independently, since the analyses in preceding chapters show that these structural elements tend to cluster in particular contexts.

Above, I characterized as subjective utterances in this corpus that are structurally and/or pragmatically anchored to the speaker. Given that the interaction between human beings and the use of language is complex, diversity in the form and function of subjective expression is not unexpected.[4] The most explicit and conventionally-recognized marker of subjectivity is the first person singular subject, *I* (Benveniste 1971). Recall that this is the most frequent subject type in the database. In the present tense, in combination with verbs of cognition and verbal process, *I* utterances appear as fixed expressions and overwhelmingly function epistemically in discourse. In these *I* utterances and others (e.g. *I* + material verbs with modal elements), the referent of *I* is the *speake*r. On the other hand, in clauses with 1s subjects with past tense material or perception predicates, the referent of *I* is not only the speaker but also the *subject* of the material process. That is, in these situations, *I* refers to the agent or experiencer with respect to the predicate, and only tacitly to the speaker within the discourse context. Utterances in which *I* refers to a participant or actor in an event are more objective than those in which *I* communicates from the speaker's consciousness (cf. Langacker 1990; Lyons 1982). Thus even the presence of a first person singular pronoun in a given utterance does not uniformly add to its subjectivity without the contribution of other structural features (e.g. modals, present tense, cognitive verbs, etc.).

In the appropriate contexts and combinations, then, 1s subjects, verbs of cognition, modal elements, and present tense structurally anchor an utterance to the speaker. Modal elements occurring with non–1s subjects also

mediate the propositionality, or informativity, of a clause. In utterances with modals, the focus is on some aspect of the speaker's stance concerning the accomplishment of the activity designated by the verb (e.g. *can shave, want to shave*) and not on the activity itself (e.g. *shaving*). When these elements appear with third person human subjects, the modal conveys the speaker's view of the subject's involvement—an example of what I have called *covert subjectivity* (see below).

A consistent finding characteristic of many utterances in the corpus is the presence of a subject that is referentially nonspecific. This includes *generic* subjects (e.g. clauses with *they*) and subjects whose referents are best described as being along a continuum of *event-state-discourse*. This latter group is typically expressed by *it* and *that*. Both generic and event-state-discourse subjects are anchored to the speaker in that, in both of these cases, discourse participants are delineating—rather than naming something in the real world —a group of people (using *they*) or an aspect or interpretation of the discourse (using *it* and *that*). This construction of discourse phenomena, as illustrated in the use of these subjects, is an inherently subjective and evaluative act on the part of the speaker.

Expression of subjectivity in relational clauses in the database manifests in two ways. In predicate adjective constructions, the relation between the subject, or starting point, and the adjective in the predicate is one of *evaluation*. In these usages, it is the pragmatics and semantics of the lexical adjective that convey the speaker's judgments and opinions. On the other hand, expression of point of view in predicate nominal clauses does not rely on the subjectivity of lexical items. In these constructions, the relation is *identifying*. But only in the most concrete cases is a specific entity being identified by the speaker (typically using the subject *this*). In most cases, the act of identifying, or viewing a subject and predicate nominal as similar, is a subjective process on the part of the speaker in that it is based on psychological and sociocultural experiences and expectations of participants. In other words, what is anchored to the speaker in predicate nominal utterances is a process of gauging similarity.

I suggested above that the subjectivity of an utterance is not an absolute characterization, but a continuous one. One way to assess the subjectivity of a locution is to note how overtly the speaker marks her contribution as point of view.[5] In English, the most explicit ownership of opinion and attitude is signaled through the use of the first person singular pronoun. To a lesser

degree, the presence of modal auxiliaries and relational clauses with value adjectives (e.g. *good, great*) may be interpreted by participants as expressing speaker stance. However, in utterances with third person nonentity subjects, in particular relational clauses, the construction of a discursive subject and its evaluation or "identification" are not structurally coded as expressing speaker point of view, even though these utterances may be as subjective as those containing *I*. These third person expressions are conventionally viewed as characterizing the world separate from the speaker. However, because of their rich subjective expression, they are not independent descriptions of entities and events. In using these frequent utterance types, speakers are projecting their evaluative discourse onto reality, in essence, expressing a *covert subjectivity*.

Given, then, the variety and complexity of structural expression of subjectivity in this corpus, what can be said about the role of objectivity in interaction? In the current study, the utterances that were found to express speaker point of view least (or most covertly) are those with subjects referring to entities, in particular human beings, appearing with past tense material predicates, and also oblique predicates. The structural and functional properties of these utterance types—human subject, past tense, a (more or less) propositional predicate—are all conceptually referential. That is, human beings are individuated in physical form, past tense perspective tends to facilitate construal of entire events as discrete units, and material verbs are often those lexical items that designate activities. With respect to the data in this corpus, then, it appears that notions of objectivity go hand in hand with participants' conception of the world as reified and separate from the speaker.

Finally, recall that in interactive discourse, the frequency of these more referential utterances is low, and they tend to cluster in conversational narratives. Kerby (1991: 59) suggests that narrative "articulates what is of value to us and why, for it essentially defines (in the first instance) who we are and what we want". Therefore, even these more objective locutions are not without subjective import; choosing what to say or what to react to is the most basic expression of subjectivity in language.

5.4. Subjectivity and interaction

As noted above, a contribution this study makes is to the collection of individual language investigations of linguistic structure in conversation for use in typological inquiries. Another area of research suggested by the findings in this study is that of examining the links between expression of speaker point of view and interaction in English discourse.

One of the limitations of the present work is that the basic unit of analysis is the utterance. This method—while conducive to investigating structural relations between subject and predicate (the organizing focus for the analyses in this study)—is not particularly well suited to the analysis of conversation. This is because individual conversational utterances are not separate from one another; rather, they are strongly contextualized. Participants are interactively engaged with one another in their talk; consequently, viewing utterances in isolation fails to represent the relations among discursive contributions.[6] Given the ubiquitous expression of speaker point of view and personalization of utterance in English conversation, however, there is good reason to believe that these phenomena are closely linked to interactional demands and processes.

Important structural sites for the convergence of subjectivity and interaction in English are second person singular utterances. As discussed in chapter 3, section 3.4., second person singular clauses are propositionally mediated in a variety of ways: (1) there are fewer past tense predicates occurring with 2s subjects than with any other human subject, (2) 2s utterances have the highest proportion of modals of all subject types, (3) 2s clauses are often preceded by evaluative or epistemic clauses (e.g. ***that's good*** *you're getting r- good rest.*), (4) 2s utterances are frequently questions, (5) in the present tense, the subject *you* commonly has generic reference (e.g. *it smells like* ***you****'re walking past a d=umpster.*), and (6) 26 percent of all 2s clauses are instances of the expression, *you know*.

Structurally, then, 2s utterances in this corpus are highly personalized by the speaker. Functionally, forms with generic *you* subjects convey opinions speakers assume might be shared by other participants. The highly frequent expression, *you know*, is used to encourage agreement with or ratification of the speaker's own assessments. Other types of mediation of assertion in 2s clauses (e.g. the high presence of modals, few past tense reports) are due to negative politeness strategies (not imposing one's will on another person). In

chapter 3, I characterized the personalized and mediated character of 2s utterances in conversation as interactively, or empathetically, subjective. Because of the considerable social and cultural conventions concerning how people interact with other people, it is not surprising that the personalizing properties of the majority of second person singular utterances are used by speakers to situate their utterances in relation to another speech act participant. This, then, provides a specific example of how subjective elements have interactive consequences.

It is not serendipitous that this study of subjectivity emerged from studies of conversation. In interactive discourse, English-speaking participants are more likely to explicitly mark their contributions as based in their own viewpoints than in more formal genres, which tend towards a covertly subjective style (e.g. Biber 1988). In higher registers, the indexing and construction of evaluation and situation using third person subjects endows the speaker's/writer's points of view with authority.[7] As discussed in chapter 4 for 3s relational clauses in conversation, the attribution of qualities and existence is typically interpreted and treated as making reference to external, nondiscursively-bound realties (objectified). It is also the case that in these more formal language styles, relative to conversation, there is less interaction among participants.

In mainstream American English-speaking culture, objectification of expression in higher registers or authoritative speech (such as avoiding first person singular pronouns or intensifiers, e.g. replacing a construction like *I'm sure* with *Clearly*) stabilizes reality by presuming consensus among participants. In such language, there is little acknowledged need for a cache of linguistic resources used to express different degrees of speaker reaction or commitment. The gap between speaker recognition of the existence of differing realities (which are negotiated regularly in casual conversation using conventions of politeness, for example) and linguistically encoding a particular reality (which is also a subjective act) is narrowed in propositionally-dense language. In objectified expression, there is little elasticity between the speaker's discursive contribution and her take on it—hence, less room for another person's input. This is in contrast to English conversation in which linguistic structure is replete with explicit markers indicating that there is *space* (acknowledgment) between speakers and their utterances. And it is in that *space* that resides the potential for negotiation of meaning. It is in this *space* that

haggling of sense and its representation occurs among participants, and these activities are facilitated by the understanding of linguistic expression as point of view.

Notes

1. It is relative frequency that is of interest here. For example, there are some analyses in this study that are based on small sets of linguistic elements in absolute terms (e.g. first person plural utterances). Consistent with other analyses, however, it is the most frequent items in these groups that are attended to.
2. However, in a recent paper on Koyukon, a northern Athabaskan language, Axelrod (2000) finds that markers of aspect and mood are not categorically distinct from one another. In part, the continuity found among members of these two grammatical classes has to do with the expression of speaker point of view.
3. Second person singular subjects will be discussed below in section 5.4.
4. These findings are compatible with the broad characterizations found in the literature on linguistic subjectivity (see chapter 1, section 1.2.1.).
5. One could also describe the degree of subjectivity of a given utterance by counting, or measuring in some way, its structural features (e.g. the presence of *I*, value adjectives, multiple uses of intensifiers).
6. One of the challenges of analyzing the data in this corpus was continually having to go back and forth between utterances in the coded database and the conversational transcripts and audiotapes.
7. For discussions of social and ideological influences of texts, see Fairclough (1995) and other works in Critical Discourse Analysis.

Appendix A: Transcription symbols

(from Du Bois et al. 1993:68–69, 88–89)

Units		
	Intonation unit	{carriage return}
	Truncated intonation unit	--
	Word	{space}
	Truncated word	–
	Speaker identity/turn start	:
	Speech overlap	[ ]
Transitional Continuity		
	Final	.
	Continuing	,
	Appeal	?
Terminal Pitch Direction		
	Fall	\
	Rise	/
	Level	_
Accent and Lengthening		
	Primary accent	^
	Secondary accent	’
	High booster	!
	Low booster	;
	Lengthening	=
Pause		
	Long	…(N)
	Medium	…
	Short	..
	Latching	(0)
Vocal Noises		
	Vocal noises	( )
	Inhalation	(H)
	Exhalation	(Hx)
	Glottal stop	%
	Laughter	@
Quality		
	Quality	<Y Y>
	Laugh quality	<@ @>
	Quotation quality	<Q Q>
	Forte: loud	<F F>
	Piano: soft	<P P>

Crescendo: gradually louder	<CR CR>
Diminuendo: gradually softer	<DIM DIM>
Higher pitch level	<HI HI>
Lowered pitch level	<LO LO>
Widened pitch range	<W W>
Narrowed pitch range	<N N>
parenthetical prosody	<PAR PAR>
Allegro: rapid speech	<A A>
Lento: slow speech	<L L>
Marcato: each word emphasized	<MRC MRC>
Whispered	<WH WH>
Breathy	<BR BR>
Creaky	<CRK CRK>
Phonetics	
Phonetic/phonemic transcription	(/ /)
Transcriber's Perspective	
Researcher's comment	(( ))
Uncertain hearing	<X X>
Indecipherable syllable	X
Duration	(N)
Intonation unit continued	&
Accent unit boundary	|
Embedded intonation unit	<| |>
False start	< >
Code switching	<L2 L2>
Nontranscription	$

Appendix B: Intermediate function verbs in the database

afford to
be able to
be going to
be gonna
be supposed to
begin
better
continue on
expect (NP) to
get
get to
go
got to
gotta
have (NP)
have to
keep
let (NP)
let's
like to
make (NP)
need
ought to
rather
start
stop
try and
try to
used to
wanna
want (NP) to
want to

References

Austin, J.L. 1997. *How to do things with words*. (2nd ed.). Cambridge, MA: Harvard University Press.

Axelrod, M. 2000. *Categories and oppositions: Tense, aspect, and modality in Koyukon Athabaskan.* Paper presented at the Linguistic Society of America 2000 Annual Meeting, Chicago, IL, January 5–7.

Benveniste, E. 1971. *Problems in general linguistics*. (Translated by Mary Elizabeth Meek) Coral Gables, FL: University of Miami Press.

Biber, D. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, D. and Finegan, E. 1989. "Styles of stance in English: Lexical and grammatical marking of evidentiality and affect". *Text* 9: 93–124.

Bolinger, D. 1972. *Degree words*. The Hague: Mouton.

Bolinger, D. 1977. *Meaning and form*. London/New York: Longman.

Boyland, J.T. 1996. *Morphosyntactic change in progress: A psycholinguistic approach.* Ph.D. dissertation, University of California, Berkeley.

Bybee, J. 1985. "On the nature of grammatical categories: A diachronic perspective". In *Proceedings of the Second Eastern States Conference on Linguistics.*, S. Choi (ed), 17–34.

Bybee, J. 1998. "The emergent lexicon". *Chicago Linguistic Society* 34*: The Panels*: 421–435.

Bybee, J. To appear. "Mechanisms of change in grammaticization: The role of frequency". In *Handbook of historical linguistics*, B. Joseph and R. Janda (eds), Oxford: Blackwell.

Bybee, J., Perkins, R. and Pagliuca, W. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.

Bybee, J. and Scheibman, J. 1999. "The effect of usage on degrees of constituency: The reduction of *don't* in English". *Linguistics* 37(4): 575–596.

Carey, K. 1995. "Subjectification and the development of the English Perfect". In *Subjectivity and subjectivization: Linguistic perspectives*, D. Stein and S. Wright (eds), 83–102. Cambridge: Cambridge University Press.

Chafe, W. 1988. "Linking intonation units in spoken English". In *Clause combining in grammar and discourse*, J. Haiman and S.A. Thompson (eds), 1–27. Amsterdam/Philadelphia: John Benjamins.

Chafe, W. 1993. "Prosodic and functional units of language". In *Talking data: Transcription and coding in discourse research*, J.A. Edwards and M.D. Lampert (eds), 33–43. Hillsdale, NJ: Lawrence Erlbaum.

Chafe, W. 1994. *Discourse, consciousness, and time*. Chicago: University of Chicago Press.

Clark, H.H. 1992. *Arenas of language use*. Chicago: University of Chicago Press.

Clark, H.H. 1996. *Using language*. Cambridge: Cambridge University Press.

Coates, J. 1996. *Women talk: Conversation between women friends*. Oxford: Blackwell.

Comrie, B. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.
Croft, W. 1995. "Intonation units and grammatical structure". *Linguistics* 33:839–882.
Dahl, Ö. 2000. "Egophoricity in discourse and syntax". *Functions of Language* 7:37–77.
Dixon, R.M.W. 1977. "Where have all the adjectives gone?". *Studies in Language* 1:19–80.
Dixon, R.M.W. 1991. *A new approach to English grammar, on semantic principles*. Oxford: Clarendon Press.
Du Bois, J., Schuetze-Coburn, S., Cumming, S. and Paolino, D. 1993. "Outline of discourse transcription". In *Talking data: Transcription and coding in discourse research*, J.A. Edwards and M.D. Lampert (eds), 45–89. Hillsdale, NJ: Lawrence Erlbaum.
Duranti, A. 1994. *From grammar to politics: Linguistic anthropology in a Western Samoan village*. Berkeley: University of California Press.
Edelsky, C. 1981. "Who's got the floor?". *Language in Society* 10:383–421.
Fairclough, N. 1995. *Critical discourse analysis: The critical study of language*. London/New York: Longman.
Ford, C.E. 1993. *Grammar in interaction: Adverbial clauses in American English conversations*. Cambridge: Cambridge University Press.
Ford, C.E. and Thompson, S.A. 1996. "Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns". In *Interaction and grammar*, E. Ochs, E.A. Schegloff and S.A. Thompson (eds), 134–184. Cambridge: Cambridge University Press.
Fox, B.A. 1987. *Anaphora and the structure of discourse*. Cambridge: Cambridge University Press.
Fox, B.A. 1995. "The category 'S' in English conversation". In *Discourse grammar and typology*, A. Werner, T. Givón and S.A. Thompson (eds), 153–178. Amsterdam/Philiadelphia: John Benjamins.
Fujii, N. and Ono, T. 2000. "The occurrence and non-occurrence of the Japanese direct object marker *o* in conversation". *Studies in Language* 24(1):1–39.
Gell-Mann, M. 1994. *The quark and the jaguar: Adventures in the simple and the complex*. NY: W. H. Freeman.
Givón, T. 1990. *Syntax: A functional-typological introduction: vol. II*. Amsterdam/Philadelphia: John Benjamins.
Goodwin, C. 1981. *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
Greenberg, J.H. 1985. "Some iconic relationships among place, time, and discourse deixis". In *Iconicity in syntax*, J. Haiman (ed) 271–287. Amsterdam/Philadelphia: John Benjamins.
Haiman, J. 1995. "Grammatical signs of the divided self: A study of language and culture". In *Discourse grammar and typology: Papers in honor of John W. M. Verhaar*, W. Abraham, T. Givón and S.A. Thompson (eds), 213–234. Amsterdam/Philadelphia: John Benjamins.
Haiman, J. 1998. *Talk is cheap: Sarcasm, alienation, and the evolution of language*. New York: Oxford University Press.

Haiman, J. and Thompson, S.A. (eds.). 1988. *Clause combining in grammar and discourse*. Amsterdam/Philadelphia: John Benjamins.

Halliday, M.A.K. 1994. *An introduction to Functional Grammar*. (2nd ed.). London: Edward Arnold.

Hanks, W.F. 1990. *Referential Practice: Language and lived space among the Maya*. Chicago: University of Chicago Press.

Hanks, W.F. 1992. "The indexical ground of deictic reference". In *Rethinking context: Language as an interactive phenomenon*, A. Duranti and C. Goodwin (eds), 43–76. Cambridge: Cambridge University.

Hanks, W.F. 1995. "When utterances become objects". *Journal of Contemporary Legal Issues* 6: 173–186 *(The crisis of text: Issues in the constitution of authority)*.

Harris, Z.S. 1951. *Methods in structural linguistics*. Chicago: University of Chicago Press.

Hatcher, A.G. 1951. "The use of the Progressive form in English". *Language* 27: 254–280.

Himmelmann, N.P. 1996. "Demonstratives in narrative discourse: A taxonomy of universal uses". In *Studies in anaphora*, B. Fox (ed), 205–254. Amsterdam/Philadelphia: John Benjamins.

Hopper, P.J. 1987. "Emergent grammar". *Berkeley Linguistics Society* 13: 139–157.

Hopper, P.J. 1991. "Dispersed verbal predicates in vernacular written narrative". *Berkeley Linguistics Society* 17: 402–413.

Hopper, P.J. 1997. "When 'grammar' and discourse clash: The problem of source conflicts". In *Essays on language function and language type*, J. Bybee, J. Haiman and S.A. Thompson (eds), 231–247. Amsterdam/Philadelphia: John Benjamins.

Iwasaki, S. 1993. *Subjectivity in grammar and discourse: Theoretical considerations and a case study of Japanese spoken discourse*. Amsterdam/Philadelphia: John Benjamins.

Jakobson, R. 1960. "Concluding statement: Linguistics and poetics". In *Style in language*, T.A. Sebeok (ed) 350–377. NY: John Wiley & Sons and The Technology Press of Massachusetts Institute of Technology.

Jespersen, O. 1965. *The philosophy of grammar*. New York: W. W. Norton and Company.

Johnstone, B. 2002. *Discourse analysis*. Malden, MA: Blackwell.

Kärkkäinen, E. 1998. *The marking and interactional functions of epistemic stance in American English conversation*. Ph.D. dissertation, University of California, Santa Barbara.

Kerby, A.P. 1991. *Narrative and the self*. Bloomington and Indianapolis: Indiana University.

Krug, M. 1998. "String frequency: A cognitive motivating factor in coalescence, language processing, and linguistic change". *Journal of English Linguistics* 26: 286–320.

Kuno, S. and Kaburaki, E. 1977. "Empathy and syntax". *Linguistic Inquiry* 8(4): 627–672.

Lakoff, G. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lampert, M.D. and Ervin-Tripp, S.M. 1993. "Structured coding for the study of language and social interaction". In *Talking data: Transcription and coding in discourse research*, J.A. Edwards and M.D. Lampert (eds), 169–206. Hillsdale, NJ: Lawrence Erlbaum Associates.

Langacker, R.W. 1987. *Foundations of Cognitive Grammar, Vol. I: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Langacker, R.W. 1990. "Subjectification". *Cognitive Linguistics* 1:5–38.

Langacker, R.W. 1991. *Concept, image, and symbol: The cognitive basis of grammar*. Berlin/New York: Mouton de Gruyter.

Laury, R. 1997. *Demonstratives in interaction: The emergence of a definite article in Finnish*. Amsterdam/Philadelphia: John Benjamins.

Levin, B. 1993. *English verb classes: A preliminary investigation*. Chicago: Chicago University Press.

Levinson, S.C. 1983. *Pragmatics*. Cambridge: Cambridge University.

Lyons, J. 1977. *Semantics*. Cambridge: Cambridge University Press.

Lyons, J. 1982. "Deixis and subjectivity: *Loquor, ergo sum*?". In *Speech, place, and action: Studies in deixis and related topics*, R.J. Jarvella and W. Klein (eds), 101–124. New York: John Wiley & Sons.

Lyons, J. 1994. "Subjecthood and subjectivity". In *Subjecthood and subjectivity: The status of the subject in linguistic theory*, M. Yaguello (ed), 9–17. Paris: Ophrys.

MacWhinney, B. 1977. "Starting points". *Language* 53:152–168.

Malinowski, B. 1923. "The problem of meaning in primitive languages". In *The meaning of meaning: A study of the influence of language upon thought and of the science of symbolism*, C.K. Ogden and I.A. Richards (eds), 296–336. New York: Harcourt Brace Jovanovich.

Martinet, A. 1991. "Expressivité". *Linguistique* 27(1):3–14.

Matthiessen, C. and Thompson, S.A. 1988. "The structure of discourse and 'subordination'". In *Clause combining in grammar and discourse*, J. Haiman and S.A. Thompson (eds), 275–329. Amsterdam/Philadelphia: John Benjamins.

McCawley, J.D. 1988. *The syntactic phenomenon of English, Vol. 1*. Chicago: University of Chicago Press.

Ochs, E. 1979. "Transcription as theory". In *Developmental pragmatics*, E. Ochs and B.B. Schieffelin (eds), New York: Academic Press.

Ochs, E. 1996. "Linguistic resources for socializing humanity". In *Rethinking linguistic relativity*, J.J. Gumperz and S.C. Levinson (eds), Cambridge: Cambridge University Press.

Ogden, C.K. and Richards, I.A. 1923. *The meaning of meaning: A study of the influence of language upon thought and of the science of symbolism*. New York: Harcourt Brace Jovanovich.

Ono, T. and Thompson, S.A. 1995. "What can conversation tell us about syntax?". In *Alternative linguistics: Descriptive and theoretical modes*, P.W. Davis (ed), 213–271. Amsterdam/Philadelphia: John Benjamins.

Ono, T., Thompson, S.A. and Suzuki, R. 2000. "The pragmatic nature of the so-called subject marker *ga* in Japanese: Evidence from conversation". *Discourse Studies* 1:55–84.

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. 1985. *A comprehensive grammar of the English language*. London: Longman.

Reddy, M.J. 1993. "The conduit metaphor: A case of frame conflict in our language about language". In *Metaphor and thought*, (2nd ed.), A. Ortony (ed), Cambridge: Cambridge University Press.

Rommetveit, R. 1974. *On message structure: A framework for the study of language and communication*. London: John Wiley & Sons.

Sacks, H., Schegloff, E.A. and Jefferson, G. 1974. "A simplest systematics for the organization of turn-taking for conversation". *Language* 50:696–735.

Schegloff, E.A. 1972. "Sequencing in conversational openings". In *Directions in sociolinguistics: The ethnography of communication*, J.J. Gumperz and D. Hymes (eds), 346–380. New York: Holt, Rinehart and Winston.

Schegloff, E.A. 1996. "Turn organization: One intersection of grammar and interaction". In *Interaction and grammar*, E. Ochs, E.A. Schegloff and S.A. Thompson (eds), 52–133. Cambridge: Cambridge University Press.

Scheibman, J. 2000. "*I dunno* ... A usage-based account of the phonological reduction of *don't* in American English conversation". *Journal of Pragmatics* 32(1):105–124.

Scheibman, J. 2001. "Local patterns of subjectivity in person and verb type in American English conversation". In *Frequency and the emergence of linguistic structure*, J.L. Bybee and P. Hopper (eds), 61–89. Amsterdam/Philadelphia: John Benjamins.

Schiffrin, D. 1987. *Discourse markers*. Cambridge: Cambridge University Press.

Schiffrin, D. 1994. *Approaches to discourse*. Cambridge, MA: Blackwell.

Shibatani, M. 1990. "Japanese". In *The world's major languages*, B. Comrie (ed), 855–880. New York: Oxford University Press.

Silverstein, M. 1976. "Shifters, linguistic categories, and cultural description". In *Meaning in anthropology*, K.H. Basso and H.A. Selby (eds), 11–55. Albuquerque: University of New Mexico Press.

Stein, D. 1995. "Subjective meanings and the history of inversions in English". In *Subjectivity and subjectivization: Linguistic perspectives*, D. Stein and S. Wright (eds), 129–150. Cambridge: Cambridge University Press.

Tao, H. 2001. "Discovering the usual with corpora: The case of *remember*". In *Corpus linguistics in North America: Selections from the 1999 Symposium*, R. Simpson and J. Swales (eds), 116–114. Ann Arbor: University of Michigan Press.

Tao, L. 1996. "Topic discontinuity and zero anaphor in Chinese discourse: Cognitive strategies in discourse processing". In *Studies in anaphora*, B. Fox (ed), 487–513. Amsterdam/Philadelphia: John Benjamins.

Thompson, S.A. 1988. "A discourse approach to the cross-linguistic category 'adjective'". In *Explaining language universals*, J. Hawkins (ed), 167–185. Oxford: Blackwell.

Thompson, S.A. and Hopper, P.J. 2001. "Transitivity, clause structure, and argument structure: Evidence from conversation". In *Frequency and the emergence of linguistic structure*, J. Bybee and P. Hopper (eds), 27–60. Amsterdam/Philadelphia: John Benjamins.

Thompson, S.A. and Mulac, A. 1991. "A quantitative perspective on the grammaticization of epistemic parentheticals in English". In *Approaches to grammaticalization, Volume II: Focus on types of grammatical markers*, E.C. Traugott and B. Heine (eds), 313–329. Amsterdam/Philadelphia: John Benjamins.

Traugott, E.C. 1988. "Pragmatic strengthening and grammaticalization". *Berkeley Linguistics Society* 14:406–416.

Traugott, E.C. 1989. "On the rise of epistemic meanings in English: An example of subjectification in semantic change". *Language* 65:31–55.

Traugott, E.C. 1995. "Subjectification in grammaticalisation". In *Subjectivity and subjectivization: Linguistic perspectives*, D. Stein and S. Wright (eds), 31–54. Cambridge: Cambridge University Press.

Traugott, E.C. and König, E. 1991. "The semantics-pragmatics of grammaticalization revisited". In *Approaches to grammaticalization, Volume I: Focus on theoretical and methodological issues*, E.C. Traugott and B. Heine (eds), 189–218. Amsterdam/Philadelphia: John Benjamins.

Tsui, A.B.M. 1991. "The pragmatic functions of *I don't know*". *Text* 11:607–622.

Vendler, Z. 1967. *Linguistics in philosophy*. Ithaca, NY: Cornell University.

Verhagen, A. 1995. "Subjectification, syntax, and communication". In *Subjectivity and subjectivization: Linguistic perspectives*, D. Stein and S. Wright (eds), 103–128. Cambridge: Cambridge University Press.

Vološinov, V.N. 1973. *Marxism and the philosophy of language*. (Translated by Ladislav Matejka and I.R. Titunik) Cambridge, MA: Harvard University Press.

Weber, E. 1993. *Varieties of questions in English conversation*. Amsterdam/Philadelphia: John Benjamins.

Wright, S. 1995. "Subjectivity and experiential syntax". In *Subjectivity and subjectivization: Linguistic perspectives*, D. Stein and S. Wright (eds), 151–172. Cambridge: Cambridge University Press.

Index

In the series STUDIES IN DISCOURSE AND GRAMMAR (SiDaG) the following titles have been published:

1. GELUYKENS, Ronald: *From Discourse Process to Grammatical Construction: On Left-Dislocation in English*. Amsterdam/Philadelphia, 1992.
2. IWASAKI, Shoichi: *Subjectivity in Grammar and Discourse: Theoretical Considerations and a Case Study of Japanese Spoken Discourse*. Amsterdam/Philadelphia, 1993.
3. WEBER, Elizabeth G.: *Varieties of Questions in English Conversation*. Amsterdam/Philadelphia, 1993.
4. DOWNING, Pamela: *Numerical Classifier Systems: The Case of Japanese*. Amsterdam/Philadelphia, 1996.
5. TAO, Hongyin: *Units in Mandarin Conversation: Prosody, Discourse, and Grammar*. Amsterdam/Philadelphia, 1996.
6. DORGELOH, Heidrun: *Inversion in Modern English: Form and function*. Amsterdam/Philadelphia, 1997.
7. LAURY, Ritva: *Demonstratives in Interaction. The emergence of a definite article in Finnish.* Amsterdam/Philadelphia, 1997.
8. MORI, Junko: *Negotiating Agreement and Disagreement in Japanese. Connective expressions and turn construction.* Amsterdam/Philadelphia, 1999.
9. HELASVUO, Marja-Liisa: *Syntax in the making: The emergence of syntactic units in Finnish conversation.* Amsterdam/Philadelpia, 2001.
10. SELTING, Margret and Elizabeth COUPER-KUHLEN (eds.): *Studies in Interactional Linguistics.* Amsterdam/Philadelpia, 2001.
11. SCHEIBMAN, Joanne: *Point of View and Grammar. Structural patterns of subjectivity in American English conversation.* Amsterdam/Philadelpia, 2002.
12. HAYASHI, Makoto: *Joint Utterance Construction in Japanese Conversation.* Amsterdam/Philadelpia, 2003.
13. ENGLEBRETSON, ROBERT: *Searching for Structure. The problem of complementation in colloquial Indonesian conversation.* Amsterdam/Philadelpia, 2003.
14. DU BOIS, John W., Lorraine E. KUMPF and William J. ASHBY (eds.): *Preferred Argument Structure. Grammar as architecture for function.* Amsterdam/Philadelpia, 2003.