

PROCESSING
COGNITIVE

Methods in Cognitive Linguistics

Edited by
Monica Gonzalez-Marquez,
Irene Mittelberg,
Seana Coulson and
Michael J. Spivey

Methods in Cognitive Linguistics is an introduction to empirical methodology for language researchers. Intended as a handbook to exploring the empirical dimension of the theoretical questions raised by Cognitive Linguistics, the volume presents guidelines for employing methods from a variety of intersecting disciplines, laying out different ways of gathering empirical evidence. The book is divided into 5 sections. *Methods and Motivations* provides the reader with the preliminary background in scientific methodology and statistics. The sections on *Corpus and Discourse Analysis*, and *Sign Language and Gesture* describe different ways of investigating usage data. *Behavioral Research* describes methods for exploring mental representation, simulation semantics, child language development, and the relationships between space and language, and eye movements and cognition. Lastly, *Neural Approaches* introduces the reader to ERP research and to the computational modeling of language.

"This book is indispensable for all researchers, whatever their discipline, working on language and cognition. Monica Gonzalez-Marquez has performed a great service for the research community."

Chris Sinha, *University of Portsmouth, England*

"The first single volume collection of its kind. *Methods in Cognitive Linguistics: Ithaca* provides an indispensable overview of the state of the art and new directions, bringing together some of the most influential and innovative scholars working in the field today. This represents indispensable reading for students and established scholars alike, and will be required reading on courses I teach in cognitive linguistics."

Professor Vyvyan Evans, author of *The Structure of Time*, and co-author of *Cognitive Linguistics: An Introduction*

ISBN 978 90 272 2372 2



HUMAN COGNITIVE PROCESSING is a forum for interdisciplinary research on the nature and organization of the cognitive systems and processes involved in speaking and understanding natural language (including sign language), and their relationship to other domains of human cognition, including general conceptual or knowledge systems and processes (the language and thought issue), and other perceptual or behavioral systems such as vision and non-verbal behavior (e.g. gesture). 'Cognition' should be taken broadly, not only including the domain of rationality, but also dimensions such as emotion and the unconscious. The series is open to any type of approach to the above questions (methodologically and theoretically) and to research from any discipline, including (but not restricted to) different branches of psychology, artificial intelligence and computer science, cognitive anthropology, linguistics, philosophy and neuroscience. It takes a special interest in research crossing the boundaries of these disciplines.

Editors

Marcelo Dascal, *Tel Aviv University*

Raymond W. Gibbs, *University of California at Santa Cruz*

Jan Nuyts, *University of Antwerp*

Editorial address

Jan Nuyts, University of Antwerp, Dept. of Linguistics (GER),
Universiteitsplein 1, B 2610 Wilrijk, Belgium.

E-mail: jan.nuyts@ua.ac.be

Editorial Advisory Board

Melissa Bowerman, *Nijmegen*; Wallace Chafe, *Santa Barbara, CA*;

Philip R. Cohen, *Portland, OR*; Antonio Damasio, *Iowa City, IA*;

Morton Ann Gernsbacher, *Madison, WI*; David McNeill, *Chicago, IL*;

Eric Pederson, *Eugene, OR*; François Recanati, *Paris*;

Sally Rice, *Edmonton, Alberta*; Benny Shanon, *Jerusalem*;

Lokendra Shastri, *Berkeley, CA*; Dan Slobin, *Berkeley, CA*;

Paul Thagard, *Waterloo, Ontario*

Volume 18

Methods in Cognitive Linguistics

Edited by Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson and
Michael J. Spivey

re
in
d
al
at
r-
ly
d
re
y
y,
s,
re

Methods in Cognitive Linguistics

Edited by

Monica Gonzalez-Marquez

Irene Mittelberg

Cornell University

Seana Coulson


UCSD

Michael J. Spivey

Cornell 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

Methods in cognitive linguistics / edited by Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson and Michael J. Spivey.
p. cm. (Human Cognitive Processing, ISSN 1387-6724 ; v. 18)
Includes bibliographical references and indexes.
1. Cognitive grammar--Research--Methodology. I. Gonzalez-Marquez, Monica.

P165.M47 2006

415--dc22

2006050840

ISBN 90 272 2371 8 (Hb; alk. paper)

ISBN 90 272 2372 6 (Pb; alk. paper)

© 2007 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

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

mgm ... for howie, claud and lari
im ... for wolfgang settekorn
sc ... for sjb
mjs ... for christopher mckinstry

r

ls

Table of contents

| | |
|--|------|
| Contributors | VIII |
| Acknowledgements | IX |
| Foreword <i>Leonard Talmy</i> | XI |
| Introduction: The many faces of research in Cognitive Linguistics <i>The Editors</i> | XXII |
| I. Methods and motivations | |
| Why cognitive linguists should care more about empirical methods <i>Raymond W. Gibbs, Jr.</i> | 2 |
| They actually said that? An introduction to working with usage data through discourse and corpus analysis <i>Irene Mittelberg, Thomas A. Farmer and Linda R. Waugh</i> | 19 |
| An introduction to experimental methods for language researchers <i>Monica Gonzalez-Marquez, Raymond B. Becker and James E. Cutting</i> | 53 |
| Inferential statistics in the context of empirical cognitive linguistics <i>Rafael Núñez</i> | 87 |
| II. Corpus and discourse analysis | |
| Multiple empirical approaches to a complex analysis of discourse <i>Linda R. Waugh, Bonnie Fonseca-Greber, Caroline Vickers and Betil Eröz</i> | 120 |
| A case for a cognitive corpus linguistics <i>Stefan Grondelaers, Dirk Geeraerts and Dirk Speelman</i> | 149 |
| III. Sign language and gesture | |
| Empirical methods in signed language research <i>Sherman Wilcox and Jill P. Morford</i> | 171 |
| Looking at space to study mental spaces: Co-speech gesture as a crucial data source in cognitive linguistics <i>Eve Sweetser</i> | 201 |

| | | |
|------|--|-----|
| | Methodology for multimodality: One way of working with speech and gesture data | 225 |
| | <i>Irene Mittelberg</i> | |
| | IV. Behavioral research | |
| | Experimental methods for studying language and space | 250 |
| | <i>Laura A. Carlson and Patrick L. Hill</i> | |
| VIII | Experimental methods for simulation semantics | 277 |
| | <i>Benjamin Bergen</i> | |
| IX | Experimental methods for studying the mental representation of language | 302 |
| XI | <i>Uri Hasson and Rachel Giora</i> | |
| | Eye movements in language and cognition: A brief introduction | 323 |
| XXII | <i>Daniel C. Richardson, Rick Dale and Michael J. Spivey</i> | |
| | Speaking for the wordless: Methods for studying the foundations of cognitive linguistics in infants | 345 |
| | <i>Amanda Brandone, Roberta Michnick Golinkoff, Rachel Pulverman, Mandy J. Maguire, Kathy Hirsh-Pasek, and Shannon M. Pruden</i> | |
| 2 | Experimental study of first and second language morphological processing | 367 |
| | <i>Kira Gor</i> | |
| 19 | V. Neural approaches | |
| 53 | Electrifying results: ERP data and cognitive linguistics | 400 |
| | <i>Seana Coulson</i> | |
| 87 | Bridging language with the rest of cognition: Computational, algorithmic and neurobiological issues and methods | 424 |
| | <i>Shimon Edelman</i> | |
| | Index | 446 |
| 120 | | |
| 149 | | |
| 171 | | |
| 201 | | |

Contributors

Leonard Talmy
State University of New York, Buffalo

Raymond Gibbs
University of California, Santa Cruz

Irene Mittelberg
European University Viadrina

Thomas Farmer
Cornell University

Linda R. Waugh
University of Arizona

Mónica González-Márquez
Cornell University

Raymond B. Becker
University of California, Merced

James Cutting
Cornell University

Rafael Núñez
University of California, San Diego

Bonnie Fonseca-Greber
Bowling Green State University

Caroline Vickers
California State University
San Bernardino

Betil Eröz
Middle East Technical University

Stefan Grondelaers
Université Libre de Bruxelles &
Radboud University Nijmegen

Dirk Geeraerts
University of Leuven

Dirk Speelman
University of Leuven

Sherman Wilcox
University of New Mexico

Jill Morford
University of New Mexico

Eve Sweetser
University of California, Berkeley

Laura Carlson
University of Notre Dame

Patrick L. Hill
University of Notre Dame

Ben Bergen
University of Hawai'i, Manoa

Uri Hasson
The University of Chicago

Rachel Giora
Tel Aviv University

Daniel C. Richardson
University of Reading

Rick Dale
University of Memphis

Michael Spivey
Cornell University &
University of California, Merced

Amanda Brandone
University of Michigan

Roberta Michnick Golinkoff
University of Delaware

Rachel Pulverman
University of Michigan

Mandy J. Maguire
University of Texas at Dallas

Kathy Hirsh-Pasek
Temple University

Shannon M. Pruden
Temple University

Kira Gor
University of Maryland

Seana Coulson
University of California, San Diego

Shimon Edelman
Cornell University

Multiple empirical approaches to a complex analysis of discourse

Linda R. Waugh, Bonnie Fonseca-Greber, Caroline Vickers
and Betil Eröz

1. Introduction

Our purpose here is to exemplify the benefits of an integrated approach to the study of discourse that examines culturally contextualized examples of authentic language use with a rich, fine grained analysis derived from various empirical approaches. By discourse we mean the actual use of language for communication. Communication has many functions (see Jakobson 1960), including referential (e.g., talking about the outside world), emotive (e.g., expressing our emotions, attitudes, beliefs), conative/directive (e.g., getting others to do things, making sure they pay attention to what we say), phatic (e.g., making sure our interlocutor can hear us), metalinguistic (e.g., talking about language), poetic (e.g., aesthetic use), and interpersonal (e.g., creating, maintaining, changing interpersonal relationships – see Halliday 1978) and we need to study language in all of these functions if we're really to have an understanding of language-in-use. Moreover, discourse and its communicative functions are intertwined with linguistic, cognitive, social, cultural, historical, ideological and biological patterns, and none of these exist separately from the others. In other words, ours is an ecological perspective on language, which views language as embedded in a complex ecological system in which all facets are co-dependent. As a result, discourse, here, will be viewed as a process that continually shapes and is shaped by its ecological embeddedness. This means that discourse influences and is influenced by all the other facets of its ecological setting, including all those factors that shape and are shaped by human cognition. This viewpoint is shared by a number of researchers (e.g., Kramsch 2002; Mittelberg et al., this volume; Lantolf 2000; van Lier 2000, 2004; for a more explicit focus on discourse and cognition, see Edwards 1997 and Virtanen 2004; for an integration of cognitive and functional approaches to language structure, see Tomasello 1998, 2003).

Our focus here will be on discourse and what it tells us about all of these factors since we feel that this is an important empirical starting point. Here we will be studying the discourse of members of specific speech communities (also called discourse communities), a speech community being defined as people who “use, value or interpret language” in shared ways (Saville-Troike 1982/1989/2003: 15). We will focus on the spoken language, since that is universal to all human beings, and will deal with only some types (genres)

of speech events. A speech event (also sometimes called a communication event – see Jakobson 1960; Hymes 1974) is a specific instance of language use (including a linguistic message, an utterance, a piece of discourse), which takes place between particular people (speaker and addressee=interlocutors), at a specific time and place, in a particular social and cultural context, using (a) particular language(s), with a specific means of communication (e.g., ordinary speech, telephone, TV, amplified sound, etc.). The speech events analyzed here are comprised of two or more participants in face-to-face dialogue with each other and engaged in ordinary conversation between family members and friends, discussion among students doing a project together, discussion in a classroom, etc. We decided on these speech event types because we wanted to focus on language use that you, the readers of this book, are familiar with and could therefore relate to. All of the participants in our speech events share a common language (French in the case of Waugh and Fonseca-Greber=W&F-G, English in the case of Vickers and of Eröz, though for some participants it is not their native language); and in many cases, they know each other more or less well and have a history of linguistic and social interaction with each other. In some cases, through repetition and shared goals, the interaction between the participants has evolved into a community of practice, “an aggregate of people who come together around some enterprise. United by this common enterprise, people come to develop shared ways of doing things, ways of talking, beliefs, values – in short, practices – as a function of their joint engagement in activity” (Eckert 2000: 35) – all of which practices are related to cognition and conceptual structures.

Most importantly, we are focusing on authentic use of language – language that is socio-culturally and cognitively defined as a form of talk (Goffman 1981) that is typical of the given speech community/community of practice (spontaneously produced and not based on artificial situations) – since we believe that this is the only way to get at what language is really like. In some cases of authentic language, the analyst is present in the speech event and may even participate in it – as is the case for some of the speech events we report on here – but in all cases the analyst is careful not to bias what the participants say and how they say it (or at least, we do our best not to bias it – you can never be sure about this, of course, but we all did our best not to be an intrusive presence). Typically, discourse analysts from many different empirical traditions collect examples of such authentic language use through audio- or video-taping (or both). These traditions include, for example, (anthropological and socio)linguistic approaches to discourse analysis (Labov & Waletzky 1967; Halliday & Hasan 1976; Givón 1979; Hopper 1982; Klein-Andreu 1983; Tannen 1984, 1989, 1993; Chafe & Nichols 1986; Hill & Irvine 1992; Chafe 1994; Bybee & Fleischman 1995; Fox 1996; Cheshire & Trudgill 1998; Du Bois 2003), conversation analysis (CA) (Goodwin & Heritage 1990; Duranti & Goodwin 1992; Schiffrin 1994: 232–281; Hutchby & Wooffitt 1998; Markee 2000), interactional sociolinguistics (Goffman 1981; Gumperz 1982, 2001; Brown & Levinson 1987; Tannen 1984, 1989; Schiffrin 1987, 1994: 97–136, 2001), and corpus linguistics (Stubbs 1996; Kennedy 1998; Biber, Conrad & Reppen 1998; Tognini-Bonelli 2001; Conrad 2002). By taping, the details of language behavior will not be lost (Goodwin & Heritage 1990: 289; Biber, Conrad & Reppen 1998; and Edwards & Lampert 1993) and can be listened to time and again. This body of data has been called a corpus by some of the theoretical approaches we use here (but not by all – indeed, some

ly of
with
e we
ions
stive
rs to
r in-
etic
ps –
ally
stive
ogith-
ther
ded
rse,
gical
ther
hu-
002;
ucus
1 of
nce
dis-
es),
' in
age,
res)

would be surprised at our use of the word corpus, but we feel that it captures nicely the fact that we're using a body of authentic data gathered systematically under rigorous conditions). Since there is a chapter in this volume on corpus-based research, we will not address further the issue of corpus-design, except to say that, while most emphasis in corpus linguistics has been on large corpora (sometimes consisting of millions of words), here we will focus on smaller corpora within well-defined speech communities in well-defined speech events. Our corpora are large enough for quantitative analysis, and small enough to allow us to concentrate on the fine details of linguistic interaction since it's in the details of everyday talk that language works – and we can then relate these details to both the specifics of the speech event and the larger community of which they are a part. In other words, we prefer these smaller corpora that we know well and were gathered for our own work specifically, to the sometimes quite anonymous corpora of the large data banks.

In order to carry out our fine-grained work, we all worked with transcripts (this is standard in corpus work). Now, transcripts are not completely objective, since they are inherently selective and interpretive, and, since there are several different transcription conventions associated with different types of research (e.g., CA, Ethnography of Communication, various researchers doing corpus linguistics – see the contributions to Edwards & Lampert (Eds.), 1993), we each had to decide on the transcription conventions that would be suitable to the analysis being carried out, including the theoretical perspectives and the research questions being addressed (see Ochs 1979; and Edwards 1993, 2001). Transcription conventions are definitely not one-size-fits-all.

The most important issue is that the transcript has to reflect what is actually said, not some cleaned-up version of what is said. Real speech is filled with unfinished sentences, partially pronounced words, incoherent utterances, people overlapping with each other, and so forth – and these are an important part of language use and have to be dealt with in some way (which is why transcripts are different from the standard written language, although they are done in writing). However, it is not enough just to have a transcript. Linguistic communication takes place in a context, and thus, in order to understand discourse, we need also to take into account the socio-cultural systems that have an effect on/are effected by language use, the historical contingencies that influence/are influenced by language use, the ideological structures that permeate the world in which language is used, etc. This means, for example, that the researcher needs to collect social and cultural/ethnographic information in order to better interpret what the participants are saying and what it means to them (as practiced, for example, in the ethnography of communication, see Coulthard 1977/1985; Saville-Troike 1982/1989/2003; Fasold 1990: 39–64, Schiffrin 1994: 137–189), and in sociolinguistics, see Labov 1972, 2001; Coupland & Jaworski 1997; Hudson 1980/1996; Fasold 1984, 1990). For the purposes of our work, all of us gathered information about the participants themselves (e.g., age, sex, education, social status, occupation, native language), the relationship of the participants to each other and their roles in the speech event (whether determined in advance or negotiated during the interaction). In addition, we identified the type of speech event involved: e.g., family meals in France and Switzerland, students meeting outside of class in America to discuss a project they are jointly working on, students in an English composition classroom in America, etc. Each of these types is defined by the particular society of which they are a

part (France, Switzerland, U.S.), as well as the type of context (familial, academic, institutional); hence, it was crucial for us to have knowledge of their socio-cultural status as well as norms and expectations about language use in each type for a complex and rich analysis of the data. Moreover, each has definite cognitive demands that need to be understood in their ecological complexity.

Now, this type of information about speech events is not in conformity with the approach taken, in particular, by CA, which “employs inductive methods to search for recurring patterns across many cases” (Lazarton 2002: 37) and in its more orthodox version claims, for example, that the analyst should know nothing about the participants, except for what can be discerned from the actual talk. However, we agree with Moerman (1998), who advocates combining CA and ethnography of communication into culturally contextualized conversation analysis (CCCA) and proposes close analysis of conversational data in combination with information about the speech community within which the conversation occurs in order to gain an understanding of the culturally specific meaning of the data for the participants. Sometimes, participants in a speech event will not disclose what’s taken for granted, because it’s so obvious to them on socio-cultural, conceptual, cognitive, historical grounds – and it’s important for the analyst to have some sense of the taken-for-granted. At the same time, the analyst, Moerman says, “tries to limit the ingredients of interpretation, the components of meaning, to ones that are locally significant and locally occasioned” (Moerman 1988:7). These abductive interpretations (Peirce 1878/1992) by the analyst are akin to what has sometimes been called pragmatic/discourse/conversational inferencing, but the analyst should take care not to over-interpret the data or bring his/her preconceived analysis to the data. This is very hard to do, but is crucial for rigorous, empirical work with authentic language use since on the one hand use is very rich and complex and on the other hand we want our analyses to come out of the data, not to be imposed on them.

Because of this richness and complexity, discourse analysis is a multi-/inter-disciplinary field made up of traditions from many different disciplines that often don’t overlap, for example, linguistics, cognitive psychology, anthropology, and sociology, to name only a few (for overviews of discourse analysis, see Coulthard 1977/1985; van Dijk 1997a, b; Schiffrin 1994; Georgakopoulou & Goutsos 1997; Jaworski & Coupland 1999; Johnstone 2002; Schiffrin, Tannen, & Hamilton 2001; McGroarty 2002; Phillips & Hardy 2002). Here, we will focus on those empirical approaches that, in combination and incorporating both qualitative and quantitative methods (see Lazarton 2002), have proven to be best for our own work – best in terms of what we’ve studied, what questions we’ve asked, and also where the data led us (see Wood & Kroger 2000 for a different combination than ours).

2. Conversation and pronouns: Linguistic approaches to discourse

We’ll start with Fonseca-Greber and Waugh’s work on spoken French, which we began separately because we were interested in the properties of spoken French and knew that it is very different from written French and indeed from what is taught in French classes here in the U.S., and then decided to do joint work given our common interests. We de-

cided to focus on the subject pronouns because we knew from our own experience that there were interesting differences between spoken and written French in this regard. Indeed, previous published work by others had already suggested that the subject pronouns of French that are presented in reference and teaching grammars and used in the written language do not reflect the everyday speech of educated, middle-class native speakers (for an overview of the French situation, see Fonseca-Greber 2000; Fonseca-Greber & Waugh 2003a, 2003b; for accessible work on English pronouns see Mühlhäusler & Harré 1990; Wales 1996). Native speakers of French, like speakers of any language, are unaware of how much the spoken language is different from the written language and are not conscious of their own linguistic usage, especially in the informal spoken language. Indeed, as is well known, much of language structure and language use is unconscious or at the very least subconscious, so it was clear to us that asking native speakers what pronouns they use, when they use them, what they mean, etc. would not work. And there was no really rigorous study of this situation based on a corpus of authentic use in ordinary conversation. We therefore set out to address this issue, with the aid of corpora of authentic language use and the general findings of corpus linguistics (although most corpus linguistics is focused on English – there is little work, even now, on French). We also were inspired by anthropological and (socio)linguistic discourse analysis, which taught us that work on any part of language – such as pronouns – needs to take into account the larger discourse of which they are a part and the many communicative functions they may fulfill. Pronouns are deictic (shifters, also called indexicality) – e.g., in their meanings, they make reference to their use in speech event contexts (e.g., *I* means the person who is uttering *I*) – and so we were also influenced by work on deixis and indexicality in linguistics. We read as well research in the fairly new field of pragmatics, especially linguistic/ functional/ discourse/socio-pragmatics – and, in particular, those approaches that study language use in relation to the discourse/textual/ social/speech event context in which it is used (see Lyons 1977; Levinson 1983; Anderson & Keenan 1985; Givón 1989; Grundy 1995; Yule 1996; Verschueren 1999). Because we suspected that we would find changes in progress for the pronouns (the written language reflecting the older stages of the language, the spoken language the newer stages), we therefore took into account the two fields that look seriously at language change: 1) historical linguistics, and especially the area of grammaticalization, that studies how grammatical categories (like pronouns) evolve in meaning and use (see Heine, Claudi, & Hunnemeyer 1991; Hopper & Traugott 1993; Bybee, Perkins, & Pagliuca 1994; Bybee 2003); and 2) sociolinguistics and sociology of language, that focuses on different varieties of language use in a given speech community and on language use, ideology and social categorization, was also relevant. This was also coupled with work on ideology and identity that addresses the social and the cognitive/psychological forces that influence both (see Downes 1984/1998; Fasold 1984, 1990; Chambers 1995; Holmes 1992/2001; Hudson 1980/1996; Labov 1972, 2001; Coupland, Sarangi, & Candlin 2001).

In order to study pronouns (and other grammatical issues), we decided to use spontaneous, naturally-occurring face-to-face everyday conversations, audio-taped under naturalistic conditions and transcribed by the authors and/or others under their guidance. There are two corpora represented here: Waugh's French corpus ($\pm 77,000$ words) of Everyday Conversational Metropolitan French (ECMF) and Fonseca-Greber's Swiss (French)

corpus ($\pm 117,000$ words) of Everyday Conversational Swiss French (ECSF), both of which were created for the purposes of this kind of investigation. In some cases, findings from either one of these corpora are reported here; in other cases, the two corpora were combined into a single larger corpus of Everyday Conversational European French (ECEP), containing $\pm 194,000$ words, spanning 15 conversations and including the speech of 27 educated middle-class speakers: 11 men, 16 women, 10 over 40 years old and 17 under 40. This larger corpus allowed more robust (significant) quantitative results. Since care must be taken before using only one corpus or pooling corpora if the results are claimed to be valid more generally, we noted that there are many French (socio)linguists who consider the French of France, Belgium, and Switzerland to have no significant differences in grammar. Moreover, we did an empirical examination of our two corpora for the features we were interested in analyzing and found that there was little/no difference between them – and in cases where there was a difference, we did not pool the results.

The inventory of pronouns for the spoken language with their meanings was established using traditional, discourse, and corpus linguistic techniques of searching for the forms and establishing their meanings through a detailed examination of their use in context. Table 1 gives the written language pronouns that are found in any reference grammar/language textbook; Table 2 gives the spoken language pronouns as we established them through our work.

Now, there are many differences between the two tables and we had two major decisions to make: 1) what to focus on for our research, and 2) what to report on in this chapter, which, given its brevity, could only capture part of our research. We decided to start with a particularly striking quantitative empirical result that had sparked our (and others') interest in subject pronouns at the very beginning: namely the status of the sub-

Table 1. Subject personal pronouns of written French

| Form of the verb | Pronouns (Singular) | Pronouns (Plural) |
|------------------|---------------------------------------|--|
| 1Sg. | je, j' 'I' | 1Pl. nous 'we' |
| 2Sg. | tu* (familiar) 'you' | 2Pl. vous (familiar Pl., formal Sg. and Pl.) 'you' |
| 3Sg. | il (masc.), elle (fem.) 'he, she, it' | 3Pl. ils (masc.), elles (fem.) 'they' |
| 3Sg. | on 'one' (indefinite) | |

* The forms in bold are discussed in this chapter.

Table 2. Subject personal pronouns/prefixes of spoken French

| Form of the verb | Pronouns (Singular) | Pronouns (Plural) |
|------------------|---------------------------------------|--|
| 1Sg. | je, j' 'I' | 1Pl. [nous 'we' rarely used] |
| 2Sg. | tu, t'* (familiar Sg., Pl) 'you' | 2Pl. vous (formal Sg., Pl.) 'you' |
| 2Sg. | tu, t' indefinite 'you' | |
| 3Sg. | il (masc.), elle (fem.) 'he, she, it' | 3Pl. ils (masc., fem.), elles (fem. Pl) 'they' |
| 3Sg. | on- 'we' | |
| 3Sg. | on- 'one' (indefinite) | 3Pl. ils- (masc.) [i, iz] indefinite 'they' |

* The forms in bold are different in the spoken language and the written language as discussed in this chapter.

Table 3. Loss of *nous* and replacement by *on* in ECEF

| n. = 1348 | Tokens | Percentage |
|-----------|--------|------------|
| Nous 'we' | 13 | 1% |
| On 'we' | 1335 | 99% |

ject pronoun *nous*, which is 'we' in the written language, but the spoken results show that the meaning 'we' is now given by *on*, not *nous*, which has all but disappeared from the spoken language as a subject pronoun (Fonseca-Greber 2000; Fonseca-Greber & Waugh 2003a, 2003b).

Example (1) shows the use of *on* to mean 'we'.

- (1) *on-s-est mariés deux fois ouais...ici et aux Etats-U – ouais* (S)
 'we.got.married twice yeah...here and in the United St – yeah'
 [note: according to our empirical results, all the subject pronouns studied here are actually prefixes and will be transcribed with a hyphen to show this status and glossed as part of the word in the English translation. This issue was explored in some depth in Fonseca-Greber, 2000, and reported on in less detail in Fonseca-Greber and Waugh, 2003.
 'United St –' indicates that in the French original, *Etats-U –* is an incomplete word]

99% replacement of *nous* by *on* is an overwhelming, highly significant empirical result – and a surprising one, since when native speakers talk about the use of *on* for 'we', they claim that it is a minor phenomenon and attribute it to uneducated, lower class speech (but, note that our corpus contains the speech of educated, middle-class speakers). Moreover, the extent of the use of *on* has also gone unnoticed by many linguists: this is no doubt because the use of *nous* for 'we' is fully acceptable and grammatical in written French, and *nous* is still used in conjunction with *on* for emphasis or contrast. Thus, when asked for grammaticality or acceptability judgments of fabricated utterances by linguists who rely on such traditional, non-empirical approaches, literate native speakers, who are aware of the conventions of written French as well as the prescriptivist norms for the spoken language, accept these forms. For these native speakers, 'acceptable' or 'grammatical' could be interpreted as 'good' (what they learned at school), not as what they actually say. Additionally, when such informants respond with *nous* for 'we', perhaps they are giving us not the *nous* of written French *nous parlons* 'we speak' but rather the former independent – but current optional, emphatic or contrastive – subject pronoun *nous* of *nous on-a* 'we we-have', since if *on-* functions as a prefix, it may not be cognitively accessible to speakers independently of the verb stem *-a*. Whichever of these explanations may more accurately reflect the cognitive processes speakers engage in when responding to a grammaticality judgment task, a corpus of actual language use allows researchers to uncover previously undetected, even major patterns of language use in a way that intuitive judgments about the language cannot. But we, as researchers must be sensitive to our own interpretive assumptions.

This brings up the issue of the small number of tokens of *nous* listed above (13 tokens, or 1%): in fact, all of these tokens occur in the French corpus (ECMF), not the Swiss one. Through an analysis of the data and correlation of the data with what we know about the participants in our corpus (gained by a close reading of the transcripts of all the conversations, as well as participant data-forms), we found that the majority of the 13 tokens of

Table 4. Relative Frequency of Use of *on-* for 'we' and 'one' in ECEF

| n. = 1749 | Tokens | Percentage |
|-----------------------|--------|------------|
| On = Indefinite 'one' | 100 | 5.7% |
| On = Vague* | 314 | 18% |
| On = Personal 'we' | 1335 | 76.3% |

* Vague=could be interpreted as either Indefinite or Personal in the context.

nous are used by educators (cf. the French Academy, with its heavy emphasis on 'correct' usage in schooling) and thus tend to be related to linguistic conservatism. However, this is true for France only: some of the Swiss speakers are also (retired) schoolteachers but they don't use *nous* (there is no French Academy in Switzerland, and the Swiss in general may be less worried about 'correctness'). In addition, some of the participants in ECMF are of North African background and their *nous* usage may reflect their desire to show their knowledge of formal/written/high prestige norms and to assimilate in French culture. So, even if they are native or all-but-native speakers of French, they may make less use of the range of registers that are available and in particular may cling to the more formal uses of the language.

These findings about *on* bring up a number of questions, since in the written language *on* is the way of giving indefinite meaning and the widespread intuition is that the situation is the same in spoken French. Indefinite pronouns are typically used when the speaker does not want to be or cannot be definite about the reference of the pronoun. In English, for example, some indefinite pronouns are 'one' in the formal language ('one shouldn't do that'), and, in informal usage, 'you' ('to go from Tucson to Phoenix, you take I-10'), 'they' ('they say it's going to rain tomorrow'), and 'we' ('we're going to make a mess of the planet if we don't start conserving energy now'). This led us to investigate a new set of research questions that we had not started out to look at but became vital as soon as we saw the quantitative results: how often is *on* used as an indefinite?; are other personal pronouns used as indefinites?; if so, what is the difference between them conceptually, pragmatically, discursively?; what would lead the speaker to use the one rather than the other?

Our first task was to establish a base-line, that is, how often *on* is used for indefinite meaning, and the answer, based on ECEF, was that it is used in only a small number of cases for indefinite meaning, as displayed in Table 4.

Now, work in linguistic semantics shows that most grammatical and lexical items in a language have multiple meanings (polysemy) as they are used in different contexts and situations. But, in grammar, especially, there is typically one meaning that is the basic (core, nuclear, prototypical) meaning (see Jakobson 1936/1982; Comrie 1976: 11; Taylor 1989/1995; Ungerer & Schmid 1996; Lee 2001), which is assumed to be contextually the least conditioned and cognitively the most salient, and is quantitatively calculated to be present in about 2/3rd (or more) of all tokens. According to our empirical findings, Table 4 shows that the basic meaning of *on* is not the indefinite (5.7%) but its new personal meaning of 'we' (76.3%), which is a complete reversal from its earlier meaning and from its meaning in the written language. Example (2) shows one type of indefinite meaning of *on*:

- (2) parce moi on-m-a fait légaliser le mariage (S)
 'cuz for.me they.had.me authenticate the wedding'

[note that the object pronoun *me* (here *m'* because it precedes a vowel) is, like the subject pronoun, a prefix. All object pronouns are transcribed as prefixes in French and glossed as part of the word in English.]

In (2), *on* means some undefined authorities, unknown to 'me'. Note that the presence of 'me' *m'* as the direct object of the verb means that it is impossible to interpret this use of *on* as 'we'. This is typical of non-basic (peripheral, marginal) meanings: they need contextual support for their interpretation and this gives empirical justification for the analysis. In example (1) above, the interpretation of *on* as meaning 'we' is obvious from the context, since the speaker is talking about herself and her husband and their two weddings to accommodate family in the U.S. and in Switzerland. Table 4 contains a third category – vague *on* – which accounts for almost a fifth of the tokens; example (3) is one case from our data.

- (3) C'est ridicule on-a même pas besoin de permis (K)
 'it's ridiculous one.doesn't/we.don't need to have a permit
 [to buy a gun in the United States]'

The general context here is that K, who lives in the U.S., is talking about violence in the U.S. This utterance could be interpreted to mean either indefinite 'one' (anyone, everyone=any American, all Americans), or as personal 'we' (we Americans) in the context. The discovery of this vague meaning is a contribution to work on both language change and linguistic interaction. On the one hand, it shows how the gradual semantic drift from indefinite to personal, seemingly so implausible at first glance, could have taken place; on the other hand, vagueness about meaning makes *on* very versatile, since it allows the interlocutors not to be specific about what they're referring to and thus offers the broadest range of possible interpretations (see Channell 1994 on vague language). Indeed, speaker and hearer could have different cognitive understandings of the intended meaning, but even so, it is obvious from a close analysis of the data that the flow of conversation is not affected by this vagueness at all.

The empirical results of Tables 3 and 4 raise many new questions, and we decided to continue our quest about indefinite meaning, since, if *on* is only marginally used now in its original indefinite meaning, it seems as if it probably is no longer the preferred way of expressing indefinite meaning, but this has to be proven empirically. The results in Table 5 show that *on* accounts for barely over a quarter of the indefinite (including vague) tokens (or only 8.5% if we leave out the vague tokens). So, how do you give indefinite meaning in French?

Table 5. Indefinite Meaning: 'one', 'you', 'they' in ECMF

| n. = 1489 | Tokens | Percentage |
|---------------------------|--------|------------|
| On (Indef. 'one' & Vague) | 414 | 27.8% |
| Tu (Indef. 'you') | 918 | 61.7% |
| Vous (Indef. 'you') | 18 | 1.2% |
| Ils (Indef. 'they') | 139 | 9.3% |

By reading the literature in French linguistics, we knew that there were others who had already said that the French equivalents of 'you' (*tu* and *vous*) were used as indefinites, so we decided to focus on these two pronouns and their uses in our corpus. What we found was that the vastly preferred indefinite is *tu* 'you [familiar/solidary]', accounting for over 60% of the indefinite tokens (or, 78%, if we take out the vague tokens of *on*); *vous* 'you [formal/plural]' is rarely used in this meaning in our corpus, and so we didn't focus on it in this study. One type of indefinite usage is exemplified by (4), where *tu* means 'one', 'anyone':

- (4) Môme dans les villes comme Cincinnati ou j-sais pas des villes qui sont un peu
 'Even in cities like Cincinnati or L.dunno cities that are a little
 plus perdues que Washington tu-trouveras beaucoup plus
 more isolated than Washington you'll.find [one.finds] many more
 de gens obèses. (K)
 obese people'

It seems to us that this preference for indefinite *tu* could be related to the rise documented by (corpus and socio-) linguists in the use of *tu* (vs. *vous*) as the personal form of address among the young across both sexes, all social classes, educational backgrounds, etc. (see the classic text by Brown & Gilman 1960 on this issue – although we should also say that a study based on corpora with many more uses of the *vous* pronoun than we have in ours might prove interesting). We also think that this rise in the use of *tu* could be related to the spread in France and in Europe more generally of an ideology of egalitarianism, less distance, more solidarity, and less formality between interlocutors. This spread has been documented as occurring since the change to democratic governments in Europe (over a few centuries), but may have received new impetus from the general social changes that have taken place since the 1960's. All of these speculations cannot be addressed by our corpus and await further empirical work, but they promise to be fruitful ground for the strong ecological tie between conceptual/ cognitive and social phenomena.

In addition, the spread of *tu* for indefinite meaning is also not just a French phenomenon, since empirical work on other European languages in corpus and socio-linguistics (especially work on the spread of linguistic features across languages in contact with each other – see Thomason 2001) shows that the equivalents of 'you' in these languages are also increasingly used as an indefinite. Now, since English, the globally dominant language with high prestige, also uses 'you' as an indefinite, there is a possible ideological influence on usage in a variety of languages in contact with English and with each other. Here, there are many avenues of research that should be explored further.

However, we decided to continue work with our corpora and, in particular, we wanted to look further at a third pronoun used for indefinite meaning, namely, *ils* 'they', with 9.3% of the indefinite tokens (or 11.8%, if the vague tokens of *on-* are excluded). Here too we noted a parallelism with indefinite 'they' in English (although we don't have the same empirical evidence for direct ideological influence from English as we do for 'you'). In Table 6, we summarize our empirical findings for *ils* in both its personal and indefinite meanings. Example (5) is one type of indefinite *ils*:

subject
sed assence
t this
need
r the
n the
dings
ory –
fromn the
very-
text.
ange
from
e; on
e in-
adest
aker
, but
s noted to
w in
ay of
ble 5
kens
ning

ntage

6

6

Table 6. Frequency of Personal vs. Indefinite Uses of *ils*

| n. = 341 | Tokens | Percentage |
|--------------------------------|--------|------------|
| <i>ils</i> = Personal 'they' | 198 | 58.1% |
| <i>ils</i> = Indefinite 'they' | 139 | 40.8% |

- (5) *Ils-te-donnent une bourse de transport de l'équivalent de 30 dinards.* (K)
 'They.give.you a transportation scholarship the equivalent of 30 dinards.'

We see, then, that there are three different indefinites in spoken French: most frequent/newer *tu*, less frequent/older *on*, and less frequent/newer *ils*. At this juncture, we made the decision to explore these three further in terms of their meanings/usage and so we did a fine-grained semantic and pragmatic analysis of their uses, and came to the following conclusions:

(a) *Indefinite on*: more formal than *tu* or *ils*; general and wide-ranging in its use; it is the unmarked of the three (Waugh 1982); it can express solidarity or distance or be neutral (see Brown & Gilman 1960); it can be positive or negative or neutral; potentially lighter, face-saving because of its possible neutrality (where face is defined as "a socially attributed aspect of self that is temporarily" construed in verbal interaction with others" (Watts 2003:125), as first explored in depth in the classic text by Brown and Levinson (1987)). Because *on-* is used in its basic meaning for 'we', it has a tendency to include 'we' in some (but not all) contexts. In example (2) above, indefinite *on-* is exclusive, distancing, possibly negative, and as noted above, does not include any nuance of 'we' due to the presence of the object prefix *m-* 'me' in the same sentence, which rules out 'we' as a possible meaning; in example (3) above, *on* could be either inclusive or exclusive.

(b) *Indefinite tu*: less formal; inclusive=potentially inclusive of the interlocutors as virtual/potential participants in the verbal process; highly personalized because it can include and express solidarity between speaker and addressee; strives for communicative involvement of the addressee in the topic of the utterance; usually shows positive face and often gives a positive evaluation of the verbal process. In example (4) above, *tu* functions as an inclusive indefinite: even though only K has been to America, by using indefinite *tu* he includes his interlocutors as virtual participants in the experience, in an act of communicative solidarity. He thereby possibly achieves their greater communicative (including cognitive and emotional) involvement in what he is talking about. The same is true of the use of *tu* as the object in (5).

(c) *Indefinite ils*: less formal; exclusive=potentially exclusive of the interlocutors as virtual/potential participants in the verbal process; expresses distance between the speaker and addressee on the one hand and the topic of the utterance on the other hand; often gives a negative evaluation of the verbal process. In example (5) above, K is referring to unknown authorities that don't include speaker and addressee. Here the distancing and possible negative evaluation is counteracted by the presence of the object pronoun *te* 'you',

and the topic (being given money to study Tunisian Arabic in Tunisia by the Tunisian authorities).

In example (6),

- (6) à Genève le résiné ils-le-font: (S)
in Geneva fruit.juice.pie they.make.it'

indefinite *ils* metonymically means 'those people who live in Geneva', from which, given the distancing of *ils*, the usual pragmatic inference would be that the speaker doesn't live in Geneva. However, given the ethnographic information about the participants and participant-observation, we know that the speaker does live in Geneva. Now, as we saw in example (4) earlier, establishing solidarity in conversational interaction is an effective rapport-building strategy. We therefore interpret this usage in this way: by framing the Genevans as the 'other' ('those people in Geneva'), her interlocutors who are from the adjoining canton of Vaud can interpret her utterance as aligning herself with them and not against them.

In the following example, the speaker uses the personal *tu* in its new plural meaning several times (and the discourse marker *t'sais* 'y'know' – discourse markers are expressions like English *well, but, oh, y'know* that have discourse, social, cognitive and expressive functions, but don't actually refer to anything concrete in the external world – see Schiffrin 1987, 2001). She uses as well the indefinite *ils* (and its object counterpart *leur* 'them'). In this interaction with several interlocutors including an American who is married to a Swiss, the speaker shows solidarity/commonality (positive face) by the use of *tu* with everyone, including the American, and the indefinite *ils* when talking about Americans in general: in this way, the American is being constructed as Swiss, not one of those finicky Americans who don't like trying new foods:

- (7) t-as déjà remarqué qu'ils-sont pas très courageux
'Have y'[plural]noticed that they [Americans] aren't very about brave
au point de vue bouffe non plus? ...t'sais quand t-essaies...quand
about food either?... y'know when ... when...y'try...when
tu-leur-prépare quelque chose d'europpéen. (S)
y'prepare.for.them something European'

In contrast with the virtual inclusiveness that indefinite *tu* creates, indefinite *ils* tends to exclude speaker and hearer and can be used to create distance from those (indefinite people) being talked about.

At this juncture, there were many empirical paths we could have taken – for example, there is much more to be done about the differences between *on*, *tu* and *ils* as used by different speakers, in different context, about different topics, etc. Indeed, there is a whole literature on indefiniteness in English that could be used as a comparison to French. However, we decided to focus next on a research question that arose by working closely with the data and that was due to the special nature of some of the participants in the French data – namely, the fact that some of them were North African in ethnicity and thus they presented very interesting issues for language use, identity, ideology, etc. In particular, we found that the *on*, *tu*, and *ils* indefinites can be used in conjunction with each other to create or shift identities (for the question of identity, see work in interactional sociolinguis-

tics cited above and in intercultural communication (Scollon & Scollon 1995/2001)) or to align oneself with or shift ideologies throughout a conversation (ideology is explored in intercultural communication, linguistic anthropology (Duranti 1997), sociolinguistics, and critical discourse analysis (CDA) (Fairclough 1995; van Dijk 1998, 2001a, b; Fairclough & Wodak 1997; Wodak & Meyer 2001). We decided to concentrate on one especially interesting conversation tape-recorded in France in 1998 between Karim and three others (Karim is a pseudonym – it is standard to use pseudonyms in this type of work in order to protect participants' confidentiality) that led us to further exploration of this issue. We should say that when we began our research we had no intention of looking at indefinites at all – that arose because of our findings about *on* – and we certainly had not envisaged that we would make a foray into the question of identity/ideology – but, as is often the case with empirical work, the data were so compelling that we felt drawn to consider these issues.

As revealed in the ethnographic data as well as in this conversation and others in which Karim participates, he regularly constructs himself as having three ethnic/national identities. He is a French citizen; his grandparents live in France, and his parents used to live in France and still have close friends there whom they visit on a regular basis. At the time of the recordings, he's visiting friends and relatives in France during the summer (sometimes with, sometimes without his mother) – in the case of this conversation, he is with friends, and without his mother. At one point in this conversation, he says "I'm from a French family". But he also makes it clear that he also considers himself to be an American; his parents have lived in the U.S. for some time, near Washington, D.C., he did some of his secondary education there, and he is an undergraduate in an elite liberal arts college in the Northeastern U.S. (he's from an upper middle class family). He gives his interlocutors lots of information about all of this. And, he is of Tunisian descent, and at one point he says "I'm Tunisian-American". In addition to his three ethnic identities, Karim's linguistic situation is also complex: he is a fluent bilingual speaker of both French and English (the other three in the conversation are native speakers of French and don't speak English), and just before this conversation he had gone to Tunisia for a month to study Tunisian Arabic (which he doesn't speak). As we will see, he shows his shifting identities and his ideology(ies) through his use of the indefinite pronouns.

Early in the conversation, Karim talks about his language training in Tunisia, and part of what he says was given in example (5) above, incorporated in example (8) here:

- (8) Karim: Ils-te-paient la moitié de ton voyage Ils-te-paient les cours.
 'They.pay.for.you half of your trip. They.pay.for.you course fees.
 Ils-te-donnent une bourse de transport de l'équivalent
 They.give.you a a transportation scholarship of about
 de 30 dinards. [...] On- on-fait pas mal avec. Et euh on-te-donne
 30 dinards. [...] One. one.can get along on that. And uh they.give.you
 une chambre dans un foyer universitaire.
 a room in a dormitory.

By the use of *tu* ('you' and the object and possessive adjective equivalents *te* 'you' and *ton* 'your' respectively), Karim indexes himself as Tunisian and at the same time creates solidarity with and involvement of his interlocutors (none of whom is Tunisian). *Ils* 'they' in (8) means the anonymous, faceless Tunisian authorities. He differentiates them from

the university people who gave him a room in the dormitory, since he uses *on* in the last sentence when talking about them and the inference is that he met them (had face-to-face interaction with them). The next to last *on* (in *on- on-fait pas mal avec 'one. one.can get along with that'* – the repetition of *on* is known as a 'false start'), rendered awkwardly as 'one' in English, is vague: it could mean 'we' (the participants in the course, or even more generally Franco-Tunisians living abroad who go there), or it could mean 'anyone visiting Tunisia'. But, as said earlier, it doesn't matter what it means exactly, since the conversation continues without any problems and it may be that Karim himself couldn't say which one he might have meant – perhaps he meant them all.

Later in the conversation, Karim constructs himself as being an expert on America, and when he talks about violence in America and how easy it is to buy a gun, he says (part of this was example (3) above):

- (9) K: C'est c'est ridicule. On-a même pas besoin de permis t'vois c'est juste
 'It's it's ridiculous. One.doesn't even need a permit y'see that's right
 <-> on-achète son arme et on-sort C'est ridicule. Et après ils-sont
 <-> one.buys the gun and one.leaves It's ridiculous. And then they're
 étonnés d'avoir des trucs, des tueries comme ça.
 surprised to have stuff, killings like that one."

Here, the use of vague/neutral *on* indexes the fact that buying a gun in America is an everyday occurrence that anyone can do (and this is in contrast with France, and his interlocutors know that the American situation is completely different from the French one). Karim shows his distance from the American situation – and from America and Americans – by stating twice "it's ridiculous" and by using the indefinite *ils* in his last sentence, referring to (those faceless, anonymous) Americans from whom he distances himself ideologically: 'not me', he implies; 'in this respect, I'm not an American'.

In a final example, Karim is talking about obesity in the U.S.:

- (10) Alors ils-font attention alors en ville. Quand tu-vois les obèses
 'So they.pay more attention [to weight] in the city. When you.see obese people
 on-peut assez facilement réparer s'ils-sont de la ville ou pas...
 one.can pretty easily figure.out if they're from the city [Washington] or not...
 Même dans les villes comme Cincinnati ou j'sais pas des villes qui sont un peu
 Even in cities like Cincinnati or I don't know cities that are a little
 plus perdues que Washington tu-trouveras beaucoup plus de gens obèses.
 more isolated than Washington you'll.find many more obese people.
 Et ça c'est clair.
 And that's for sure.

Here, again, Karim reveals his shifting identity and his ideology through his use of the indefinite pronouns. When he uses *ils* in the first sentence, he means 'those Americans' ('not me'), even though elsewhere in this conversation he very much identifies himself as an American; in this same sentence he makes reference to 'the city' (meaning Washington, D.C.), which he has talked about earlier at length as being 'my city' (because he grew up there, his parents live there, and he goes there often). In the next sentence (and in the next to last sentence), he uses *tu* as a conversational involvement strategy with his interlocutors,

who could imagine themselves there. But then his use of the more neutral *on*, which could be rendered better in English through the impersonal – ‘it’s pretty easy to figure out if they’re from the city or not’ – suggests that anyone could tell whether people are from the city or from somewhere else (e.g., American tourists visiting Washington) by how obese they are.

In these two excerpts about America, we can see that when it comes to some of the stereotypical characteristics of Americans that the French (like his interlocutors) don’t like about Americans (buying guns and being obese), Karim makes it clear that while he lives in America and in certain ways he is an American, he is ideologically not American – he’s not one of them.

We started this section with the issue of what the pronouns of spoken French are and we have seen that by using a variety of different qualitative and quantitative empirical methods inspired by many different traditions, most of them related to linguistics in some way, we have not only been able to make discoveries about forms, meanings and uses of the pronouns but also to address a variety of social and cultural issues important for cognitive analysis, such as establishment of solidarity with or distance from the addressee by the speaker, positive and negative face, indexing of linguistic and cultural identity, and showing one’s ideology – all through a fine-grained analysis of the use of indefinite pronouns in conversational interaction. Now, given our data, there are many more empirical issues we could have raised – and indeed we have addressed some of them in other contexts (see Fonseca-Greber 2000; Fonseca-Greber & Waugh 2003a, b), but there is much more that needs to be done. A corpus of authentic data is like a treasure-trove and working with one can take years. However, if we wanted to address questions having to do with the use of specific lexical items (one of the traditional ways in which larger corpora are used), we would have to gather more data. We should also note that Waugh’s corpus contains some spoken academic discourse (university lectures) and what are called ‘service encounters’ (e.g., talk between a cheese store owner and customers, between a pharmacist and customers) – and these too can be used for a variety of different empirical questions, which have yet to be explored.

3. Discourse and interactional accommodation: CCCA

The discussion in this section concentrates on my own (Vickers’) English corpus ($\pm 150,000$ words), focusing on conversational interactions between native speakers (NS) and nonnative speakers (NNS) of English in the U.S. academic community (see Vickers 2004). My study demonstrates ways in which NSs and NNSs interactionally accommodate to each other as they try to perform a task together, and, in particular, instances in which interactants successfully come to work within a shared interpretive frame, i.e., ways of knowing and of experiencing the world (Gumperz 1982). Furthermore, the discussion will demonstrate that coming to work within a shared frame has implications in terms of the power relations between NSs and NNSs in this speech community. I will discuss in this section the research process, including the challenges of collecting a corpus for the purpose of studying a specific set of research questions, as well as the final product of my study.

To begin, I will discuss the challenges involved in finding the appropriate context in which to collect data related to a specific set of research questions. It is not always easy to find a context friendly to the everyday presence of the researcher, and it is crucial that the researcher be present on a regular basis when collecting a culturally contextualized corpus of this sort. In the case of my study, I originally wanted to do research in the context of a college of business in a major American university. I had hoped to investigate Chinese and Korean students' interactions with American students during teamwork in this context. Unfortunately, though the college of business was open to my conducting experimental research in which students had contact with me in a laboratory type situation, they were not comfortable with me naturalistically observing their teams. Therefore, I had to find a different context in which to address my research questions. The department of Electrical and Computer Engineering (ECE) as well as Family and Consumer Sciences enthusiastically welcomed my research. The ECE department offered me access to a large undergraduate course, while Family and Consumer Sciences could only offer me access to a small graduate seminar.

After considering these options, I decided to work with the ECE department and thus the corpus I'll be discussing was gathered in the context of a large undergraduate design course, in which typically 70% are NSs and 30% are NNSs. Though there were quite a few NNSs in the course, I was unable to recruit multiple Chinese and Korean participants. As a result, rather than focusing on Chinese and Korean students' interactions with American students, I was forced to revise my research questions to encompass a study that would examine NS-NNS interaction in the ECE speech community, regardless of country of origin. Therefore, my research questions had to be revised to reflect the population that I was actually studying, one in which the NNSs came from a variety of cultural backgrounds. The actual data that I had access to led to the revision of my research questions.

In the ECE design course, all students are required to work together throughout the academic year on teams, formed by the professors according to the students' technical interests and background, to design operable electronic devices. The situation is high stakes, since students' graduation is dependent on the design of a successful device. As such, the ECE team meetings represent a naturalistic context in which to study NS-NNS teamwork. The focus of this study was shaped by the context of the study, the data, and thus the salient issues in this high stakes, competitive, and quite male context.

To collect data, I employed microethnographic methods, as developed in the ethnography of communication. Besides video and audio taping of the team meetings, I engaged in participant-observation: I observed the classes, I was at the team meetings and I was present at the ECE portfolio evaluations. I also conducted participant playback sessions (where I showed the video tapes to the participants and asked them to comment on certain factors of the interaction). The objective was to gain an understanding of the speech community into which the ECE students were being initiated and of the ECE students' participation within that speech community, so that I could understand what was happening in the team meetings, which themselves constituted communities of practice. Observation of the course lectures allowed me a certain amount of insight into the forms of talk (Goffman 1981) in the ECE speech community – these forms of talk are quite specialized, including extensive use of technical jargon. Observation was also informative in

terms of the course objectives and the standards of professionalism in communication that were expected in the ECE team meetings.

One important consideration when engaging in ethnographic work, particularly in an unfamiliar speech community, as in the case of my study in the ECE community, is to allow the salient interaction types and implications of those interaction types in that particular community to emerge from the data. Especially when much is unfamiliar as the research begins, initial research question must be a) quite broad and b) flexible. I had to be willing to revise the research questions based on the data. The data drove the research questions, and revision of research questions was a recursive process that happened in combination with data collection and analysis.

As I collected and began to analyze the data, it became clear that there were particular interaction types that allowed teammates to accommodate to each other so that they could engage in negotiation. It also became clear from the data that NSs and NNSs had different ways of accommodating. Thus, one of my research questions became, "What strategies do NSs and NNSs employ to accommodate to each other?" Another became "What implications do these strategies have for the hierarchical structure of the team?" I formulated these research questions based on what the data was showing me. As I engaged in data collection and analysis in investigation of these questions, it was evident that I would need to involve multiple theoretical approaches to fully analyze the data and address my research questions, as I will demonstrate below.

In conducting data analysis, at the micro level, consisting of fine-grained analysis of participants engaged in face-to-face interaction, I incorporated analytic methods from CA, interactional sociolinguistics, and corpus linguistics. At the macro level – analysis of the larger context within which the participants operate – the work was inspired by ethnography of communication and in particular CCCA (Moerman (1988) discussed above). The collection of multiple instances of the same speech event in the same cultural context within and across teams allowed for the kind of systematization conducive to both quantitative and qualitative corpus analysis.

On the level of micro interaction, the most important influence from CA is the necessity to examine the interdependency between utterances in order to understand how utterances are contextualized within a conversation and how conversants' contributions are constructed within the conversation at hand. By examining utterances in relation to surrounding utterances in the context of an ongoing conversation, I was able to locate conversational sequences, in this case, accommodation sequences. The accommodation sequences explored in this analysis contain in the first utterance either the solicitation (explicit request) for a clarification by a participant in the speech event or an explanation by a participant in the speech event that was not solicited by the other participants. The clarifications and explanations are of four types: an assumption, an opinion, technical content, and linguistic content. I made decisions about coding these sequences in light of the conversational context in which the speaker made the utterance and according to the way the utterance was understood (uptaken) by the interlocutor. In other words, while in some cases the intention of the speaker may be different from the understanding (uptake) of the addressee, the flow of the discourse is determined more by how the addressee understands what the speaker has said and this in turn determines how the addressee responds to what

the speaker has said – and thus, the focus here is on the co-construction of what the discourse is about, since that is what the data show (for a discussion of co-construction, see Jacoby & Ochs 1995). In other words, from a cognitive perspective, what is important is how the interactants understand what was said and how it should be understood, not just what the speaker might have meant or intended.

Examples 11 and 12 are solicited clarifications; they are solicited clarifications rather than unsolicited explanations because in (11) P and in (12) R ask for clarification through their direct questions – italicized in the example. (Other clarification sequences not exemplified here are Solicited Clarification of Opinion=SCO, and Solicited Clarification of Linguistic Content=SIC).

(11) *Solicited Clarification of Assumption Sequence (SCA)*

G: I don't know if it's worth it maybe he'll give me some of the old parts and maybe that'll cut it for what we need and it'll at least give us something to test in the lab

P: *well what does that mean for like part one b does that mean if you can't get the parts until January that we're putting it off until January*

G: yeah that's what that means

(12) *Solicited Clarification of Technical Content Sequence (SCTC)*

G: I mean it wouldn't be that bad if we could run this off five volts I don't know

R: *what what's the matter with five volts*

G: [it's]

P: [it's] not a good battery voltage

G: *yeah you got a 1.5 volt battery so you need four of 'em*

R: yeah that's true

G: *three volts you only need two double A*

[Notice that the brackets around G's [it's] and P's [it's] means that they overlap with each other=talk at the same time]

Example 12, involving two NSs, George (G) and Peter (P), and one NNS, Ramelan (R) (all pseudonyms), shows the necessity of looking at the interdependency between utterances to identify an accommodation sequence. In their response to Ramelan's initiation of a clarification sequence, George and Peter show that they have uptaken it as SCTC (not SCA): Peter provides the first part of the technical content, and George the second part and thus they are both able to display their technical expertise. Ramelan accepts these responses. This demonstrates that interactants do not come to a particular interaction with all aspects of the interpretive frame predefined (as would be assumed in a pure interactional sociolinguistic analysis) but that the interpretive frame itself is co-constructed in face-to-face interaction, as CA claims, and that these sequences constitute strategies that interactants employ to formulate a shared interpretive frame. This is also the basis for some who claim that cognition is social in nature (see Edwards 1997; Lantolf 2000; Virtanen 2004, among others). Examples 13 and 14 are unsolicited explanations:

(13) *Unsolicited Explanation of Assumption Sequence (UEA)*

B: I'm talking about the design you're talking about the design so it's like we are [repeating ourselves]

A: [yeah therefore] we understand ourselves but when you said we should do this I thought with this like I'm specializing hardware I will just this mostly [about the hardware I don't]

B: [no no right so you'll do all]

(14) *Unsolicited Explanation of Technical Content Sequence (UETC)*

P: I thought that big one looked like it did but

G: no

P: no it's [surface mount]

G: [surface mount] so what you do is be able to board with a surface mount lay it [on to the pad]

P: oh yeah right

In example (14), George is able to display his own technical knowledge by providing an unsolicited explanation of technical content to Peter. The display of technical knowledge along with an acceptance of that technical knowledge by team members has the effect of shaping the team's interpretive frame. Thus, the micro-level analysis provides strong evidence that displays of technical knowledge are associated with high status on the ECE team, since being able to shape the interpretive frame gives a participant high status.

This brings up the issue of who has high status on these teams. Table 7 demonstrates the percentage occurrence of each type of accommodation sequence, as well as percentage of types of accommodation sequences used by NSs and NNSs respectively.

By examining the frequency of occurrence of accommodation sequences by NS and NNS, I was able to determine major differences in the ways the NSs and NNSs engage in interactional accommodation. In general, NNSs are far more likely to engage in solicited clarifications in general, and in particular they favor SCTC (42% of all their accommodation sequences) and SCA (31%). Native speakers, on the other hand, are more equally divided between the solicited clarifications and unsolicited explanations, and while they do engage in SCTC (20%) and SCA (16%), it is at a lower rate than NNSs and they favor UETC (41% of all their accommodation sequences), while for NNSs the percentage of UETCs is only 8%. This difference becomes important to the construction of the status of NSs and NNSs on the team.

Ethnographic data from the ECE speech community then informs the micro-level analysis because in the ECE speech community, the ability to display technical knowledge is important in achieving high status in the community. Ethnography of communication provides a framework for examining the ECE team meeting speech event as embedded in the ECE speech community. Therefore, coming to know the speech community sheds light on the forms of talk in the ECE team meeting, and how they operate in the speech community. In particular, it is clear that UETC is a powerful, high status move – because

Table 7. Frequency of All Accommodation Sequences

| Initiator | SCA | SCO | SCTC | SCL | UEA | UEO | UETC |
|-----------|-----|-----|------|-----|-----|-----|------|
| All | 21% | 4% | 27% | 9% | 2% | 6% | 31% |
| NS | 16% | 6% | 20% | 7% | 2% | 8% | 41% |
| NNS | 31% | 1% | 42% | 15% | 1% | 2% | 8% |

it is unsolicited and because it deals with technical knowledge, which is given high prestige both in the ECE community and in the community of practice of team meetings. On the other hand, SCTC is quite face threatening for speakers, because, by soliciting technical knowledge from their interlocutors, speakers construct themselves as lacking technical competence, and, as a result, as potentially having low(er) status in this speech community/community of practice. Recurring instances in which NSs display (solicited or unsolicited) technical knowledge and NNSs solicit technical explanations then results in the construction of NSs as high status members and NNSs as low status members of the team, not because of their linguistic competence *per se* (although this may be involved) but because of their ability to engage in high status forms of talk within the ECE speech community.

Furthermore, ethnographic data demonstrates as well that it is not simply the NS-NNS distinction that necessarily gives NSs higher status than NNSs, but it is a matter of who has access to high status forms of talk in the ECE speech community in general. In particular, playback session interviews in which corpus participants discuss team interactions while watching videotapes of the interactions shed light on other contextual information, such as the fact that NSs and NNSs have differential access to high status forms of talk within the ECE speech community: NS ECE students typically have the opportunity to gain professional internships which typically last several years and allow them to gain access to high status forms of talk; NNSs do not have this opportunity because of their visa status and thus are less exposed to high status forms of talk.

The integrated analysis of the corpus data shows that people from different cultural backgrounds typically accommodate to each other in the process of face-to-face interaction in the ECE team meeting speech event in order to establish a shared interpretive frame and that there are shared strategies for interactional accommodation that allow participants to construct this frame. However, people from different cultural backgrounds and with different access to high status forms of talk (and experience) tend to employ strategies associated with interactional accommodation with different frequencies and in qualitatively distinct ways. These differences become important to the development of social hierarchy on the ECE team and in the ECE speech community because, in the process of accommodation, much of the construction of team members as possessing technical expertise in particular areas and as being able to make technical expertise relevant to the team's project is accomplished. As such, micro-level interactions construct, co-construct, and re-construct NS team members as high status and NNS team members as low status members of the team and ultimately of the speech community.

4. Interactional patterns of international students: Triangulation

This section is based on my own work (Eröz 2003), which focuses on the interactional patterns of international students with each other and with Americans, in order to determine if international students interact in ways that are different from Americans and if these ways of interacting are a reflection of classroom interactional patterns in their home country. The speech community or communities of practice I analyzed for this project

this I
t the

on to

g an
edge
effect
rong
ECE

rates
it age

and
ge in
cited
mo-
ually
they
y fa-
ge of
us of

level
edge
tion
lded
heds
each
ause

JETC

1%
1%
%

consisted of two mandatory freshman composition courses at a large southwestern American university. The participants were 35 international and 9 American students, some enrolled in an all-international section (IS) consisting of non-native speakers of English (NNS), the others in a mixed section (MS) consisting of both American native speakers of English (NSs) and international students.

In this research study, I focused on the common interactional patterns within four non-American culture groups: (Asian) Indian, Japanese, Chinese, and Middle Eastern. I also investigated international students' perspectives on studying with native speakers in content versus English classes, as well as their sociocultural adjustment difficulties as newcomers in the United States. Furthermore, I analyzed classroom interaction patterns common to three participants from each of the culture groups and compared these to American classroom patterns. My three data collection techniques were ethnography (combined with ethnography of communication), CA, and CCCA. The database for this corpus was compiled through participant (classroom teacher) and non-participant observation as well as video and audio taping of classroom discussion (15.5 hours) and of individual (7.5 hours) and group teacher-student (6 hours) conferences. In addition there were one-on-one interviews with most of the international participants and questionnaires using both Likert-scale and open-ended questions about their experiences in America. I also used my own written data: systematic notes taken during classroom observations in the MS, reflective and retrospective Teacher's Journal entries made after IS class sessions, notes taken during multiple viewings of the individual interviews with participants asking for their reactions, and topic transcriptions of the video-taped data noting the time of the interactions, direction of the interactions (e.g., teacher-to-student), and content of the interactions (e.g., teacher asks if there are any questions). My aim was to do an integrated analysis of multiple data sources – triangulation – in order to come up with relevant, objective, and trustworthy results.

The claim that I addressed through this triangulation was that there are interactional similarities among students from the same culture in terms of how and how much they participate in class, and how they relate to each other and to the teacher in whole class and small group discussion in the classroom, in teacher-student conferences and in interviews. My hypothesis was that these similarities are due to the fact that these students transfer attitudes about the role of the teacher and traditional classroom behavior from their home country to the American classroom. I explored these claims by making comparisons across all the various types of data, searching for commonalities among same-culture groups and differences across cultures.

In order to explain the procedures of triangulation for this data set to address the research questions, I will exemplify the nine stages of my data analysis and interpretation: 1) Question, 2) Database Overview, 3) Assumption, 4) Close Look at Database, 5) Data Source(s) Selection, 6) Example Selection, 7) Interpretation, 8) Literature Comparison, and 9) Conclusion. To illustrate these stages I now focus on the Chinese students in the Asian group and the research question pertinent to investigating similarities among students from the same culture and the cultural influences of this behavior.

1. *Question:* Are there interactional similarities among students from the same culture in terms of how and how much they participate in class and how they relate to the

teacher and each other in communicative activities (e.g., small group discussions and conferences)?

2. *Database Overview*: To address this question, I looked through classroom notes, topic transcriptions of lectures and student-teacher conferences to identify culture-specific patterns of the culture groups represented by three or more participants – in this database, these groups were Asian Indian, Chinese, Japanese, Middle-Eastern, and American. After this, I determined one specific group to focus on: Chinese students.
3. *Assumption*: After the initial overview of data, I observed that Chinese students tended to be quiet during classroom discussions, but collaborative in small group discussions. They seemed modest and respectful towards the teacher; they didn't interrupt or disagree openly with the teacher.
4. *Close Look at Database*: At this stage, I consulted written data sources to note the general classroom interaction behavior of Chinese students; reviewed instances of Chinese students talking in classroom discussions and small group work; and examined their attitudes in individual and group conferences. I revisited the class videotapes when necessary.
5. *Data Source(s) Selection*: I checked and validated assumptions through the use of these following data sources: videotaped classroom sessions; topic transcription of taped sessions and conferences; interview information gathered from or about Chinese students; and questionnaire responses of Chinese students in general.
6. *Example selection*: In order to illustrate the points implicated by the assumptions, I selected striking and clear examples from various data sources:

Data Source 1: Videotaped conferences (examples 15 and 16):

- (15) Adam (Chinese NNS in IS): individual conference with the teacher

Teacher (T): You have to give specific examples from the story {writes on his essay}

Adam: A-hm

T: From the story, and then you have to say... you have to explain what this means about the culture

Adam: Okay

T: What you think it means about the culture. And then you need to refer to your sources and say...you know... in India traditionally the relationships between mothers and sons are like this.

Adam: Okay

T: That is coming from your outside sources {Adam nods} supporting the illustrations that you gave {Adam nods} and then the point that you made.

So this is support for both of those things.

Adam: A-hm

T: Okay?

Adam: Yeah

T: But if you're not specific, if you're not direct...then it can be...it may not give the desired effect

Adam: Okay

{Brief silence}

- T: What else? Anything else?
 Adam: {Shakes his head}
- (16) Ted (NNS in IS): group conference with teacher and Dan (NS)
 Teacher: {To Dan} Did you read Ted's essay?
 Dan: Yeah.
 T: {To Ted} Did you read Dan's essay?
 Ted: {Nods}
 T: We'll grill the essays. Who wants to go first?
 Dan: Ted's essay is good, it's great.
 T: Come on that's not fun [laughs].
 Ss: {Smile}
 T: {To Ted} What did you think of Dan's organization and content?
 Ted: OK. ...Very nice.
 [Brief silence]
 T: {To Ted} Did you like his points?
 Ted: {Silent}
 T: Did you find them?
 Ted: {Silent, looking at essay}
 T: Sometimes I couldn't find them.
 [Brief silence]
 T: {To Ted} What do you think?
 Ted: OK. Better than me.

Data Source 2: Questionnaire responses

In her questionnaire, Rosy summarized her perspective on appropriate student conduct in the classroom like this: "Chinese thought: Respect. Teacher always right. If you have questions, ask personally or ask after class. Never interrupt or make teacher embarrassing." In the questionnaires, almost all of the students who responded "strongly agree" or "agree" to the statement, "The cultural norms I grew up with prevent me from being critical of my classmates' work in peer revision", were Chinese.

Data Source 3: Interview Information

In our one-on-one interviews Lee admitted that "[Chinese students] don't really talk, they are quiet. They aren't active," and in his interview Adam explained this behavior by stating that "Chinese students aren't comfortable criticizing other students' essays. It is rude. Criticizing brings a lot of shame. In Hong Kong, I never criticized others so I don't know. Criticism from teacher is OK, that is what happens in Hong Kong. . . In Hong Kong whatever teacher say is right. No questioning the teacher."

7. *Interpretation:* The video recordings present the Chinese students as one of the quietest groups in classroom discussions. However, they seem to be collaborative and even talkative in small group tasks. Adam, for example, doesn't speak up much in class, but he seems really engaged in small group work. He has smart comments and helpful ideas while performing the task, but he is never the group spokesperson. An-

other example is Paige and Jeanie who did a class presentation together. They were soft-spoken and hesitant while addressing the whole class, but they are more active in group work. Looking at the examples from various data sources and the comments of the Chinese students, it is possible to confirm Chinese students' behavior as consistently respectful rather than passive. By their own report, they consider teachers' suggestions and comments to be essential, so they listen without interrupting or asking questions, which they say is a result of their cultural background and the Chinese education system. Their discomfort in criticizing others is confirmed through group conference behavior and interview responses.

8. *Literature Comparison:* All of these results were confirmed by comparison with the literature on Chinese education, which shows that Chinese classrooms are traditionally teacher-oriented; students don't ask questions or work on collaborative tasks like small group or dyadic activities (Su & Su 1994; Liu 2001; Hammond & Gao 2002). The teacher is seen as the authority figure who lectures for the most part and doesn't encourage students to ask questions or express their opinions; students are expected to stay quiet without asking questions or challenging the teacher (Su & Su 1994). Two Asian participants in Liu's (2001) study suggested that asking very simple and basic questions in class was wasting class time. Therefore, they preferred to talk to the teacher and ask questions after class. They were surprised and annoyed to see Americans asking every question they had during class time and thereby taking up the teacher's lecture time.
9. *Conclusion:* Through triangulation of the various data sources as well as confirmation from the literature about Chinese education and other studies of Chinese students in American classrooms, it is possible to interpret Chinese students' quiet behavior and their reluctance to do peer review as related to Chinese cultural and educational norms – which are then transferred to the American setting. Thus, we have seen that triangulation is challenging but worthwhile; it enables researchers to look at speech communities from various angles and helps them to come up with relevant, objective, and trustworthy results.

5. Conclusion: Discourse and cognition

In this chapter on the study of discourse, we have shown the importance of working with a number of different methodological approaches, for all of which it is necessary to have a corpus of tape- and/or video-recorded language use produced within a particular socio-cultural context that can serve as the empirical data for a rich, complex analysis. In addition to the corpora themselves, the researchers collected ethnographic and sociolinguistic information through interviews, questionnaires, and/or observation, in order to become familiar with the cultural context within which the data was collected and to understand the participants themselves, including their conceptualization of the language they used and their style of interaction. They also performed fine-grained qualitative and/or quantitative analyses of discursive and interactional patterns by working with the transcripts of the language use in combination with the ethnographic and soci-

duct
have
ing.”
gree”
al of

talk,
avior
. It is
don't
Kong

qui-
and
h in
and
An-

olinguistic information. In addition, the sources of data were triangulated to obtain rich, multiplicitous analyses. The use of a corpus of authentic language enabled the researchers not only to address their initial questions but also to explore those that emerged from the data themselves during analysis. The discussion of the corpora collected by Waugh, Fonseca-Greber, Vickers, and Eröz demonstrate that the study purpose drives the specific data collection method(s), and that the data drives the specific analytic approach(es) so that the end result is a rich analysis that captures the true complexity of human interaction in specific speech communities and communities of practice.

As said in the introduction, it is our belief that discourse and cognition are intertwined with each other and thus that the study of one of these facets necessarily touches on the study of the other: (individual) cognition arises out of, and is influenced by, language use in its multifarious social and cultural contexts. Further, cognitive representations speakers have of their language may evolve over time through a combination of factors, e.g., language change and socio-cultural change. Discourse, cognition, social and cultural activities – all are mutually implicating, and none is understandable without serious consideration of their interdependent integration. In other words, we agree with Tomasello that language itself is a “complex mosaic of cognitive and social communicative activities closely integrated with the rest of human psychology” (1998:ix). Hence, in order to be complete, the study of cognition needs to be correlated with work on discourse, social and cultural structures, and the functional foundation of language. Thus, work on discourse, especially insofar as it is multidisciplinary and integrates perspectives that take into account not only its linguistic but also its social and cultural aspects, provides a complementary, indeed a necessary, perspective for an integrated understanding of cognition. As discourse analysts, we are interested in the nature of human interaction through language in its social and cultural setting, and it is from that perspective that we can draw conclusions about, and provide insight into, cognition. Thus, as we have seen, the grammatical structures that participants use index social status, identity, attitudes, ideologies; the particular strategies chosen in the process of interaction construct hierarchy among individuals and hence create group dynamics; and the linguistic, cultural and cognitive patterns that have been developed in one culture may be difficult to change in another one and may give rise to specific behaviors that are inappropriate or misunderstood in a different cultural context. The social and cultural settings and the communicative tasks at hand influence not only what is said but also the construction, co-construction, and re-construction of the ways that individuals are perceived by others, how they perceive themselves and how groups of people interact. Thus, as said earlier, the main thinking underlying this chapter is that language use (discourse), language structure, cognition, social and cultural structures, and biological, historical and ideological patterns are all interconnected in a complex ecological system, which therefore necessitates multiple empirical approaches that can give an understanding of their relational and interdependent complexity.

References

- Anderson, S. & Keenan, E. (1985). Deixis. In T. Shopen (Ed.), *Language typology and syntactic description III: Grammatical categories and the lexicon* (pp. 259–308). Cambridge: Cambridge University Press.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Brown, R. & Gilman, A. (1960). The pronouns of power and solidarity. In T. Sebeok (Ed.), *Style in language* (pp. 253–276). Cambridge, MA: MIT Press.
- Brown, P. & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Bybee, J. (2003). Cognitive processes in grammaticalization. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. 145–167). Mahwah, NJ: Erlbaum.
- Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar*. Chicago: Chicago University Press.
- Bybee, J. & Fleischman, S. (Eds.). (1995). *Modality in grammar and discourse*. Amsterdam: John Benjamins.
- Chafe, W. (1994). *Discourse, consciousness, and time*. Chicago: University of Chicago Press.
- Chafe, W. & Nichols, J. (1986). *Evidentiality: The linguistic aspect of epistemology*. Norwood, NJ: Ablex.
- Chambers, J. (1995). *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. Oxford: Blackwell.
- Channell, J. (1994). *Vague language*. Oxford: Oxford University Press.
- Cheshire, J. & Trudgill, P. (1998). *The sociolinguistics reader: Volume 2, gender and discourse*. New York, NY: Arnold.
- Comrie, Bernard (1976). *Aspect*. Cambridge: Cambridge University Press.
- Conrad, S. (2002). Corpus linguistics approaches for discourse analysis. In M. McGroarty (Ed.), *Annual review of applied linguistics, 22, Discourse and dialogue* (pp. 32–51). Cambridge: Cambridge University Press.
- Coulthard, M. (1977/1985). *An introduction to discourse analysis*. London: Longman.
- Coupland, N. & Jaworski, A. (1997). *Sociolinguistics: A reader and coursebook*. New York: St. Martin's Press.
- Coupland, N., Sarangi, S., & Candlin, C. (2001). *Sociolinguistics and social theory*. Essex, England: Pearson.
- Downes, W. (1984/1998). *Language and society*. Cambridge: Cambridge University Press.
- Du Bois, J. W. (2003). Discourse and grammar. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. 47–87). Mahwah, NJ: Erlbaum.
- Duranti, A. (1997). *Linguistic anthropology*. Cambridge: Cambridge University Press.
- Duranti, A. & Goodwin, C. (Eds.). (1992). *Rethinking context: Language as an interactive phenomenon*. Cambridge: Cambridge University Press.
- Eckert, P. (2000). *Linguistic variation as social practice*. Oxford: Blackwell.
- Edwards, D. (1997). *Discourse and cognition*. London: Sage.
- Edwards, J. (1993). Principles and contrasting systems of discourse transcription. In J. Edwards & M. Lampert (Eds.), *Talking data: Transcription and coding in discourse research* (pp. 3–32). Hillsdale, NJ: Erlbaum.
- Edwards, J. (2001). The transcription of discourse. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 321–348). Malden, MA: Blackwell.
- Edwards, J. & Lampert, M. (Eds.). (1993). *Talking data: transcription and coding in discourse research*. Hillsdale, NJ: Erlbaum.
- Eröz, B. (2003). *An Ethnographic examination of international students in English composition classes: Cultural patterns, classroom dynamics, and adjustment difficulties*. Unpublished doctoral dissertation. Tucson, AZ: University of Arizona.
- Fairclough, N. L. (1995). *Critical discourse analysis: The critical study of language*. Harlow, UK: Longman.

- Fairclough, N. L. & Wodak, R. (1997). Critical discourse analysis. In T. van Dijk (Ed.), *Discourse as structure and process. Discourse studies: A multidisciplinary introduction* (pp. 258–284). London/Thousand Oaks/New Delhi: Sage Publications.
- Fasold, R. (1984). *Sociolinguistics of society*. Cambridge, MA: Blackwell.
- Fasold, R. (1990). *Sociolinguistics of language*. Cambridge, MA: Blackwell.
- Fonseca-Greber, B. (2000). *The change from pronoun to clitic to prefix and the rise of null subjects in spoken Swiss French*. Unpublished doctoral dissertation. Tucson, AZ: University of Arizona.
- Fonseca-Greber, B. & Waugh, L. R. (2003a). On the radical difference between the subject personal pronouns in written and spoken European French. In P. Leistyna & C. F. Meyer (Eds.), *Corpus analysis: Language structure and language use* (pp. 225–240). Amsterdam/New York: Rodopi.
- Fonseca-Greber, B. & Waugh, L. R. (2003b). The subject clitics of European conversational French: Morphologization, grammatical change, semantic change, and change in progress. In R. Núñez-Cedeno, L. López, & R. Cameron (Eds.), *A Romance Perspective on Language Knowledge and Use* (pp. 99–118). Amsterdam & Philadelphia, PA: John Benjamins.
- Fox, B. (Ed.). (1996). *Studies in anaphora*. Amsterdam: John Benjamins.
- Georgakopoulou, A. & Goutsos, D. (1997). *Discourse analysis*. Edinburgh: Edinburgh University Press.
- Givón, T. (1979). *Syntax and semantics: Volume 12, discourse and syntax*. New York: Academic Press.
- Givón, T. (1989). *Mind, code, and context: Essays in pragmatics*. Hillsdale, NJ: Erlbaum.
- Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.
- Goodwin, C. & Heritage, J. (1990). Conversation Analysis. *Annual Review of Anthropology*, 19, 283–307.
- Grundy, P. (1995). *Doing Pragmatics*. London: Arnold.
- Gumperz, J. J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.
- Gumperz, J. J. (2001). Interactional sociolinguistics: A personal perspective. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 215–228). Malden, MA/Oxford, UK: Blackwell.
- Gumperz, J. & Cook-Gumperz, J. (1982). Introduction: language and the communication of social identity. In J. Gumperz (Ed.), *Language and social identity* (pp. 1–21). Cambridge: Cambridge University Press.
- Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Halliday, M. A. K. (1978). *Language as social semiotic*. Baltimore: University Park Press.
- Hammond, S. C. & Gao, H. (2002). Pan Gu's paradigm: Chinese education's return to holistic communication in learning. In X. Lu, W. Jia, & D. R. Heisey (Eds.), *Chinese Communication Studies. Contexts and Comparisons* (pp. 227–243). Westport, CT: Ablex Publishing.
- Heine, B., Claudi, U., & Hunnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.
- Hill, J. & Irvine, J. (1992). *Responsibility and evidence in oral discourse*. Cambridge: Cambridge University Press.
- Holmes, J. (1992/2001). *Introduction to Sociolinguistics*. Essex, Eng: Longman.
- Hopper, P. (1982). Aspect between discourse and grammar: An introductory essay for the volume. In P. Hopper (Ed.), *Tense-aspect: Between semantics and pragmatics* (pp. 3–18). Amsterdam: John Benjamins.
- Hopper, P. & Traugott, E. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- Hudson, R. (1980/1996). *Sociolinguistics*. Cambridge: Cambridge University Press.
- Hutchby, I. & Wooffitt, R. (1998). *Conversation analysis: Principles, practices, and applications*. Cornwall: Polity Press.
- Hymes, D. (1974). *Foundations in sociolinguistics*. Philadelphia: University of Pennsylvania Press.
- Jacoby, S. & Ochs, E. E. (1995). Co-construction: An introduction. *Research on language and social interaction*, 28(3), 171–183.

- Jakobson, R. (1936/1982). Contribution to the general theory of case: General meanings of the Russian cases. In L. Waugh & M. Halle (Eds.), (1982), Roman Jakobson. *Russian and Slavic grammar: Studies 1931–1981*. Berlin: Mouton.
- Jakobson, R. (1960). Linguistics and poetics. In T. Sebeok (Ed.), *Style in language*. Cambridge, MA: MIT Press.
- Jaworski, A. & Coupland, N. (Eds.). (1999). *The discourse reader*. London/New York: Routledge.
- Johnstone, B. (2002). *Discourse analysis*. Oxford: Blackwell.
- Kennedy, G. (1998). *An introduction to corpus linguistics*. New York: Addison, Wesley, Longman.
- Klein-Andreu, F. (Ed.). (1983). *Discourse perspectives on syntax*. New York: Academic Press.
- Kramsch, C. (2002). 'How can we tell the dancer from the dance?' In C. Kramsch (Ed.), *Language acquisition and language socialization: Ecological perspectives* (1–30). London/New York: Continuum.
- Labov, W. & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helm (Ed.), *Essays on verbal and visual arts* (pp. 12–44). Seattle: University of Washington Press.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W. (2001). *Principles of linguistic change: Volume 2, Social factors*. Malden, MA: Blackwell.
- Lantolf, J. P. (2000) Introducing sociocultural theory. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning*. Oxford: Oxford University Press.
- Lazarton, A. (2002). Quantitative and qualitative approaches to discourse analysis. In M. McGroarty (Ed.), *Annual Review of Applied Linguistics, 22, Discourse and dialogue* (pp. 32–51). Cambridge: Cambridge University Press.
- Lee, D. (2001). *Cognitive linguistics: An introduction*. Oxford: Oxford University Press.
- Levinson, S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Liu, J. (2001). *Asian students' classroom interaction patterns in U.S. universities: An emic perspective*. Westport, CT: Ablex Publishing.
- Lyons, J. (1977). *Semantics 2*. Cambridge: Cambridge University Press.
- Markee, N. (2000). *Conversation analysis*. Mahwah, NJ: Erlbaum.
- McGroarty, M. (Ed.). (2002). *Annual review of applied linguistics 22: Discourse and dialogue*. Cambridge: Cambridge University Press.
- Moerman, M. (1988). *Talking culture: Ethnography and conversation analysis*. Philadelphia, PA: University of Philadelphia Press.
- Mühlhäusler, P. & Harré, R. (1990). *Pronouns and people: The linguistic construction of social and personal identity*. Cambridge, MA: Blackwell.
- Ochs, E. (1979). Transcription as theory. In E. Ochs & B. Scheffelin (Eds.), *Developmental pragmatics*. New York: Academic Press.
- Peirce, C. S. (1878/1992). Deduction, induction, and hypothesis. In N. Houser & C. Kloesel (Eds.) (1992), *The essential Peirce: Selected philosophical writings. Volume 1 (1867–1893)*. Bloomington, IN: Indiana University Press.
- Phillips, N. & Hardy, C. (2002). *Discourse analysis: Investigating processes of social construction*. Thousand Oaks, CA: Sage.
- Saville-Troike, M. (1982/1989/2003). *The ethnography of communication*. London: Blackwell.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.
- Schiffrin, D. (1994). *Approaches to Discourse*. Oxford: Blackwell.
- Schiffrin, D., Tannen, D., & Hamilton, H. E. (Eds.). (2001). *The handbook of discourse analysis*. Malden, MA/Oxford, UK: Blackwell.
- Schiffrin, Deborah (2001). Discourse markers: Language, meaning, and context. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 54–75). Malden, MA/Oxford, UK: Blackwell.
- Scollon, R. & Scollon, S. W. (1995/2001). *Intercultural communication*. Malden, MA: Blackwell.

- Stubbs, M. (1996). *Text and corpus analysis. Computer-assisted studies of language and culture*. Oxford, UK/Cambridge/MA: Blackwell.
- Su, Z. & Su, J. (1994). Teaching and learning science in American Chinese high schools: A comparative study. *Comparative Education*, 30(3), 255–271.
- Tannen, D. (1984). *Conversational style: Analyzing talk among friends*. Norwood, NJ: Ablex.
- Tannen, D. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.
- Tannen, D. (Ed.). (1993). *Framing in discourse*. New York: Oxford University Press.
- Taylor, J. (1989/1995). *Linguistic categorization: Prototypes in linguistic theory*. Oxford: Oxford University Press.
- Thomason, S. (2001). *Language contact: An introduction*. Washington, DC: Georgetown University Press.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam/Philadelphia: John Benjamins.
- Tomasello, M. (1998). Introduction: A cognitive-functional perspective on language structure. In Tomasello 1998 (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. vii–xxiii). Mahwah, NJ: Erlbaum.
- Tomasello, M. (Ed.). (1998, 2003). *The new psychology of language: Cognitive and functional approaches to language structure*, Vols. 1 & 2. Mahwah, NJ: Erlbaum.
- Ungerer, F. & Schmid, H. (1996). *An introduction to cognitive linguistics*. London: Longman.
- van Dijk, T. A. (Ed.). (1997a). *Discourse as structure and process. Discourse studies: A multidisciplinary introduction*. Volume 1. London/Thousand Oaks: Sage.
- van Dijk, T. A. (Ed.). (1997b). *Discourse as social interaction. Discourse studies: A multidisciplinary introduction*. Volume 2. London/Thousand Oaks: Sage.
- van Dijk, T. A. (1998). *Ideology: A multidisciplinary study*. London: Sage.
- van Dijk, T. A. (2001a). *Critical discourse analysis*. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 352–371). Malden, MA/Oxford, UK: Blackwell.
- van Dijk, T. A. (2001b). Multidisciplinary CDA: A plea for diversity. In R. Wodak & M. Meyer (Eds.), (2001), *Methods of critical discourse analysis* (pp. 95–120). London: Sage.
- van Lier, L. (2000). From input to affordance: Social-interactive learning from an ecological perspective. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 245–259). Oxford: Oxford University Press.
- van Lier, L. (2004). *The ecology and semiotics of language learning: A sociocultural perspective*. Boston/Dordrecht: Kluwer Academic.
- Verschueren, J. (1999). *Understanding pragmatics*. London: Arnold.
- Vickers, C. (2004). *Interactional accommodation and the construction of social roles among culturally diverse undergraduates*. Unpublished doctoral dissertation. Tucson, AZ: University of Arizona.
- Virtanen, T. (Ed.). (2004). *Approaches to cognition through text and discourse*. Berlin: Mouton de Gruyter.
- Wales, K. (1996). *Personal pronouns in present-day English*. Cambridge: Cambridge University Press.
- Watts, R. (2003). *Politeness*. Cambridge: Cambridge University Press.
- Waugh, L. R. (1982). Marked and unmarked – a choice between unequals in semiotic structure. *Semiotica*, 38, 3/4, 299–318.
- Wodak, R. & Meyer, M. (2001). *Methods of critical discourse analysis*. London: Sage.
- Wood, L. & Kroger, R. (2000). *Doing discourse analysis: Methods for studying action in talk and text*. London: Sage.
- Yule, G. (1996). *Pragmatics*. Oxford University Press.