Cognition and Language: A Series in Psycholinguistics

Daniel C. O'Connell
Sabine Kowal

Communicating with One Another

Toward a Psychology of Spontaneous Spoken Discourse

Springer

Communicating with One Another

COGNITION AND LANGUAGE
A Series in Psycholinguistics • Series Editor: R. W. RIEBER

Recent Volumes in this Series:

THE BILINGUAL MIND:
Thinking, Feeling and Speaking in Two Languages
Rafael Art. Javier

THE COLLECTED WORKS OF L. S. VYGOTSKY
Volume 1: Problems of General Psychology
Volume 2: The Fundamentals of Defectology (Abnormal Psychology and Learning Disabilities)
Volume 3: Problems of the Theory and History of Psychology
Volume 4: The History of the Development of Higher Mental Functions
Volume 5: Child Psychology
Volume 6: Scientific Legacy

EXPERIMENTAL SLIPS AND HUMAN ERROR:
Exploring and Architecture of Volition
Edited by Bernard J. Baars

LANGUAGE FOR THOSE WHO HAVE NOTHING
Mikhail Bakhtin and the Landscape of Psychiatry
Peter Good

LANGUAGE, THOUGHT, AND THE BRAIN
Tatyana B. Glezerman and Victoria I. Balkoski

PSYCHOENVIRONMENTAL FORCES IN SUBSTANCE ABUSE PREVENTION
Lorand B. Szalay, Jean Bryson Strohl, and Kathleen T. Doherty

THE PSYCHOPATHOLOGY OF LANGUAGE AND COGNITION
Robert W. Rieber and Harold J. Vetter

TIME, WILL, AND MENTAL PROCESS
Jason W. Brown

VYGOTSKY'S PSYCHOLOGY-PHILOSOPHY: A Metaphor for Language Theory and Learning
Dorothy Robbins

VYGOTSKY'S SOCIOHISTORICAL PSYCHOLOGY AND ITS CONTEMPORARY APPLICATIONS
Carl Ratner

COMMUNICATING WITH ONE ANOTHER: Toward a Psychology of Spontaneous Spoken Discourse
Daniel C. O'Connell and Sabine Kowal

Daniel C. O'Connell • Sabine Kowal

Communicating with One Another

Toward a Psychology of Spontaneous Spoken Discourse

Foreword by Donelson E. Dulany

Springer

Daniel C. O'Connell
Georgetown University
Washington, DC
USA
doconnell@jesuits-mis.org

Sabine Kowal
Technische Universität
Institut für Sprache und Kommunikation
Berlin
Germany
kowal-berlin@t-online.de

ISBN: 978-0-387-77631-6 e-ISBN: 978-0-387-77632-3
DOI: 10.1007/978-0-387-77632-3

Library of Congress Control Number: 2008924420

AN ACADEMIC DEBT OF RECOGNITION

Dedicated to the memory of Professor Hans Hörmann, mentor to both of us, and to future generations of scholars, that they may have the courage he endeavored to teach to us by his example: to be honestly and steadfastly open to empirical evidence. He was ever mindful of what our task is: to learn more about how people use language to communicate with one another – to work toward a psychology of language use.

A PERSONAL DEDICATION

To two women who have inspired us with their integrity and kindness and wisdom: to our sisters, Mary Virginia Sabin and Christiane Oberg

Foreword

A reader with any background in psycholinguistics should find this book intellectually stimulating and important, as well as perhaps provocative in some ways. My own interest in psycholinguistics came early before I turned to analysis of the roles of conscious states and development of a mentalistic metatheory – a metatheory I see as quite consistent with the authors' theoretical proposals. What I see throughout this book is the authors' vision of a need to move psycholinguistics to the analysis of more psychologically significant uses of language. That indeed is a very significant and fundamental aim when so much of cognitive psychology has sacrificed significant content in the search for lawfulness and rigor. Indeed, what could be more fundamental than to "reinstate genuinely psychological considerations to the analysis of spontaneous spoken discourse" (Chapter 23, p. 217)? Psycholinguistics is, as the very term reveals, interdisciplinary in substance, and the Chomskyan (1957) revolution in linguistics has long inspired mainstream psycholinguistics to focus conceptually on the way in which linguistic structure functions in the use of language, with the contribution from psychology being largely methodological – a focus on the individual speaker or writer within a controlled experimental environment. So a summary of the mainstream psycholinguistic focus is given in a few terms: monologue, linguistic concepts, and an artificial experimental environment.

What then is the contrast proposed throughout, in the theory they develop and in their own and others' empirical research presented here? The authors provocatively lay out what they see as limitations of mainstream psycholinguistics. There should then be a shift of emphasis to an exploration of more significant psychological conceptions of intentions, feelings, beliefs, and meanings – significant states of awareness – that can animate the speaker and the hearer as well. Indeed, the authors lay out the groundwork for the further development of theory: intersubjectivity, perspectivity, open-endedness, and verbal integrity. For these psychologically more significant aspects of language use to be revealed, there would then be a shift of empirical focus from monologue to dialogue. In fact, we should look for a shift from the artificiality of our insignificant little experimental tasks to real spontaneous dialogue in field

observation – as, for example, in political interviews. That kind of shift is also intrinsically a focus on oral rather than written language.

How could these phenomena be revealed empirically? The authors draw on the work of others and on 40 years of their own research together. What is psychologically significant can often be revealed in rhetoric – and as the authors put it, "Prosody is one of the most important tools of oral rhetoric" (Chapter 5, p. 50): pauses, stress, intonation, intensity, duration, and articulation rate. These, as well as semantic content and assertion forms, can convey meaning and a sense of reference from the speaker as a thematic "center of gravity". And how should we regard pauses and fillers such as "uh" or "um"? As departures from a model of the "ideal speaker," on a mainstream assumption, or as sometimes facilitating the articulation of the speaker as well as the understanding of the hearer? For that matter, when is an interjection an interruption and when is it only natural turn-taking? And we should not forget that laughter and even booing are communicative. All of these things are topics elaborated by the authors conceptually and in the empirical evidence they review.

Now here is what I think would be an especially provocative and intellectually stimulating questions for the reader: Is what the authors are proposing a classical Kuhnian paradigm shift in psycholinguistics? Kuhn originally (1962) used the term "paradigm" in enough different senses to occupy scholars for decades, but the core and influential meaning has been that of generally accepted metatheory, theory, and methodology that guides research in a field – in Kuhn's term, a "consensually validated modus operandi." It is important to ask, then, whether the conceptual and empirical approaches they illustrate and promote are complementary to mainstream approaches or irreconcilable. For that matter, could some of the proposed shifts be complementary and some others irreconcilable? On my reading that is closest to the authors' view, a view that seems reasonable to me. The shift from a behavioristic to a cognitive paradigm came with some evidence that was irreconcilable with general stimulus–response theory – for example, evidence for complexity of memory and action structures as well as evidence against direct and automatic strengthening by reinforcement. But some see connectionist models (a version of S–R theory re-embodied) as complementary to information processing models, and behaviorists and cognitivists alike have employed manipulation response experimental designs and interpreted subject reports within a physicalistic data language. What is to be said of the relation of the mainstream and proposed psycholinguistic paradigms?

After his early theological studies in Austria, Dan O'Connell entered the PhD program here at the University of Illinois – where we collaborated on an experimental and theoretical analysis of the causal role of verbally reported conscious states in concept development (Dulany & O'Connell, 1963). This was before he left for a postdoctoral appointment at Harvard Center for Cognitive Studies, further work in Germany, and faculty appointments at St. Louis University (including President), Loyola University of Chicago, and Georgetown University (including Department Chair). Sabine Kowal studied at the

Free University of Berlin, received her PhD at St. Louis University, and has been for many years at both the Technical University of Berlin and the Anna Freud Oberschule in Berlin. O'Connell and Kowal have been long time collaborators in psycholinguistic research. As Editor of the *American Journal of Psychology,* I published the authors' 2003 "Half Century of Monologism," their description of mainstream psycholinguistics and the approach much further elaborated in this present book.

I believe the reader will see what I see in this work – a thesis that is intellectually stimulating and significant, and in fact, provocative in ways that can be important for the discipline when examined.

By Donelson E. Dulany

Department of Psychology,
University of Illinois
Champaign/Urbana, Illinois

Preface

We wish the academic dedication of our book to be a recognition of Professor Hans Hörmann, mentor to both of us at the Free University of Berlin. The first author, Daniel C. O'Connell, was a Humboldt Fellow there during the 1968–1969 academic year. It was a year of academic protests, strikes, and endless debate over rights and duties. Altogether, it was a time of chaos and scholarly stalemate, not the best year to appear in a Berlin beleaguered by both the cold war and now a student revolt.

It must have been January of 1969 that Professor Hörmann approached Sabine Kowal, a student at that time, in the corridor of the Psychological Institute of the Free University of Berlin with a suggestion that she work with him and O'Connell on a research project, since O'Connell was free to carry on his research unimpeded. The reason for this freedom was that none of the protesters wanted to confront the American or risk reprisals.

Sabine Kowal did go to see O'Connell. Within a year, the new team had attended two international psycholinguistic conferences (Schloss Lehen near Heidelberg, Germany, and Bressanone, Italy) and had published twice (O'Connell, Kowal, & Hörmann, 1969, 1970). That was almost 40 years ago, and O'Connell and Kowal are still engaged in research together.

Hörmann (1967) had quickly become a critic of the new psycholinguistics that emerged at the middle of the twentieth century. And it was precisely his attitude of questioning and criticizing that both O'Connell and Kowal learned from him. But there was more: Beneath his deep personal reserve, Hans Hörmann offered to us a sort of fatherly expectation of excellence and exemplified for us a disciplined and stubborn dedication to empirical discovery. Both of us are deeply grateful to him for his encouragement and his mentoring of our fledgling efforts to learn in a genuinely psychological framework about how people use language to communicate with one another.

Many others have contributed along the way to the growth of this project. Notable among them was Frieda Goldman-Eisler (1968), whose interest in pauses was our own starting point. Dozens of students and colleagues have worked with us. And a number of institutions and foundations have invested in us. Among the universities that have supported our research are St. Louis University (alma mater to us both), Loyola University of Chicago, Georgetown

University, the Free University of Berlin, and the Technical University of Berlin; among the foundations and agencies are the Alexander von Humboldt-Stiftung, the Deutsche Forschungsgemeinschaft, the Deutscher Akademischer Austauschdienst, the National Science Foundation, the National Institutes of Health, and the Fulbright Commission. There is one more group of colleagues who deserve special thanks. Donelson E. Dulany of the University of Illinois (Champaign/Urbana), Robert W. Rieber of CUNY and Fordham Universities, Camelia Suleiman of Florida International University, and Kevin Weinfurt of Duke University have plowed through the preliminary ms. with us and offered both encouragement and constructive criticism along the way. Don Dulany has also graciously agreed to write a Foreword for our book.

Responsibility for the shortcomings of the book is entirely ours, but let it also serve as a sign of our gratitude to all who have contributed in any way to its existence. And may it subserve the greater good by clarifying issues, providing empirical ways and means of addressing them, and coming up with a few useful conclusions for the overall good of a psychology of language use. The reader will quickly note, especially since we begin our chapters with Part I, A Critique of Mainstream Psycholinguistics, that our book is occasionally agonistic. We find this inevitable as we venture into unknown territory. And it also entails the use of many direct quotations in order to accurately reflect the positions we wish to criticize. For the cumbersomeness of such a multitude of citations, we beg our readers' indulgence.

There is one more reflection that finds its proper place here at the beginning. This book was written dialogically. That is to say that it was literally written down from our daily spontaneous spoken discourse that sometimes lasted as long as three hours. With immense gratitude to whoever thought up the technology of the SKYPE phone system, we can report that our daily regimen of work between Berlin, Germany and St. Louis, Missouri, USA was 8 a.m. and 2 p.m. in terms of CST or CDT and correspondingly 15 *Uhr* and 21 *Uhr* in *MEZ*. Another aspect of our dialogical approach is that it is also an English-language/German-language collaboration. American psycholinguistics has clearly neglected important contributions to a psychology of language use on the part of European psychologists, even though a large portion of their contributions have been written in the English language.

Washington, DC, USA — Daniel C. O'Connell
Berlin, Germany — Sabine Kowal

About the Authors

The authors are experimental psychologists who have been engaged in research together for 40 years now. Dan O'Connell studied at St. Louis University and did doctoral work at the University of Illinois (Champaign/Urbana); Sabine Kowal studied at the Free University of Berlin and did doctoral work at St. Louis University. O'Connell's career was at St. Louis, Loyola of Chicago, and Georgetown Universities, while Kowal's was at both the Technical University of Berlin and the Anna Freud Oberschule in Berlin. For many years, the team was oriented toward mainstream psycholinguistics and experimental research on speech production. Throughout the last decades of the twentieth century, their interest shifted to spontaneous spoken discourse under field observational conditions. This shift had as its origin their observation that professional speakers known for their eloquence in public dialogue violate both ideal delivery and syntactic well-formedness – concepts established in mainstream psycholinguistics as norms for effective communication. O'Connell and Kowal have ascribed the use of these norms to a written language bias and have accordingly turned their attention – both empirically and theoretically – to the use of genuine spoken discourse. Radio and TV political interviews have provided much of the empirical database for their recent research, and their emphasis on spontaneous spoken discourse has led to the investigation of neglected speech phenomena such as fillers, pauses, interjections, and laughter in both English- and German-language corpora.

Contents

Chapter 1
The Problematic

Psycholinguistics is the name given to the study of the psychological processes involved in language . . . psycholinguistics is closely related to other areas of cognitive psychology, and relies to a large extent on the experimental methods used in cognitive psychology (Harley, 2001, p. 3 f.).

Human language is language because it is used by people for a purpose, namely, to live with other people (Hörmann, 1981, p. 303).

Chapter Prospectus

Chapter 1, *The Problematic*, sets forth the state of the question. For the most part, our research has developed over the years as a critique of mainstream psycholinguistics. Mainstream psycholinguistics has engaged primarily written discourse, has based its limited engagement of spoken discourse on a written language bias, has accepted the autonomy of language throughout the entire realm of discourse, has not taken into account the communicative context of language use, and has largely neglected the research of its own historical precursors. Spontaneous spoken discourse (i.e., genuinely meaningful use of spoken language among human beings for purposes of communication) is looked upon in such a tradition as deficient in comparison with written discourse. In the present book, spontaneous spoken discourse is considered to be lawfully structured in its own right, and therefore eligible for the premier research role in language use insofar as it is the universal, original, and basic form of human discourse.

What's It All About?

How does one get at what people do with spoken language, at why and how they use it in the act of communicating with one another? Scholars have fussed and fumed over these questions ab initio. Their importance is quite simple: The best

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_1, © Springer Science + Business Media, LLC 2008

way to know what people are thinking and feeling is to listen to them and interrogate them – in words – regarding what is being communicated. And so, how people use language becomes crucial as soon as we begin to operate on a human level. Such commonplace verbalizations as "Would you mind repeating that?" and "Just what do you mean by that?" give eloquent voice to our everyday interest in how people use language when they address one another.

And yet, just accepting speaking and listening as eminently human activities does not throw light on why and how people engage either of these activities. Why do they begin to speak at all in the midst of their other preoccupations with everyday life? Why do people tell stories? Why do mothers use words to babies who don't understand words? How do people terminate conversations? What does "Oh!" mean in a specific context, and what does it add to the ongoing discourse? How do we know when someone is lying? Or can we really? Every one of these questions is a worthy candidate for research on the use of language. But how do we – precisely as psychologists – go about research on language use?

Historical Beginnings of Mainstream Psycholinguistics

At mid-twentieth century, a convergence of historical developments contributed to a new birth of interest in these questions. There was something of a post World War II malaise within psychology, a discontent over behaviorism that left a vacuum to be filled. As it turned out, the filler came not from within psychology – at least as regarded the questions of language use – but from the neighboring science of linguistics. Noam Chomsky (1957) embodied the movement that came to be known as transformational or generative linguistics. His influence on a psychology of language use came to be conceptualized under the title psycholinguistics. The word itself predated Chomsky and is traceable back to the 1940s (see Rubenstein & Aborn, 1960), but the key insight was Chomsky's, namely that "an understanding of grammar must be central to any serious understanding of the human mind" (American Psychological Association [APA], 1985, p. 286). This is a hugely bold and radical statement of principle. The APA simply accepted it as "demonstrated" by Chomsky, but what that demonstration consisted in has never been clarified. On the face of it, it seems to be diametrically the opposite of Ong's (1982, p. 7) statement that:

> Computer language rules ('grammar') are stated first and thereafter used. The 'rules' of grammar in natural human languages are used first and can be abstracted from usage and stated explicitly in words only with difficulty and never completely.

Is then grammar really the appropriate starting point for the investigation of the human mind, or even more specifically, as it was to be billed in the emerging discipline of psycholinguistics, for a psychology of language use? Or did Chomsky have the cart before the horse?

The More Remote Background of Modern Psycholinguistics

Perhaps it would help to step back from the immediate origins of modern mainstream psycholinguistics for a moment to consider the ways in which the question of how people use language has been answered through the ages. For indeed, an inveterate tradition of observing and collecting linguistic frequencies, mistakes, commonalities, exceptions, specific forms of language use, and, yes, even rules has been in use by students of language for millennia. The study of language did not spring full grown from the head of Chomsky at mid-twentieth century.

A considerable store of systematic knowledge about language use actually predates modern psycholinguistics and the mid-twentieth century. In fact, many of the great psychologists of the nineteenth and early twentieth centuries engaged language use. These would include Karl Bühler, Jaymond M. Cattell, Sigmund Freud, William James, O. Hobart Mowrer, Jean Piaget, Clara and William Stern, Lev S. Vygotsky, and Wilhelm Wundt. Much of this historical background has been summarized in Blumenthal (1970), and developmental aspects of it have been reviewed by McCarthy (1954). But most of these precursors are not even mentioned in recent mainstream psycholinguistic texts: Cutler (2005) and Field (2004) have mentioned none of them; and Harley (2001) has mentioned only Vygotsky. Suddenly at mid-century, the Zeitgeist shifted and the principles used by these early scholars, along with their findings, became somehow irrelevant and were ignored. In fact, Knobloch (2003, p. 29; our translation), in his review of the history of psycholinguistics, has concluded that "it is overflowing with promising, but then abandoned research programs . . . . The mountain of unsolved problems and untested models left in its wake by psycholinguistics notably continues to grow."

The Cognitive Revolution

Evidently, more than just a convergence of Chomsky and weariness with behaviorism took place in mid-twentieth century with respect to the way in which questions were to be asked about language use. Baars (1986) has referred to it as *The cognitive revolution in psychology*, although it still remains unclear whereunto this revolution has come. In fact, much of the vaunted abdication of behavioristic methods has amounted to mere posturing; many of the breakthroughs have evaporated or been reduced to meaningless abstractions; and in fact, the very meaning of the terms *cognitive* and *cognition* has become so eviscerated as to now be a mere shiboleth. As one cognitive scientist (Fodor, 2000, p. 100) has put it:

> What our cognitive science has done so far is mostly to throw some light on how much dark there is. So far, what our cognitive science has found out about the mind is mostly that we don't know how it works.

Graumann (1987, p. 58; our translation), a social psychologist, has criticized the individualism that a cognitive approach has produced within social psychology:

> This "individualism" of psychological [rather than sociological] social psychology is not at all limited to the methodological; the favorite theories and models are almost all centered on the individual. The prototype of the social relationship is the dyad; however, the other is often reduced to only a representation within the one: the other as "cognition."

According to Quasthoff (1995, p. 4), herself a linguist: "*Psychology* focuses on the cognitive processing of knowledge as the basic function of communicative processes." Her narrow focus on "the cognitive processing of knowledge" also described quite well the limited spectrum of psychology engaged by mainstream psycholinguistics.

Presently, we are left with a cognitive science in which "meaning is a core unsolved problem" (Fitch, 2005, p. 395), in which "psycholinguistic data is irrelevant to formal linguistic theory" (Boland, 2005, p. 23), in which neurolinguistic research "has not advanced – in an *explanatorily significant* way – the understanding of either linguistic theory or of neuroscience" (Poeppel & Embick, 2005, p. 104), in which there exists "a deep dissatisfaction about the psycholinguistic quality of most neuroimaging studies on language" (Hagoort, 2005, p. 157), and in which "consensus at any concrete level has been elusive" (Crocker, 2005, p. 363) for extant models of language comprehension. It should be noted that all these expressions of criticism originate not in opponents to mainstream psycholinguistics, but from contributors to a single volume (Cutler, 2005) with its own origins deeply rooted in the mainstream psycholinguistics of the Max Planck Institute (MPI) of Psycholinguistics in Nijmegen, The Netherlands. If current publications are any indication, then both the recent *Cognitive linguistics* (Evans & Green, 2006) and the *Handbook of latent semantic analysis* (Landauer, McNamara, Dennis, & Kintsch, 2007) have provided additional evidence, respectively, of a continuing popularity for formal and mathematical approaches to meaning in the cognitive tradition.

Some Current Criticisms of Mainstream Psycholinguistics

If one looks further, outside the tradition of mainstream psycholinguistics, additional criticisms must be taken into account. Sabin and O'Connell's (2006, February 15) review of Cutler's (2005) *Twenty-first century psycholinguistics: Four cornerstones* has criticized its failure to engage "what a speaker intends to say or means" (p. 8) and "the concepts of intention, meaning, consciousness, purpose, perspective, dialogue, social role, culture, affect, and finality" (p. 9). It is difficult, if not impossible, to imagine how communicative language use could ever be comprehensively and profitably engaged without taking these concepts into account. However, this litany of perhaps seemingly disparate elements can be traced to a more unitary objection to modern

mainstream psycholinguistics, namely its overwhelming emphasis on the language system itself, on the syntax and well-formedness of sentences. Herrmann (2005, p. 78; our translation) has commented on this emphasis in somewhat different terms as follows: "The psycholinguistic *mainstream* hardly ever gets beyond investigating the mental and neural processes involved in the production and reception of individual sentences."

The Centrality of Grammar in Mainstream Psycholinguistics

What went wrong? Or was mainstream psycholinguistics wrongheaded from the beginning? We wish to argue that it has been wrongheaded from the very beginning of the psycholinguistic era in mid-twentieth century and from the inauguration of the cognitive revolution as well. The erroneous rationale upon which the whole edifice was based can be seen in a passage from Chomsky's (1957, p. 106) foundational *Syntactic structures*: "The most that can reasonably be expected of linguistic theory is that it shall provide an evaluation procedure for grammars.... Grammar is best formulated as a self-contained study independent of semantics." Such a proclamation should have been enough to scare away any self-respecting psychologist of language use, but it was not. Quite the contrary, given the malaise over behaviorism and a fresh interest in higher processes, and given the penchant for meticulous experimental analysis on the part of some folks at Harvard University, just down the road from Massachusetts Institute of Technology, it became a heady, seductive potion. Notable among these folks were George A. Miller and his students. We are convinced that the real attraction of this veritable mystique was an almost magical reductionistic clarity. But the baby was thrown out with the bath; semantics – what people actually talk about and listen for – was reduced to a role derivative from and subordinate to grammar.

The problem was not grammar itself; language use does involve grammar, very importantly. The problem lay in the very reductionism that came to constitute the gospel of the new discipline. The reductionistic clarity did not derive from syntax alone, but from the arbitrary restriction to *syntax on the written page*. It is easy to work with well-formed sentences that appear on the written page precisely because sentences on the written page have already been formulated out of the richness of our overlearned, alphabetized literacy – the product of years of schooling. And so, the well-formed written sentence became the unit of empirical analysis for the modern psycholinguist, and the cognitive processing of such materials by experimental subjects became the focus of psycholinguistic theory. That such civilized literacy, not the use of oral language as such, was leading the parade for psycholinguists has never, to our knowledge, been acknowledged by them. To modern scientists, enveloped in our scientific approach to language use, how an analphabetic speaker or listener can use language seems hardly imaginable. Or, as Ong

(1982, p. 2) has expressed it: "We – readers of books such as this – are so literate that it is very difficult for us to conceive of an oral universe of communication or thought except as a variant of a literate universe." And we are also so literate, as Miller and Weinert (1998, p. 378 f.) have insisted, that we find it hard to acknowledge that children do not learn Chomsky's "magnasyntax" (p. 378) as the basis of their first language: "As their first language children do not acquire the written variety of their native tongue but the structures and vocabulary that they hear in the spontaneous speech around them" (p. 379). Thus, the science derived from generative grammar is chained to a literate bias and cannot see beyond it. And yet, millions of speakers and listeners use language quite articulately and eloquently everyday without the benefit of alphabetization or literacy of any kind.

Ideal Delivery: A Corollary of Syntactic Well-formedness

These real, but implicit underpinnings of modern psycholinguistics are shown more clearly in a corollary to well-formedness in written discourse. Thus, Chomsky's (1965, p. 3) classical ideal speaker speaks like a written page:

> Linguistic theory is concerned primarily with an ideal speaker/listener, in a completely homogeneous speech community, who knows his language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance.

This linguistic principle was then translated into psycholinguistic theory by Clark and Clark (1977, p. 261):

> For there to be a speech "error" there must be a "correct" way of executing a sentence, and this will be called the *ideal delivery*. When people know what they want to say and say it fluently, they are giving an ideal delivery.

Chomsky's and the Clarks' conditions for the legitimate use of oral language place both linguistic and psycholinguistic theory squarely in a never–never land of well-formed sentences on the lips of ideal speakers and do not enable us to learn about how people actually speak and listen effectively in a real world. Or, as Harris (1981, p. 33) has put it: "The 'ideal speaker–hearer,' it might appear, is in fact a communicational cripple." Even earlier, Abercrombie (1965, p. 1) had emphasized the differences between conversation and what he referred to as "spoken prose" (i.e., a written text read aloud). After invoking a litany of all the grammatical, indeed illiterate horrors that appear when conversation is committed to paper, he concluded:

> But of course it *should* be illiterate – literally. It should be different from written language. We are so used to deriving our notions of what is correct and logical in language from prose, that we find it hard to realize that a quite different set of standards must be applied to conversation. (p. 6)

More than a third of a century later, spontaneous spoken discourse is still notoriously conceptualized as flawed and inefficient: "The language of dialogue is disorderly compared to the straightforward sentences of monologue" (Garrod & Pickering, 1999, p. 10).

The fact of the matter is that spontaneous spoken discourse is orderly in multitudinous ways that transcend the sterile lawfulness of syntax or grammar – or well-formedness and the ideal delivery. It is neither necessary nor even possible for an idealized grammar to carry the entire burden of orderliness in the communicative use of spontaneous spoken discourse. Furthermore, these other-than-grammatical forms of psychological orderliness or lawfulness are discoverable by means of the traditional principles and tools of the scientific method. The question still remains: Whence comes the bias in favor of the laws dictated by grammar, to the exclusion of the richest treasures of orderly, intelligible data derivable from empirical analyses of genuine language use – spontaneous spoken discourse?

In 1982, Per Linell wrote a book, entitled *The written language bias in linguistics*, in which he claimed that modern language scientists have concentrated almost exclusively on written language as the prototype of language use. More recently, Linell (2005) has published an entirely new book under the very same title and with a much more detailed criticism of the written language bias. Once again, his claim is that syntactic well-formedness has assumed a disproportionate importance. Ong (1982, p. 75) has summed up succinctly the problematic that arises from this emphasis on the written language: "It is impossible for script to be more than marks on a surface unless it is used by a conscious human being as a cue to sounded words, real or imagined, directly or indirectly." As it turns out, then, Chomsky's (1965, p. 3) characterization of the speaker's "grammatically irrelevant conditions" that arise "in applying his knowledge of the language in actual performance" should be the other way around: Speaking and listening are the primary psychological evidence, whereas the autonomous grammatical system is an abstraction derived from them. But from this autonomous grammatical system, a psychological understanding of our use of language can never be derived, precisely because it has not only been shorn of the "memory limitations, distractions, shifts of attention and interest, and errors," but, even more importantly, because it has been isolated from context, prosody, dramatis personae, and a host of other paralinguistic and extralinguistic phenomena that are constitutive of the reality of all interactive behavior. Hence, Chomsky's characterization has nothing to do with the actual bodily embeddedness of spontaneous spoken discourse.

Recently, there has been much ado about research carried out by Daniel Everett on the language of the Pirahã, a remote Amazonian tribe (Colapinto, 2007, April 16; Grossman, 2007, June 10). The issues have even been dated in jest as B.C. (Before Chomsky) and A.D. (After Dan). The claim made by Everett's followers is that Chomsky's requirement of recursiveness as a universal property of human languages has now been proven false insofar as the Pirahã language does not make use of syntactic recursiveness at all. It appears that Everett has made his case, despite the protestations of Chomsky's

followers. We wish here to make the additional point that Everett's claim misses a more important deficiency in Chomsky's system: As we have outlined above, his theory is incapable of dealing with spontaneous spoken discourse; in fact, the inability of such a theory to deal with spontaneous spoken discourse has been swept under the rug of twentieth-century linguistic fad. More than a quarter of a century ago, however, Hörmann (1981, p. 315) had already rejected Chomsky's claim that generative grammar is relevant as an explanation of the meaning and understanding characteristic of everyday spoken discourse:

> The claim has revealed itself as an idle promise; the fact is that a theory of language developed on purely rational grounds, as is the case with generative transformational grammar, discloses its inadequacy as soon as it is exposed to the crucial test of its predictive power, i.e., the power to predict events as they occur in everyday life.

Chomsky and the modern mainstream psycholinguists are – at best – performing an autopsy on a cadaver, rather than dealing with the actual involvement of interlocutors in the bodily liveliness of spontaneous spoken discourse. The cadaver metaphor, interestingly enough, has a history of its own. Linell (2005, pp. 9, 196, footnote 11) has traced it back to von Humboldt's (1841–1852/1969, p. 419; our translation) comment that "dissection into words and rules is nothing more than a defunct concoction, the consequence of scientific dismemberment" and his claim (p. 186; our translation) that "genuine language is to be found only in articulated speech; the grammar and the lexicon are hardly comparable to a lifeless skeleton of speech." Linell has also cited Bakhtin's (1981, p. 292) description of the written: "All we have left is the naked corpse of the word, from which we can learn nothing at all about the social situation or the fate of a given word in life." Voloshinov (1973, p. 71) and Firth (1968, p. 47) have added their own descriptions of the written and the linguistic analyses thereof, respectively, as cadaverous. We would like to include in this history the description by Rommetveit (1974, p. 61), credited by him to Birdwhistell (1971):

> Birdwhistell argues that what is preserved in typed transcripts of face-to-face dialogues is in fact only 'the cadaver of speech'. And an essential part of what is lost in the transcription has to do with what Roman Jakobson refers to as meta-linguistic operations, i.e., with shifting premises of communication conveyed by, for example, body movement, gesture, facial expression, and tone of voice.

Or, as the essayist Siri Hustvedt (2006, p. 102) has put it quite bluntly: "In every book, the writer's body is missing."

The Users of Language

Bühler (1934/1982) has designated his theory of language itself as a *Sprachtheorie* and his theory of language use as an *Organon* or tool theory. Banal as it may sound, "Language is a tool, defined as to its use by the people who use it" (O'Connell, 1988, p. 62 f.). *Meaning and understanding* exist only in the psyche

of man, as Hörmann (1976, 1981) knew full well when he so entitled his book to emphasize these absolutely basic processes of language use; on paper alone, words remain "mere semantic potentialities" (Rommetveit, 1974, p. 87). The confusion of the two realms – potentiality on paper vs. actuality in the psyche – has for many years now pinpointed the locus of the problematic we are here discussing. Isolation of words and sentences in a theory that relies on the claim of an autonomous syntactic structure has yielded nothing but confusion about a genuine psychology of language use, or, as Hörmann (1981, p. vii) has written more than a quarter of a century ago: "The models and theories current in linguistics tend to approach a level of complexity at which extreme sophistication borders on folly." What have been forgotten are simply the agents in all this, speakers and listeners. And hence, we should acknowledge at the outset that the rest of this book should be read not as a treatise on spontaneous spoken discourse as some free-standing phenomenon, but as a treatise – as should be every psychological essay – on *people* speaking and listening interactively, meaning and understanding. In our own view, such an emphasis constitutes the only way imaginable of integrating language use comprehensively into psychology. Apart from "*an intersubjectively established, temporarily shared social world*" (Rommetveit, 1974, p. 29) among people, no speaking and listening can occur. This is precisely why Graumann (1984, p. 247; our translation) has insisted that a psychology of language use must be subsumed within a social psychology: "Every science of signs must articulate and axiomatize the domain of social living, where alone signs exist and are of consequence."

Modern mainstream psycholinguistics, on the other hand, has become dependent to such an extent on the discipline of linguistics as to be considered by Herrmann (2005, p. 12 f.; our translation) itself a linguistic discipline in contradistinction to a psychological discipline:

> The psycholinguists (as linguists) generally take as their point of departure the language system itself and subordinate the development of their theories regarding language use to linguistic conceptualizations: How is language instantiated in people? On the other hand, the psychologists of language use (as psychologists) see as their primary task the integration of language use into psychological events.

Fifteen years ago, Reyna (1993, p. 23), in assessing the relevance of linguistics for psychology, expressed the following warning: "For psychologists, therefore, the ultimate concern is that linguistic theory might bear no relation to behavioral reality – that it is just an abstract game with symbols." One more example of perhaps inordinate dependence on linguistics can be found in Quasthoff's (1995, p. 3) reductionistic definition of human communication as "in essence – albeit not exclusively – the mutually oriented vocal production and reception of linguistic signs." As we proceed through the pages of this book, it will become more and more evident that it is quite questionable whether the two approaches can ever be reconciled to one another as complementary contributions to a unitary psychology of language use.

The Autonomy of Language

We have referred already to the overwhelming emphasis on the language system itself and to the claim of autonomous syntactic structures. These themes are conceptualized by mainstream psycholinguistics as extremely important in the search for an answer to the question: "How is language instantiated in people?" (Herrmann, 2005, p. 12; our translation). The claim that the language system is autonomous is in turn related to the claim of nativism (language as somehow biologically determined) and the doctrine of telementation, as Taylor (1997, p. 3) has named it: "Language must 'give' us what we 'get' through communication." Almost 30 years ago, Reddy (1979, p. 290) criticized the concept more thoroughly as the *conduit* metaphor:

> (1) Language functions like a conduit, transferring thoughts bodily from one person to another; (2) in writing and speaking, people insert their thoughts and feelings in the words; (3) words accomplish the transfer by containing the thoughts or feelings and conveying them to others; and (4) in listening or reading, people extract the thoughts and feelings once again from the words.

O'Connell (1988, p. 52) has summarized the conduit metaphor as follows: "Words simply *carry* information from speaker to hearer." This doctrine constitutes a fundamentalistic application of Shannon and Weaver's (1949) *The mathematical theory of communication* to language processing. Linell (1982, p. 146) has renamed it as the *translation* or *recoding* theory, according to which meaning can be derived "by applying a linguistically correct analysis to these linguistic products." But, meaning cannot simply be milked from linguistic structure, because meaning is not entirely pre-existent to the utterance itself, and understanding never exhausts intended meaning. More importantly, speaking and listening both involve a multitude of other systems, not just the linguistic system. And accordingly, meaning and understanding are far more than the two ends of a pipeline through which autonomous linguistic structures travel unscathed by the human psyche. We wish to take aim at this conviction of many linguists and psycholinguists alike, namely, that words are simply *carriers* or *media* of information from speaker to listener.

In all this, the problem of the interactive agents as units of analysis remains. For it is precisely the interaction, not the action of any individual interlocutor as such, that is the core principle of a dialogical theory of spontaneous spoken discourse. How this transcendence of the individual in the dynamic interaction should be conceptualized and empirically investigated is still an open question.

Chapter 2
Empirical Methods

The development of psycholinguistic methods over the past 50 years reveals a clear tendency: a shift away from methods involving the collection and analysis of uncontrolled or genuine data toward experimentation (Dietrich, 2002, p. 14; our translation).

Chapter Prospectus

Chapter 2, *Empirical Methods*, is a corollary to Chapter 1 insofar as the methods of research characteristic of mainstream psycholinguistics have been largely limited to controlled, artificial, laboratory-based experimentation and quantitative analyses of findings. Field observational research involving spontaneous spoken discourse in naturalistic communicative settings has been radically neglected, and the relevance of qualitative analyses has correspondingly been downplayed.

Homo Loquens et Audiens

We have suggested at the end of Chapter 1 that modern mainstream psycholinguistics and a psychology of language use may be, in their present incarnations, irreconcilable. The rationale given for such a statement has been the basic principles, goals, and epistemology of the two approaches to language use. Currently, mainstream psycholinguistics concentrates on the language system itself, as instantiated in verbal expressions of cognitive processes. But it is our contention that a genuine psychology of language use must begin with a concentration on the people whose speaking and listening are being investigated in meaningful communicative contexts.

The last phrases of the foregoing paragraph have a subtle importance that is frequently overlooked. In the present book, we set out to engage not *homo linguisticus*, but *homo loquens et audiens*, or, as Herrmann (1985, p. 41; our translation) has expressed it, not "the *human being as language processor*," but

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_2,

"the *human being who also understands language and also speaks*." It should be noted that Herrmann has specifically credited both Bloomfield (1933) and Vygotsky (1934/1962) for these insights (see also O'Connell, 1988, p. 54). The point to be made here is that human beings speak and listen *only occasionally*, not – as a function of their nature – uninterruptedly. Hörmann (1981, p. 191) has stated it as follows: "Speaking must be seen as the continuation of action by alternative means." Such an intermittent recourse to the use of language is thus part of a comprehensive pursuit of intelligibility and relevance in the ambient world. As Bühler (1934/1982, p. 158) put it many years ago, as explication of his concept of the "*empraktische* Gebrauch von Sprachzeichen" ("*empractical* use of language signs" [Bühler, 1990, p. 179]):

> Islands of language emerge from within the sea of silent but unequivocal communication at the places where a differentiation, a *diacrisis*, a decision between several possibilities has to be made, and easily can be made by interspersing a word. (p. 176)

Bühler (p. 176) has provided as an example a railroad passenger's one-word request for a transfer. The ellipsis suffices precisely because it is embedded in a rich context, which carries the burden of meaning that is carried by words in a written text. In other words, the social context involved in this speaker/listener exchange is amplified by a nonverbal context: the very presence of the speaker on a train addressing the conductor – whose response is most likely to be nonverbal. A literary example of such language usage is to be found in a novel by Philip Roth (1959; cited in Page, 1988, p. 8):

> There was not much dinner conversation; eating was heavy and methodical and serious, and it would be just as well to record all that was said in one swoop, rather than indicate the sentences lost in the passing of food, the words gurgled into mouthfuls, the syntax chopped and forgotten in heapings, spillings, and gorgings.

In Chapter 1, we have mentioned Graumann's (1987) critique of the individualism of cognitive social psychology. More relevant in a methodological context is his claim that a psychology of language use must be subsumed under the broad rationale of social psychology. We too are convinced that the proper local habitat for a psychology of language use must be social psychology rather than cognitive psychology. However, we wish to emphasize as well that social psychology itself must be somehow compatible with the psychological understanding of the individual. Research on the social aspects of language use must never disregard the principles that formulate the dynamic activity of the individual speaker or listener, precisely because the social activities of the individual must be built upon the capacities and properties of the individual psyche. Hence, research and theory regarding spontaneous spoken discourse must be integrated with an all-embracing, comprehensive, general psychology of the individual. In other words, speaking and listening, meaning and understanding, can be properly contextualized, investigated, and understood only within the larger ambit of man's psychological, social, and cultural life. Short of that context, both research and theory gravitate back to the protective cove of

an isolated language system once again. In fact, Graumann (2006) has harbored a fear that an American version of social psychology is still inadequate to deal with language use. He has argued that American social psychology traces its heritage back all the way to Floyd H. Allport's (1924) emphasis on the individual. Or, as Graumann (2006, p. 56 f.; our translation) has put it: "What we call psychology seems to be inseparably associated with the concept of the individual." Hence, he has preferred the designation "cultural psychology" rather than social psychology as the local habitat for a psychology of language use.

Transmittal of Data by Transcribers

The social and general venues of a psychology of language use specifically influence the type of question that can legitimately be asked in research as well as the applicable methods available for answering the questions. Since spontaneous spoken discourse is essentially ephemeral, research depends on the availability of accurate transcripts. But such availability depends in turn upon the intermediate step of transcribing, which is itself an example of a very specific type of language use and can be problematic. In a study of transcription, O'Connell and Kowal (1994, p. 129) have analyzed transcripts made by transcribers in various settings. One such transcript was made by a native speaker of and teacher of German. The transcript had been prepared for an American colleague, a non-native speaker of German, who was also a teacher of German. The transcript was intended for use in his German classes in America. Without adverting to or informing the researchers of her decision, the transcriber corrected errors and redundant repetitions as she transcribed:

> Her expressed, self-instructed purpose was to produce a transcription of correct German for language instruction, and she described her procedure as involving deletion of much that was colloquially repetitious. Interestingly enough, her transcriptional principles violated the original instructions given to her by the American professor of German who had collected the data.

In this instance, the original purpose of the data collection – access to genuinely colloquial German for American university students – was actually violated and prevented by the well intentioned transcriber. In doing so, she made her transcripts excellent examples of what we have referred to as the written language bias (Linell, 1982, 2005). Her adherence to the principle that spoken discourse must be transferred to paper as well-formed sentences dramatically changed her principles of transcription from those she had been instructed to use. But she was unaware that she was complying with such a bias. Hence, her performance in this setting is evidence that the written language bias is not at all limited to the discipline of linguistics or the broader area of language and communication sciences; it affects literate people as such, most of whom have had little or no contact with linguistics, and it becomes criterial for their

judgments of the proper usage of language on paper (for a similar observation in anthropological fieldwork, see Urban, 1996, p. 27ff.).

The example serves also to bring us closer to the topic of this chapter, the empirical methods that characterize mainstream psycholinguistics on the one hand and a genuine psychology of language use on the other.

The Experimental Method

Psycholinguistics has its historical roots in both the experimental psychology (general psychology or in German *Allgemeine Psychologie*) of the nineteenth century and the behaviorism of the twentieth century. Experimental psychology and the behaviorist tradition both set a pattern of laboratory research in which careful control of extraneous variables and meticulous, systematic variation of relevant variables are essential. This tradition has led to a severe limitation of research to what could be thus subjected to laboratory conditions. In its most objectionably extreme form, it has led to the limitation of investigation to what can be most easily engaged rather than what is of the most relevance to human life. Ebbinghaus (1885/1992) has provided the prototypical example of this extreme: The very concept of rote learning as well as Ebbinghaus's use of nonsense trigrams to elicit such learning totally disregard, respectively, the nature of human learning and the meaning to be found in the most nonsensical of such materials. And despite Skinner's unsuccessful effort to return to the paradise of *Walden two* (1948), behaviorism's reduction of human learning to conditioning has also ended in a thoroughgoing irrelevance of his laboratory research to genuine language use. More recently, Foppa (1994, p. 147) has noted critically that one reason for the neglect of dialogue by psychologists "has certainly to do with the discipline's almost obsessive fixation on the experimental paradigm and with its belief that only by experimental methods can scientifically valid results be obtained."

Fellow Travelers of Psycholinguistics

What were added to these traditions by the psycholinguists of the mid-twentieth century included transformational grammar and a concomitant dependence on linguistics for experimental hypotheses and for underlying principles, along with an emphasis on information transfer and the computer instantiation of such as a model of human communication and the psychological processing of language. And all these problems contributed to a Zeitgeist that was very much associated with Noam Chomsky's and several other key researchers' charismatic leadership and visibility. Prominent among these researchers were George A. Miller and many of his colleagues and students at the Harvard Center for Cognitive Studies in Cambridge, Massachusetts.

An Example of Experimental Psycholinguistics

Allow us to analyze as exemplary of these various influences on methodology a well known experiment that has been cited in the literature many times since its publication, namely Levelt's (1981, 1989) linearization research, a conceptually simple experiment in which the subject describes aloud for audio recording a visually presented network of lines and colored nodes. As experimental laboratory research, it is clearly traditional. It is also in a tradition that Levelt (1989, p.1) himself has referred to as a cognitive/informational point of view that considers "the speaker as a highly complex information processor." Furthermore, the tradition is one in which the language system of the speaker, his or her "cognitive skill," rather than the speaker as such, is of interest: "a reasoned dissection of the system into subsystems, or processing components."

But methodologically, of even greater significance is the artificiality of the experiment itself. After the instructions, no further questions are entertained. Since the experimenter who is addressed by the experimental subject has already seen the visual materials being described, and since the "next subject" for whom the experimental subject is putatively articulating a description does not exist, there is no genuine recipient of the communication – it is all a fabrication that the subject is asked to accept and go along with. Nor is there any clear intention evident on the part of the speaker, any reason for him or her to speak, other than the request of the experimenter that he or she do so. Furthermore, without knowledge of the real purpose of the experiment, the experimental subject has no way of knowing how to address effectively the apocryphal "next subject." In other words, in a normal engagement of two people in spontaneous spoken discourse, the speaker would know *why* he or she is telling something to a listener and *how* the listener is expected or intended to use the description; and the listener in turn would give some kind of feedback, verbal or nonverbal. This gradual increment to the listener's knowledge should dictate radical differences in the speaker's description. The point may be further clarified by an example: If one is standing in Times Square, New York City, the question "How do I get to Central Park?" is clear enough; but if one is standing on the deck of a cruise ship in the North Sea, the same question would likely elicit some counter questions, perhaps along with laughter: "Starting from where?"; "By foot?"; "In this kind of weather?"

It should be carefully noted that all the experimental details described above are thought of in the experimental tradition precisely as proof that careful controls have been established. Levelt (1981, p. 309) has felt that it was sufficient to instruct the experimental subjects "to describe the figure in such a way as to enable the next subject to correctly draw it on the basis of the tape-recorded description." O'Connell (1992, p. 60 f.) has raised a number of specific objections to this experimental setting:

> An experimental subject now begins to speak. Do we know what his or her intention is in speaking? No, we do not. We know only the demand characteristics of the instructions. To make this clear, let us assume that, for every successful description on the part

> of a "next subject," $1,000 is to be shared evenly between the subject and the next subject. How would the subject best optimize this business enterprise? A prudent subject would clearly wish to know what "on the basis of the tape-recorded description" means. Does it mean that the next subject will have to execute a correct drawing: (1) *during* the playing of the tape-recording; (2) after only *one* playing of the tape-recording; (3) by some time *deadline*; (4) some combination of the above; or (5) with no limiting conditions? Of course, we all know realistically that there is no "next subject."
>
> The reason this little scenario seems so strange is not at all that the questions are absurd, but that the trivialization of speaking in such experiments is so common. If intention is really important, then we must be concerned about why a person speaks – to what end. Speaking because an experimenter has asked me to is indeed an intention, but not an exclusive nor an adequately determinative one. Intention is, in fact, simply neglected in the linearization experiment.

Finally, quite in keeping with the neglect of intentionality, the visual materials are patently trivial and meaningless, not even a puzzle that might be intellectually challenging to the experimental subject.

This is, of course, not the way human beings use language, except, as stated, under the most unusually artificial and demanding conditions. In addition, if indeed such a tape-recording were actually to be used by a "next subject" listener, the prosody of the spoken description would become of primary importance, simply as a transmitter of clarity and precision. But Levelt's failure to analyze prosodic variables indicates that not even he considered the experiment realistically relevant to the pragmatic application described in the instructions. All these elements of artificiality become very evident if one begins to imagine how the experiment would change if there really were a "next subject" whose successful redrawing of the visual materials would actually yield a substantial monetary reward or would have some other impact on the outcome of a conversation or on the completion of some practical enterprise beyond the conversation.

And yet, this research has been seriously mined for evidence on how speakers edit their speaking. Suffice it to say that the generalizability of such evidence cannot go beyond the artificial conditions of the experiment itself. A realistic situation would have made the description and its attendant editing serious elements of a genuinely communicative situation that was expected to make some difference in the life of the "next subject." Such an expectation would grant the whole procedure a truly legitimate intentionality. Unfortunately, one would then have to expect that the description and editing would both be quite different from the results reported and interpreted so generally for this experiment.

The Demand for Continuity in Speaking

It is our conviction that the ultimate rational for Levelt's experiment is his implicit theoretical demand for continuity. It is obvious that repairs and editing do indeed occur in spontaneous spoken discourse. A problem arises, however,

when they are made into an evidential base for the continuity required by the ideal-speaker theory. Taylor (1997, p. 61 f.) has rejected this inference regarding the need for continuity:

> Without the notion of discontinuity as an error in speech – a notion rhetorically derived from scriptist premises – there would be no special reason to search for the causes of discontinuities. Nor would any question arise concerning the speaker's and hearer's management of the communicational obstacles created by discontinuities. Indeed, without scriptism, the notion of discontinuity might well prove to be incoherent. It is, after all, only by comparison with the notion of continuity that it is possible to identify a set of phenomena as instances of discontinuity. Discontinuity, however, has been taken to include precisely those features of a speaker's performance which would not occur in the continuous spoken performance of a practice reader reading out loud. The practical identification of discontinuity, like its theoretical definition, is dependent upon its opposition to a fundamentally written language notion of continuity. Should these assumptions underlying the notion of continuity be withdrawn, it is not at all clear on what basis the important dialectical opposition between continuity and discontinuity could be justified.

Scriptism thus appears to be Taylor's rendition of Linell's (1982, 2005) written language bias. The fact that there can be no such thing as a pure case of continuity or fluency in human speaking makes Taylor's argument the more cogent. There is no reason to penalize what Heinrich von Kleist (1806/undated, p. 975; our translation) has referred to as "the gradual working out of one's thoughts in the process of speaking." This gradual process – with all its stops and starts – is the only way human beings can possibly engage in spontaneous spoken discourse. Chafe (1980b, p. 171) has another way of referring to discontinuity in speaking. He has quoted William James (1891/1981, p. 243): "Like a bird's life, it seems to be made of an alternation of flights and perchings." In passing, one might note, as the quotation clearly indicates, that James's famous continuous stream of consciousness applies only to thought, not to speech: Speech is of its very nature *intermittent*.

Methodology in Psycholinguistic Textbooks: The Relationship of Data and Theory

Another way of looking at the problems of method and methodology in mainstream psycholinguistics is to investigate how these two concepts are used in standard textbooks. In both the classic textbook (Clark & Clark, 1977) and in more recent textbooks (e.g., Carroll, 2004, 2007; [Cutler, 2005 has no subject index]; Field, 2003; Harley, 2001, and see 2008; Tartter, 1998), the concepts simply do not appear in the subject index. To judge from Harley's subtitle *From Data to Theory,* one starts with data and ends with a theory. How then are we to know what data to collect and what they mean if we have no proto-theory? In his chapter on language production, Harley (2001, p. 351) does use the terms method and methodology: "The methodology behind speech error analysis is a

simple one. The most common method is to collect a large corpus of errors by recording as many as possible." But, such a methodology is only reasonably applicable along with at least a preliminary definition of error, based on an at least implicit preliminary theory. Additionally, Harley's description does not actually fit with what is done empirically in the collection of speech errors. In fact, there are instead antecedent expectations based on a theory, and the theory is that of continuity in the ideal speaker's speech production. What Harley, and the textbook writers generally, do is to present the evidence collected in investigations such as these speech error ones and to incorporate them into the current theorizing. By no means does the research begin, as claimed, simply with the collection of a maximum number of speech errors.

By and large, the textbook writers have bought into the Zeitgeist. There is a certain implicit common-sense certainty about scriptism or the written language bias. The well-formed sentence is a joy to work with, and the controlled experiment is king. The consequent neglect of field-observational research – where genuinely intentional language use is to be found in abundance – has been a great loss to a psychology of language use. The virtual ban on qualitative analytic methods as an adjunct to quantitative methods and to inferential statistics has also been disappointing. In short, the methods available to mainstream psycholinguistics have been curtailed by an underlying bias toward the literate, written mode, and toward the traditional approaches of the laboratory. A number of recent handbooks have provided further details regarding the investigation of language use: Ball, Perkins, Müller, and Howard (2008), Graesser, Hernsbacher, and Goldman (2003), Schiffrin, Tannen, and Hamilton (2003), and Traxler and Gernsbacher (2006).

Monologism

Because of its overarching relevance for mainstream psycholinguistics, it is important to segregate *monologism* as a major determinant of method and methodology. Monologism is simply the concentration on monologue as the principal source of empirical material for research on language use (see O'Connell & Kowal, 2003), along with an implicit generalization of findings to dialogue. It has indeed been dominantly characteristic of mainstream psycholinguistics from its beginnings in the mid-twentieth century.

There are certainly monological components in human language use. Shakespeare's Hamlet and Joyce's Molly Bloom both give their soliloquies. But both are somehow surreptitiously engaging in dialogue: Their discourse is intended by Shakespeare and Joyce, respectively, for the audience. One can argue that language use is in principle dialogical, and that even the most private monologues have an element of dialogical otherness and distancing from self as a self-dialogue. In fact, such is precisely the rationale for the establishment in 2006 of the *International Journal for Dialogical Science*. In any event, the vast majority of

speaking, writing, listening, and reading transpires among people who are somehow present to one another – either physically or only mentally; all of these genres of language use are dialogical, not monological. Hence, it is no less than astounding that the perennial and overwhelming bias of mainstream psycholinguistics has been monologue. In his current German psycholinguistics textbook, Dietrich (2002, p. 140; our translation) has acknowledged this bias as regards research on the phase of conceptualization in speech production: "All the findings have as their source observations of monological spontaneous oral language production. They do not tell us anything about the dynamic of communicative activities."

To be sure, as with the written language bias, so too with monologism, the bias preceded the mid-twentieth century. Lazarus (1879/1986) was ignored in the late nineteenth century when he tried to introduce the study of conversation (*Über Gespräche*) into psychology. Instead, words (e.g., Cattell, 1886) and consonant–vowel–consonant trigrams (Ebbinghaus, 1885/1992) held sway in the laboratory setting. One can readily acknowledge, however, that "monologistic psycholinguistics has indeed engaged dialogue, but precisely monologistically, by simply concentrating on an individual language user as the entity of analysis" (O'Connell & Kowal, 2003, p. 195). Or as Foppa (1994, p. 148) has expressed it in his own critique of mainstream psycholinguistics: All social phenomena are explained therein "on the basis of one's knowledge of the processes in the participating individual members. In other words, there is no other 'entity' of analysis than the individual person."

A Monologistic Approach to Dialogue

Part of the problem for mainstream psycholinguistics is the fact that it is indeed "extremely difficult to have any experimental control over normal conversation and this makes it difficult to investigate dialogues in a rigorous way" (Garrod & Pickering, 1999, p. 10). But the rationale given by Garrod and Pickering for this difficulty is not convincing: Rigorous experimental control is not the only way to gather scientific data; this is simply a traditional bias of experimental psychology inherited from the last century. In addition, as we have already mentioned in Chapter 1, there is an underlying bias against dialogue on the part of Garrod and Pickering: "The language of dialogue is disorderly compared to the straightforward grammatical sentences of monologue" (p. 10). These authors seem to have confused disorderliness with complexity – a complexity based on organizing principles far different from and far beyond the grammatical organization evidenced in unrealistic monological sentences. And so, they have concluded with their own methodological desideratum: a controlled syntactic well-formedness in dialogue. As for methods genuinely applicable to conversation and dialogue in general, the fact of the matter is that "genuine dialogue is never a simulated experimental task" (O'Connell & Kowal, 2003, p. 200). The collocation of monologism and dialogism within mainstream psycholinguistics

is reminiscent of Rex Harrison's famous line in the film version of *My fair lady*: "Why can't a woman be like a man?" (Warner & Cukor, 1964/1986). Indeed, why can't a dialogue be like a monologue?

Turn-Taking

The reader may find this section somewhat more fine-grained than the other sections of this chapter. One need not seek far afield for the reason: Turn-taking comes close to being the core concept of dialogue; it is in turn-taking that the interactive process between interlocutors actually takes place and is made observable. Hence, it is of supreme importance that the monologistic well-formedness principle becomes an issue once again as we confront the analysis of turn-taking methodology. Two assumptions must be rejected: (1) that the criterion for the success of a conversation is to be found in "the smooth interchange of speaking turns" (Cutler & Pearson, 1986, p. 139) and (2) that, insofar as conversation is "organized around establishing consensus" (Garrod, 1999, p. 392; see also Clark & Brennan, 1991), "the fundamental goal of dialogue" (p. 393) is consensus. At a more operationalized level of turn-taking, Wilson and Wilson (2005, p. 966) have recently provided "a mechanistic account of how timing is coordinated between conversational partners." The sleeper here is simply that they have used data that, according to the original researchers themselves (Wilson & Zimmerman, 1986, p. 384), contained "substantial measurement error," included 25% between-speaker silences that were "not between-turn silences," and disregarded overlaps and both successful and unsuccessful interruptions on the grounds that they did not occur at "transition-relevance places" (p. 379). These conversations were also elicited dialogues, about which Taylor and Cameron (1987, p. 52) have commented: "There is no guarantee that data obtained in this way is representative of talk produced in non-experimental contexts." The fact that smooth transitions (with or without a pause) sometimes account for less than 50% of the turns in dialogue (see, e.g., Suleiman, O'Connell, & Kowal, 2002, p. 277) makes the turn data of Wilson and Wilson (2005) even more problematic. And the fact that the "*projectability*" (Wilson & Zimmerman, 1986, p. 379) claimed for such transition-relevance places has been challenged as an unreasonable and unrealistic psychological expectation (see O'Connell, Kowal, & Kaltenbacher, 1990) also goes unmentioned. And so, the population of turns in their experimental corpus is not validly represented and not legitimately analyzed in Wilson and Wilson's research.

Methodological Individualism

The monologism characteristic of mainstream psycholinguistics is essentially asocial, and such concentration on the individual is indeed methodological individualism (see Foppa, 1994, p. 148; Clark, 1985, p. 179). It is gratifying to

read in Pickering and Garrod (2005, p. 85): "The study of dialogue provides a radically different conception of psycholinguistics from the traditional study of language comprehension and language production in isolation." And yet, they still conceptualize dialogue in the traditional mode "as a largely automatic process of alignment between interlocutors." And once again, for Pickering and Garrod, this alignment of "situation models" (p. 87) constitutes the criterion for the successful execution of a dialogue. In reality, however, there is nothing at all automatic about whatever alignment occurs in the course of genuine conversations. In fact, we ourselves have found the assumption of an automatic process of alignment as the finality of dialogue to be an ever present temptation in our own research.

The Need for Normalization of Data

Another challenge to the appropriate analysis of data has to do with counts of various response measures. Biber, Conrad, and Reppen (1998, p. 263) have emphasized the importance of "normalization" for the sake of comparability of counts. In other words, the comparability of counts of response measures from one corpus to another depends on normalization whenever the corpora are of varying lengths (written or spoken, in syllables) or durations (spoken, in seconds). For example, Lakoff (2001) wanted to compare George W. Bush and Al Gore on the use of plural second-person pronominals in their speeches. She was unable to make any exact comparisons even within her own database because she failed to normalize her data. It should be noted that mainstream psycholinguistics does not generally violate this requirement of normalization of data; however, the same cannot be said for all the sciences that deal with language use. The danger arises particularly when corpora from field observational research are to be analyzed.

Access to Corpora of Spontaneous Spoken Discourse

Taylor and Cameron (1987, p. 15) have claimed: "It is a relatively straightforward task to collect conversational data, and that data lends itself to statistical analysis particularly well." Would that it were true! The Scylla and Charybdis of surreptitious audio recordings on the one hand and in-your-face microphones on the other are omnipresent. The former is unethical and the latter distorts the data. This is the basic rationale for our own turning to media discourse, since it is of the best acoustic quality, is in the public domain, is characteristically about nontrivial matters, and is spoken by articulate, intelligent, public figures. One must, however, be alert to the danger of over-generalizing media discourse to other, more informal types of spontaneous spoken discourse.

Recordings of media discourse still pose challenges of transcription: The more vigorous and spontaneous (and correlatively, the more interesting) they become, the more difficult they can become to transcribe. Very rapidly enunciated particles, spoken in sub-second durations, very softly, and overlapping another interlocutor's speech, can at the same time be of great importance. Examples of methodological difficulties with transcription of rapid spoken discourse can be found in Redder and Ehlich (1994; for a critical review, see Kucharczik, 1996) and in Schegloff (2007; see also our Chapter 16). Both publications provide the audio recordings used for their transcripts. Re-analyses of samples from these audio recordings have suggested that the published audio/transcript correspondence is deficient, partly because the original audio recordings are of poor quality. Additionally, the use of an interlocutor's entire body as a communicative tool lends nuances and sometimes even negates what is being articulated verbally. Hardly a straightforward task! Nonetheless, this is where the action is.

Use of Transcripts Prepared by Others

Finally, an important methodological problem can be pinpointed in Clark and Fox Tree's (2002) research on using *uh* and *um* in spontaneous speaking. They have made use of the London-Lund corpus (Svartvik & Quirk, 1980, p. 74) to hypothesize that *uh* and *um* are "conventional English words" (Clark & Fox Tree, 2002, p. 73) with basic meanings that signify a speaker's intention to initiate, respectively, a minor or major delay, i.e., to introduce a shorter or longer pause. The problem arises insofar as Clark and Fox Tree did not measure the physical duration of pauses after *uh* and *um* in the corpus, but relied on the professional coders in Svartvik and Quirk's analyses who perceptually identified duration units. Or as Clark and Fox Tree (p. 81) have put it: "So it is ultimately the *perception* of pause length and prolongation that we are studying here." O'Connell and Kowal (2000) and Spinos, O'Connell, and Kowal (2002) have physically measured the durations of pauses following both *uh* and *um* in samples of the London-Lund corpus and found no evidence for the correctness of Clark and Fox Tree's hypotheses. In addition, O'Connell and Kowal (2005b) have replicated Clark and Fox Tree's investigation of pause durations after *uh* and *um* with a corpus of media interviews by Hillary Clinton and found the same negative results. Because of a multitude of both false-positive and false-negative identifications, perceptually identified silent pauses do not accurately reflect the actual physical occurrence thereof (these issues are discussed in greater detail in our Chapter 13).

There is quite another way in which transcripts prepared by others are made use of in research and teaching. Transcripts from published research projects are frequently reproduced in textbooks and in further research publications for purposes of comparison. O'Connell and Kowal (2000) compared 41 original transcript excerpts with the derivative reproduced transcript excerpts and found an extraordinarily high rate of change in the reproduced excerpts. Kitzinger

(1998) has claimed that such an unusually high rate is the result of carelessness on the part of researchers. We are of the opinion that it is a consequence of the very high density of transcript notations, which puts undue pressure on the ones responsible for the reproduction of the transcript excerpts (e.g., printers and copy editors). What is perhaps even more surprising – and disconcerting – is the fact that many of the notations in the original or reproduced transcripts prove to be totally superfluous, insofar as only a few of the notations are actually used for analyses that are carried out on the transcripts. There is a very simple solution to this problem: It is for researchers to make use of only those notations that are intended to enter into their analyses of the transcripts. Transcribing everything from a spoken passage is impossible; transcribing everything one can find to transcribe is not at all scientific, but smacks instead of an effort to appear scientific with an abundance of impressive notations (for a more detailed discussion see our Chapter 10).

Back to Issues of Control

We began this chapter with an epigraph in which Dietrich (2002, p. 14; our translation) has described modern mainstream psycholinguistics as characterized by *experimental controls*. We wish to call attention here to a much broader sense in which the language use that we ourselves have subsumed under the category of spontaneous spoken discourse is controlled – not experimentally, but *situationally*. Our empirical work in recent decades has concentrated upon media discourse and, in particular, political interviews. It should be noted that this genre of human discourse constitutes only a tiny corner of the spontaneous spoken discourse that itself constitutes the vast majority of human spoken discourse. Furthermore, political media discourse, like all media discourse, has a number of built-in external *controls* that limit its scope and style. Basically, it is a setting that *demands* speaking; one is not free to arrange an interview and then sit there while others do all the speaking and the manner of speaking is also limited: Shouting or whispering into a live microphone can be quite counterproductive; good conduct – e.g., the avoidance of vulgarity and obscenity if one is so inclined – is required; time limits are stringently constraining, even to the point where interviewers must interrupt interviewees in order to stay within time limitations; no allowance is made for extraneous forms of behavior such as catching a snack or greeting a passer-by in the studio; the agreement is to a dialogical encounter, but to a very peculiar one, in which the interlocutors speak with one another for the sake of a third party, a media audience, and the format is generally question and answer. All of this is in sharp contrast with what we have already discussed above under the heading of Bühler's (1990, p. 179) "*empractical* use of language signs" and what a recent German-language dissertation has referred to as "Knappes Sprechen" (Baldauf, 2002, p. 1) – incidental, intermittent speaking: It is highly situational, elliptical, deictic, concrete, and often quite idiosyncratic.

A typical example can be heard in the commentaries and reactions of someone watching TV, e.g., *wow, huh, good for her, yeah, I thought so, high time*.

The danger for ourselves as researchers is to forget that we are occupying this tiny corner, whereas the vast amount of spontaneous spoken discourse is quite other than political media interviews. At the same time, it would be a serious mistake to think that any genre of spontaneous spoken discourse exists that is *without* controlling factors. There is always some concrete situation in which an utterance must be embedded, and the specifications of that situation are always such as to control, in one way or another, what is essentially spontaneous spoken discourse.

To return for a moment to the concept of *spontaneity* in this context, one could well argue that the controls listed above exclude any sense in which genuine spontaneity is exhibited by political interviews. The point is well taken, but the stark contrast between reading aloud or reproductive speaking on the one hand and productive speaking on the other must be maintained: Political interviews clearly fall into the latter category (see Kowal, 1991). Furthermore, as we will see in our empirical chapters (11–18), the relative differences between the literacy of interviewers (who do indeed typically make use of written notes) and the orality of interviewees is quite notable. In any event, spontaneity must be conceptualized as a broad-ranging continuum rather than as a narrow-band distribution of qualities.

Concluding Remarks

At mid-twentieth century, the new mentalists were convinced that the sterility of behaviorism was about to be overcome, and that a whole new world of psychological breakthroughs was at hand. Costall's (1991, p. 163; cited in Linell, 1998, p. 58, footnote 8) comment on such triumphalism is telling:

> Cognitive psychologists have perhaps been too busy congratulating themselves on not being behaviourists to notice that they themselves treat people as machines (Skinner 1974: 110; Morris 1991). The mechanistic scheme, and computer metaphors in particular, lead us to regard the problem of cognition as nothing other than the internalized re-presentation of the environment.

In the same footnote, Linell (1998, p. 58) has referred back to his own earlier comments, in which "Linell (1979) points out that Chomskyan mentalism is not incompatible with a behaviorism that builds upon internal mediating variables." Such is the psycholinguistic methodological legacy out of which a viable psychology of language use must be formulated anew, i.e., with an emphasis on a *genuinely psychological* study of language use. We are well aware that the present volume can only call attention to some of the directions needed to construct a psychology of language use that will prove to be a comprehensive investigation of oral communication with one another – spontaneous spoken discourse. The edifice is yet to be constructed.

Chapter 3
Fluency and Hesitation

Uh Walter that uh I know you're running out of time and here I am hemming and hawing... (Ronald Reagan in an interview with Walter Cronkite, March 3, 1981; cited in Kowal, 1991, p. 147).

In all spheres of life, hasty engagement and unreflective enactment are evidence of brutalization. Hesitation alone is human (Werner, 2006, p. 92).

Chapter Prospectus

Chapter 3, *Fluency and Hesitation*, examines another theoretical bias of mainstream psycholinguistics – the theory of the ideal speaker. This theory is a corollary of the written language bias and at the same time is basic for the understanding of the following empirical chapters. In accord with the theory of the ideal speaker, written language is considered well formed and the reading aloud thereof is considered fluent insofar as the reading is in accord with the sentential syntax. By contrast, spontaneous spoken language is considered to be both deficient and disfluent. But in the sense of an absolute continuity, perfect fluency of sequential ordering is impossible by reason of the need to breathe. The concept of fluency itself has evaded any realistic, useful definition, even though it is basic to the theory of the ideal speaker and to the concept of syntactic well-formedness. Correspondingly, a flawed concept of *disfluency* underlies much of the research on hesitation.

Chicken or Egg?

We know quite well that, both ontogenetically and historically, spoken discourse precedes written discourse. Hence, in both the language-learning child and in the course of cultural development, speaking and listening precede writing and reading. And yet, somehow, the seductive orderliness of the well-formed, written lettering and the elegant functions thereof grant both an authority and dignity to the written mode. True enough, the transmission of

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_3,

culture has become dependent more on writing than on oral tradition, and the definitive rule of law, recourse to history, liturgical, dramatic, and musical forms, and even the proverbial shopping list are trophies of written language. Nonetheless, it remains true that a huge proportion of our everyday engagements with one another is transacted via the oral mode. We spend our time speaking with one another from breakfast all the way through the day to the last "good night." This all-day engagement is a universal human phenomenon that extends equally to analphabetic and illiterate as well as to literate speakers and listeners. It should be added that the communicative intelligibility of such discourse is not an issue; that is to say that this oral activity does not generally constitute a problem of everyday life.

Some comparisons are in order, but are not as simple as might at first seem to be the case. When it comes to fluency and hesitation, the comparison to be made always involves the use of time. But this use of time is radically different in writing and reading and in speaking and listening. In writing, time is not at all of the essence for the act of communication itself. One may take two years to write a poem – or two minutes. The poem will be published in definitive form for others to read, and the use of time on their part is entirely independent of the time one spent in composing the poem in the first place. However, it should be emphasized that there is a vast difference between writing, as a temporal *process* of language use on the part of the writer, and the written, as *product* of that process: The process takes place in time, the product exists in space. All too often, when one speaks of writing, the written product is meant (see also our Chapter 4). One could, of course, publish a day-by-day or hour-by-hour time line of the actual writing process. It would resemble successive page proofs for the stages of composition. But such a document would still not provide a legitimate comparison with the oral production of speech. For example one could write as follows:

Example 3.1
The boy stood on (5 s) (5 h of reflection without writing) the burning deck (5 s).

This example might well be intended to reflect the fact that the writer took, respectively, 5 s to write, 5 h to reflect, and again 5 s to write the sentence. But such a convention is entirely forced and unnatural. Readers deal with the *product* – and they can do so at their own pace. The temporal characteristics of the composition *process* are nearly always of no interest to the reader, and a record of the time line of composition or writing is available to the reader only in very rare cases. The closest example we could find in our own data is Hannah Arendt's description of her process of writing as told in her famous German TV interview with Günter Gaus. The following quotation is taken from the published version of the interview (Arendt, 1996, p. 47; our translation):

Example 3.2
GAUS: Do you write with ease? Do you formulate easily?
ARENDT: Sometimes yes, sometimes no. But in general I can say that I never write until I copy so to speak.
GAUS: After you have already thought it through.

> ARENDT: Yes. I know exactly what I want to write. Until then, I don't write. Most of the time I write only one version. And then that takes place relatively quickly, since it is actually dependent only on how fast I type.

It is as if Arendt is here formulating a psychological theory of the temporal organization of her own writing process. But the process referred to by Arendt yields a product in which the actual temporal course of writing is lost to the reader. In passing, it should be noted that Arendt's introspective memory about the tempo of her writing may not be correct; that is, it may not be in accord with the actual temporal organization of her writing. She too is subject to the written language bias as expressed in her own statement: "I copy so to speak." Copying is from a written model – here in Arendt's own mind.

In speaking and listening, however, we are dealing with a real-time transaction across interlocutors – a time span open to their immediate observation and constituting an essential element of their interaction – whereas the 5 h of reflection in the middle of writing an otherwise 10-s long sentence would effectively remove any potential readers from the presence of the writer. Quite obviously, then, how time is used in speaking and listening becomes important for the speaker, the listener, and the researcher. And, in fact, time becomes the fundamental carrier or medium of spontaneous spoken discourse and of reproductive speaking (reading aloud) as well. The need for and the use of time become evident not only in pause and rate, but also in the way successive words are produced (e.g., syntactically well-formed or not), and in the way interlocutors use time successively or simultaneously in their respective turns. In fact, as a general principle, all the spoken phenomena presented in the following empirical Chapters (9–18) must be analyzed in terms of real time as an important component of their use. To provide the reader with a dramatic example of how time becomes important in the production of dialogical spoken discourse, we reproduce here our unpublished transcript of Hannah Arendt's and Günter Gaus's interview excerpt *as spoken* (our translation from the German) that was given above in its published version (*Example 3.2*); pause duration is given in seconds within parentheses:

Example 3.3

Gaus:	yes (0.40) yes once more regarding writing (0.43) do you write with ease (0.13) do you formulate easily
	(1.68)
Arendt:	sometimes yes sometimes no but in and of itself I never write (0.41) until I not (0.14) so to speak (0.55) copy
	(1.01)
Gaus:	mm-hm
Arendt:	uh (0.10) that is
Gaus:	copy from what you've already thought through
Arendt:	I know yes I know exactly what (0.24) and until then I don't write I uh therefore uh uh (0.32) most of the time I write only one version (1.06) and then that takes place relatively quickly since it is actually dependent only on how fast I type

A comparison of *Examples 3.2* and *3.3* yields a number of interesting differences. The published version, intended for readers, not for researchers, is in well-formed sentences; there is no such thing as overlap of turns or prolongation of any kind, whereas the unedited transcript, intended only for researchers, indicates both overlap and a variety of prolongations by means of pauses. This transcript also contains all the hesitations and revisions that were spoken and were thus "laid bare before the listener" (Chafe, 1986, p. 13), but were obviously considered inappropriate for a reader of the published version. In terms of the ideal-speaker theory, one should keep in mind that Gaus was among the foremost TV interviewers in Germany throughout the post World War II era; his interview with Arendt won for him one of his three Grimme Prizes. His own assessment (Gaus, 2004, p. 201f.; our translation) of his interview with Arendt was as follows:

> In the more than two hundred interviews that I've conducted in four decades, the one with Hannah Arendt, televised on October 28, 1964 on *ZDF* [the Second German TV Channel], is the one that has remained for myself the most deeply moving, impressive, and expressively powerful.

For her part, Arendt's assessment of the interview was articulated in a letter to Karl Jaspers (Arendt, 1992, p. 568; October 25, 1964; our translation), shortly after the recording session: "I had the feeling that I spoke too spontaneously, because I like Gaus a lot." This scenario comes very close to expressing the ideal of *successful hesitant* spontaneous spoken discourse, presented in Part IV of this book.

Ideal Delivery vs. Discontinuity

Mainstream psycholinguistics tells us in no uncertain terms how to manage the available time we have for speaking: The ideal speaker must avoid off-time except as dictated by the syntax of a well-formed sentence, that is to say, he or she must speak continuously. Hence, discontinuity "falls squarely within the domain of the now famous Chomskyan notion of the 'performance error'" (Taylor, 1997, p. 45). Some of the performance errors that contribute to this discontinuity would include "numerous false starts, deviations from rules, changes of plan in mid-course" (Chomsky, 1965, p. 4), and many others.

Strictly speaking, an absolute continuity of speech is impossible: Pauses for breathing are not optional; normally we breathe once every 5 s, or 12 times per minute. In the resting state, inhalation and exhalation require roughly equal amounts of time; but even during speaking, a minimum of about 15% of the breathing time is devoted to inhalation and so constitutes down time or pause time for the speaker (see Denes & Pinson, 1963, p. 42). Hence, the demands of mainstream psycholinguistics are not really for a literal and absolute continuity, but rather for a rule-governed, orderly allocation of the time available for speaking; and this orderly allocation of time constitutes "a 'correct' way of

executing [articulating] a sentence, and this will be called the *ideal delivery*" (Clark & Clark, 1977, p. 261). Operationally, this means that "all breaks will be at grammatical junctures between sentences or major clauses" (p. 262).

Taylor (1997) has insisted that there is no evidence whatsoever "that speakers constantly *strive* for an ideal delivery" (p. 50) and also that there is no proof for "the claim that deviations from the ideal delivery constitute an obstacle to effective communication" (p. 51). The underlying rationale of Clark and Clark's insistence on the ideal delivery is, according to Taylor, the following:

> The articulatory execution of a planned utterance is held to be the reading off of a mental text. Scriptism tends to push the difference between spoken language and written language farther and farther out to the periphery of the communicative act. The essential features of speech and writing are assumed to be the same. (p. 52)

This position is exactly the opposite of Saussure's: "Speech was the ideal form of language and ... writing consisted of a (frequently infelicitous) attempt to imitate speech" (p. 54). For Taylor, then:

> The evidence reveals (i) that discontinuities can be intentional features of a speaker's performance, (ii) that they are only disruptive from the blinkered perspective of the scriptist conception of the communicative act, and (iii) that they may serve interactional aims in speech which, due to scriptism, have too often been ignored. (p. 62)

The Use of Time for the Sake of the Speaker

Over 200 years ago, a short essay was written in German: *On the gradual working out of one's thoughts in the process of speaking* (von Kleist, c. 1806/ undated, p. 975; our translation):

> If you want to understand something and can't figure it out by pondering, I would advise you, my dear ingenious friend, to speak of it to the next acquaintance who happens by. He certainly doesn't have to be a bright fellow; that's hardly what I have in mind. You're not supposed to ask him about the matter. No, quite the contrary; you are first of all to tell him about it yourself.

The key to this transactional act of discovery and clarification is presented in the title of the essay. The process is precisely a gradual working out of one's thoughts in the very transaction of speaking to someone else – *in real time*. Nor is the process to be a request for information; it is a creative discovery for oneself of meaning and intelligibility through the very act of formulating and articulating a communicative utterance in real time and – importantly – *in the presence of an attentive listener*. All the elements are necessary: the presence of another person who understands one's language, the effort to gradually hone the thoughts into a meaningful communication through the very act of articulation, and the gradualism of the entire process in real time.

Such a process – the very creation of meaningful discourse – is an important characteristic of all spontaneous spoken discourse as we have conceptualized it in this book. The creation of meaning is a critical component of the impetus to

dialogue with one another: We speak in order to discover what we know and what we mean to say. There is no getting around it: The influence of the listener in this setting is rather mysterious, but nonetheless very real. He or she is needed for the transaction to take place at all.

Such gradualism could well be faulted as a serious violation of the ideal delivery, since indeed speakers must "know what they want to say and say it fluently" (Clark & Clark, 1977, p. 261). The paradox has been stated quite clearly by Chafe (1980b, p. 170; see also Chafe, 1980a): "The fundamental reason for hesitating is that speech production is an act of creation." Were it really only "the reading off of a mental text" (Taylor, 1997, p. 52), speech production could well obey the mandate of an ideal delivery, but at the same time it would necessarily be a speech production limited to those who can indeed read off – the literate or the alphabetized. Again, we are dealing here with a scriptist metaphor. Consequently, it must be quite clear that speech production is in fact far more and far different from a reading off. We leave it to our readers to discern whether the "reading off of a mental text" in an ideal delivery is preferable – in terms of the use of time – to the creativity of a "*gradual working-out of one's thoughts in the process of speaking*."

The Use of Time for the Sake of the Listener

Mainstream psycholinguistics conceptualizes the use of time only in terms of speakers, not listeners. In other words, the ideal delivery requires only that speakers know "what they want to say and say it fluently"; this is exactly what Clark himself (1985, p. 179) has later referred to as "the *individualist* view of language use." But an efficient use of time for communicative purposes on the part of a speaker must also embrace the needs of listeners, despite "modern linguistic theory's neglect of the role of the hearer in verbal communication" (Taylor, 1997, p. 82). And beyond this bias of scriptism, there is much evidence that discontinuity can and does serve a positive function for listeners. Hesitations such as repeats and pauses reduce the density of an utterance over time, and thus may serve to increase the communicative efficiency of the speaker rather than disrupt it. Taylor has commented insightfully on the paradox of discontinuity:

> Ordinary discontinuous conversational speech seems to work perfectly well for the communicational tasks of everyday conversational interaction. Discontinuous speech appears to raise no obstacle to the success of communication. So the proponent of the theory of the ideal delivery is asking us to believe not only that speakers constantly try to do what they almost never do but also that, if they fail to succeed, it makes no manifest difference to the success of the speech act. (p. 51)

There is a certain absurdity in all this insistence on an unrealistic continuity. For ordinary speakers of English, "the topic of discontinuity is not a feature of their ordinary metalinguistics" (p. 35). That is to say that they do not generally make such discontinuity an issue or topic of commentary or discussion or even a

matter of deliberate awareness. In fact, such unawareness may contribute to the embarrassment that speakers experience when confronted with a transcript of their own spontaneous speech. By the same logic, our chronic unawareness of discontinuity may indeed contribute to the difficulty of transcribing spontaneous spoken discourse accurately. An exception can be noted in certain cases of truly pathological discontinuity; for example, we have observed one public speaker over several decades and have found that he articulates a filler such as *uh* or *um* 27 times per minute across the board. Such usage is strikingly salient to listeners. The opposite extreme, of course, is a conversation between two speakers who articulate their discourse in perfectly well-formed, unhesitating sentences. Such discourse may prove to be oppressively boring and totally lacking in spontaneity. As such, it fails in the very requisites for dialogue.

Fluency

Is it possible, then, to accomplish fluency in spontaneous spoken discourse? The present book is actually a chronicle of empirical work that has as its proximate goal the demonstration of fluency – and indeed eloquence – in spontaneous spoken discourse. Part III of the book, Empirical Research on Spontaneous Spoken Discourse, consists of 10 chapters or roughly 50% of the book dedicated to this purpose. As we shall endeavor to show the reader, the very tools of a genuine fluency are the actual hesitations, pauses, fillers, repeats, false starts, interruptions, simultaneous speech, and overlaps that mainstream psycholinguistics categorizes as disfluency, discontinuity, disruption, and inefficiency. It should be noted that many of our empirical corpora are from expert and experienced public speakers. Were all these hesitation devices in fact detrimental to rhetorical efficiency, they would surely be the first to carefully avoid them. Our approach to fluency is clearly at odds with a quite recent negative association of the notion of fluency with speech pathology in Uppstad and Solheim (2007, p. 79): "The notion of 'fluency' is most often associated with spoken-language phenomena such as stuttering."

Even Herbert H. Clark, one of the original creators of the ideal delivery (Clark & Clark, 1977, p. 261), has meanwhile come to the conclusion that the same phenomena he had once categorized as "unsystematic noise" and "performance errors" must now be thought of as "systematic and essential to the successful use of language" (Clark, 1996, p. 389). These phenomena reflect, in fact, the effort on the part of the speaker to maintain or re-establish the coherence and fluency of spontaneous spoken discourse in real time. And all this reflects in turn the intention of the speaker to flexibly tailor his or her speech to the communicative needs of the listener.

In summary, whatever is contributed by the speaker or the listener(s) to the genuine communicative progress of an ongoing dialogue must be thought of as a contribution to fluency.

Chapter 4
The Written

Our conception of language is deeply influenced by a long tradition of analyzing only written language (Linell, 1982, p. 1).

Chapter Prospectus

Chapter 4, *The Written*, engages once again *The written language bias in linguistics* (Linell, 1982, 2005) on the part of mainstream psycholinguistics – this time with respect to the written itself. There are two problem areas: (1) the use of written materials as if they were the act of writing itself (the substitution of a product for a process); and (2) the assumption that written linguistic materials themselves can serve as the empirical evidence for a scientific explanation of language use as such. In fact, the written materials used in much of mainstream psycholinguistics often do not engage language use at all; written materials frequently manifest little or nothing of the settings and conditions under which they were produced, and they themselves remain inert and merely potentially meaningful until touched by the human psyche.

The chapter is also transitional in that it discusses additionally a very different sort of the written – written transcripts prepared for purposes of research on spoken discourse. Insofar as such transcription generally makes the transcript into the proximate database for a spoken corpus, systems of notation for transcription pose a problem – but this time more frequently for conversation-analytic research rather than for mainstream psycholinguistics. As we will discuss in more depth in Chapter 10, none of the notation systems in current use meet a reasonable set of criteria for reliability, validity, and user friendliness.

Verba Volant, Scripta Manent

People are no longer impressed by this venerable Latin adage: The spoken flies away, the written abides. From time immemorial, and up until the late nineteenth century, there was literally no access whatsoever to the acoustic past; spoken discourse was absolutely ephemeral, leaving not a trace of its history to

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_4, © Springer Science + Business Media, LLC 2008

posterity. This means that the human family, including all the language scientists, has no archival record of anything spoken before the beginning of the twentieth century. Acoustic records of speaking on the part of Bismarck, Christ, Demosthenes, Patrick Henry, and Lincoln, as well as of the nursery prattle and everyday kitchen chatter from innumerable families, are nonexistent. And even though we are in danger of forgetting these facts of history, in the long run, the invention of audio recording will in all likelihood prove to be far more important for the psychology of language use than the advent of either generative grammar or of the cognitive revolution in mid-twentieth century.

The historic hegemony of the written has been due to its ability to archive important transactions of the human family. Reliance on oral tradition for epic history and family tradition alike proved to be unwise; such record keeping from mouth to mouth was found to be far too selective and personalized and incapable of coping with the burgeoning volume of detail. Hence, the written mode took over the task of archiving the most important transactions of the human family for posterity. But even the written deteriorates. Stone surfaces wash away, parchment crumbles, paper burns, and printing plates can be smashed. Permanent preservation always remains problematic (see Posner, 1990).

The electronic age has finally brought us the possibility of preserving the voice characteristics, the acoustic history of any spoken corpus to which one has immediate access, that is to say, during the original enunciation of which a microphone can be activated to record the spoken discourse. However, this technological advantage can at the same time be a problem for the researcher: The temptation is to assume that the researcher's access to video recordings, DVDs, and transcripts for purposes of analysis puts him or her in a position analogous to that of the interlocutors' experience in the original setting.

Even the electronic age has not solved in any definitive sense the problem of preservation of the auditory signals of spoken discourse: Electronic record keeping also involves eventual deterioration. And even prior to the electronic age, the motion-picture industry had found that over time celluloid deteriorates and is dangerously flammable. Deterioration happens to be a part of the human condition to which there is no absolute answer, despite all the hubris and the combined efforts of the technical world.

Generation and the Written

The concept of a *generative* grammar clearly misled some psycholinguists into thinking that it had something to do with the *production* of sentences in writing. But in fact, generative grammar has nothing to do with language production or use. The generative grammarians have analyzed sentences *in abstracto* and *in vacuo*. To put it simply, the way words are aligned in a sentence was supposed to tell language users and psycholinguistic researchers alike something about the way people process language. But as Hörmann (1986, p. 69) has pointed out regarding Chomsky's view:

> His concept of generating was not meant psychologically; he does not maintain that his model describes the temporal relationships of the conception or the understanding of a sentence. The clearer this becomes for the psychologists, the more the importance of Chomsky's model for psycholinguistics is weakened – a process which has of course taken years.

With a similar logic, Rommetveit (1979) has voiced objections to the more specific notion of deep structure. In this connection, the entire discussion of disambiguation of sentences such as "The police were ordered to stop drinking after midnight." has put the cart before the horse by treating an isolated sentence printed on paper somehow as if it were an antecedent instance of oral language use. It is instead the isolation itself that ambiguates the sentence, not the other way around. It is literally impossible to *use* the above sentence in this isolated way, except precisely as a joke! If a speaker did use such a sentence, it would have to be about some definite police ("The police") about whom there was reason to say something at that precise moment in time. That is to say that in real time there is always a context, a setting, a reason for speaking or writing, and an intelligent audience that shares that ambience. Under these circumstances, the likelihood of a listener's even recognizing a potential ambiguity is minimal; and the likelihood of his or her being led astray into actual ambiguity is nil.

It was to Hörmann's credit that he saw through all this abstraction regarding the sentence as such and the autonomy of language that allowed the written demo sentence to become psychological data. Innis (1986, p. 4 f.; see also Derwing, 1979, p. 165 f.) has credited Hörmann with recognizing the following fallacy:

> Not only does language as an independent system of signifiers exist independently of the language user but also the study of its formal structure in itself supplies us with the guiding principles, chief questions, and heuristic clues for a specifically psychological study of language.

Mainstream psycholinguists of the twentieth and twenty-first centuries might well claim that this state of affairs no longer exists. It decidedly does. O'Connell (1988, p. 35) has summarized the situation toward the end of the twentieth century as follows:

> Despite all the emphasis on *speaking*, the sentences that have found their way into the analyses of psycholinguists have been for the most part *written* ones, isolated ones, and demonstrational ones at that (i.e., sentences that ordinary people would *not* be heard speaking).

More recently, Linell (2005, p. 30) has described the situation similarly in terms of "a *paradox* in modern linguistics":

> One claims the absolute primacy of spoken language, yet one goes on building theories and methods on ideas and experiences of a regimented, partly made-up language designed for literate purposes and overlaid with norms proposed by language cultivators, standardisers and pedagogues. All this amounts to a deeply ingrained contradiction based on a veritable reversal of priorities.

In case the reader has any doubt that mainstream psycholinguistics in the twenty-first century is still basically about the written, an examination of the most recent textbooks is to be recommended (e.g., Carroll, 2004, 2007; Dietrich, 2002; Field, 2003; Harley, 2001, 2008; Langenmayr, 1997; Rickheit, Sichelschmidt, & Strohner, 2002). Harley (2001, p. 311 f.) has stated it as follows: "**Discourse** is the spoken equivalent of text. . . . most of the research has been carried out on text comprehension rather than discourse comprehension." Rickheit et al. (2002, pp. 153–158; our translation), even in a chapter on the application of psycholinguistics to "optimizing human communication," deal almost exclusively with written texts. It seems clear that, for mainstream psycholinguistics, the autonomy of language structures is closely related to the isolation of those same language structures in written materials. The written sentence reveals structures immanent to language itself, but it can never suffice as a database to explain how human beings mean and understand when communicating orally with one another.

The Written as the Cadaver of Speech

We have already introduced the metaphor of the written as the cadaver of speech in Chapter 3, and here it will also serve us well. The written does indeed remain merely potential and inert until brought to life by speaker or reader. But this is not to deny the relevance of written materials as contrast to the spoken, whether the written was originally produced as written text, or was produced as a transcript of spontaneous spoken discourse for research purposes. In texts originally written to be read, we find words, phrases, clauses, and sentences demarcated by both capitalization and punctuation. They are essentially a static spatial array with no relationship to real time and only vestigial clues as to how the passage might be uttered in terms of prosody (variations of loudness, temporal organization, and intonation contour), nonverbal behavior, setting, ambient audience, or pragmatic involvement. In the case of the written, the heavy burden of making up for the absence of all the richness surrounding the spoken falls largely upon the verbal elements alone. The novelist must paint a scene in words and their syntactic structure alone, unaided by scent and sight, movement and sound – in short, without the surround of a perceptual reality. In a transcript, however, the goal is precisely to incorporate as much of the richness of the spoken as possible. And this proves to be a formidable task, as we shall detail later in this chapter and in Chapter 10.

But to return to the written as such, an example that throws light on its cadaverous nature can be found in Funkhouser (1979a, 1979b). Literary critics of modern poetry have traditionally based their reviews solely on the printed versions of poems. To examine this systematic limitation on the part of literary critics, Funkhouser had recourse to audio recordings made of poets themselves. One of her analyses (1979b) is particularly interesting. Although Randall

Jarrell's (1969) *The death of the ball turret gunner* was frequently anthologized and commented on by literary critics after World War II, none of them ever singled out the third line of the poem for special consideration or analysis, and none of them commented on it as in any way of special thematic importance. The third line reads as follows:

Example 4.1
Six miles from earth, loosed from its dream of life,

In Jarrell's own reading of his poem aloud, the word "loosed" was isolated by the longest and second longest pauses in the entire reading, was emphasized as the longest in duration and the loudest syllable in the entire reading, and was deliberately mispronounced by the poet himself so as to include two meanings: the freeing of the ball turret gunner from earth (the correct pronunciation of "loosed," with an unvoiced *s*) and the separation of the ball turret gunner from earth (the poet's actual pronunciation as "loozed," with a voiced *s*). Compared to such richness, the written poem is indeed a cadaver. An analysis of the written poem alone could never have unearthed such meanings – as the multitudinous commentaries of the literary critics eloquently witness – because they are essentially dependent upon the oral realization of the poem.

Transcripts, on the other hand, begin with the spoken and transform it into the written; they are cadaverous in a way very different from written text. One must, then, ask the following: How much of what is portrayed in a given instance of spontaneous spoken discourse can be and should be incorporated into the written as transcripts for research? How can transcripts serve as tools for the understanding of spontaneous spoken discourse?

Transcription

Why is transcription a very important issue in the theory and research regarding spontaneous spoken discourse? The answer is that its ephemeral nature demands that some sort of permanent record be derived, with the help of which analyses may be undertaken.

A transcript is a written version of a spoken corpus (see *Merriam-Webster's collegiate dictionary*, 11th ed., 2003, p. 1327). Interestingly enough, the *APA dictionary of psychology* (2006) has no entry whatsoever for *transcript*, despite the fact that a vast amount of modern psychological research – not only in the psychology of language use, but also in many other areas of psychology – depends on the use of transcripts. It is quite obvious that many researchers have assumed that methods of preserving spoken discourse that have been in use for millennia should be taken for granted as reliable procedures. But since, for purposes of analysis, the transcript necessarily becomes the proximate database of the original spoken data, it is important that transcripts be as close to the original as possible. Hence, for almost two decades now, we have been examining various notation systems for the transcription of spoken

discourse (e.g., Kowal & O'Connell, 1995, 2000, 2003a, 2003b; Lindsay & O'Connell, 1995; O'Connell & Kowal, 1990a, 1990b, 1994, 1995a, 1995b, 1998, 1999; Spinos, O'Connell, & Kowal, 2002; for further details, see our Chapter 14).

Over the years, a number of variant notation systems have been used by researchers. With only minor variations, the notation system for transcription devised by Jefferson (1989; see also Atkinson & Heritage, 1984; Schenkein, 1978) has become the most popularly used, especially by conversation-analytic researchers. Herrmann and Grabowski (1994, p. 32 ff.) have observed that frequently notated categories have reached a certain level of standardization across various notation systems. If this were the case, it would make comparisons from one research project to another quite feasible. However, our analyses of these categories have shown a considerable amount of variation from system to system. Across eight such systems most of which are still in current use for English and German corpora of spoken discourse, there has been complete agreement in only 30% (46/152) of the categories examined (Kowal & O'Connell, 2003a, p. 102). This level of agreement hardly reflects comparability across transcripts prepared according to the conventions of these eight current notation systems. In terms of research applications, data thus derived must be considered mutually incompatible and hence not comparable with one another.

Several decades ago, Deese (1984, p. 21) complained about the general quality of transcripts prepared for research and other important purposes:

> I have examined a number of transcriptions, some made for research purposes, others as official or legal records. I have found them all to be inaccurate to varying degrees. Those serving as official records for legal or institutional purposes are often grossly inaccurate. Those made for research purposes are often difficult to interpret and frequently ambiguous.

Our own experience has been similar to that of Deese. As our publications have recorded, even the *Congressional Record*, the *Washington Post*, and the *New York Times* make systematic errors in transcription of sources such as political speeches, comments, and interviews (see also Walker, 1986). In fact, we have found that not one of the notation systems in current use for the transcription of spoken discourse satisfies all the following criteria required for valid, reliable, user friendly transcription. One should note here that problems arise primarily with *published* transcripts; in these instances, the reader generally has no access to the original audio recordings and must rely exclusively on the user-friendliness and accuracy of the transcripts themselves. But all the systems in current use have one or more than one inbuilt flaws that may yield systematic errors in transcribing (O'Connell & Kowal, 1995a, p. 98 ff.):

1 *Only those components of spoken discourse which are to be analyzed should be transcribed, and only what makes analyses intelligible should be presented in transcripts for the reader.*

Instead one frequently finds densely notated, unreadable transcripts, not all the details of which enter into the accompanying analysis. The problem seems to be researchers' desire to transcribe *everything* in a spoken corpus – a goal that is in principle impossible. What is to be analyzed depends on the specific finality of a given research project. The claim that one standard notation system should be mandatory for all transcription of spoken discourse is neither scientifically sound nor practicable. For example, a transcript made for research on the numerousness of syllables in several corpora need not be comparable to a transcript prepared for research on the use of pronouns across several corpora. There is absolutely no necessity to attempt to transcribe as much as possible.

2 *Graphemes should be used only for the segmental representation of lexical items, and punctuation marks should be used only for their conventional purposes.*

A brief transcript from Jefferson (1984, p. 349; cited in O'Connell & Kowal, 1995a, p. 98) will suffice to exemplify a problematic usage of both graphemes and punctuation marks:

Example 4.2

→ M: eh Not the floo:r one ehh:: h euh he h-heh-he h
→ G: [eh h [h e : h [he:h

Our suggestion may sound preposterous, but do try to read it aloud, using all the notations to guide your enunciation of the passage. To say that it is not a user friendly passage is an understatement. We do know from Jefferson's text that some laughter was involved, but was *he*, for example, laughter, a pronoun, or a filler in the original utterance of M? Why is *Not* printed only with an initial capital rather than in all capital letters, if the word itself "is spoken much louder than the surrounding talk" (Atkinson & Heritage, 1984, p. xii)? Is it true that the prolongation of *floo:r* affects only the vowel and not the following continuant consonant? And does *::* = 2 × *:* in terms of temporal duration, i.e., do two colons indicate twice the time of prolongation as one colon? At the very least, such usage of colons may be confusing to readers who are accustomed to the conventional use of the colon as "a punctuation mark: used chiefly to direct attention to matter (as a list, explanation, quotation, or amplification) that follows" (*Merriam-Webster's collegiate dictionary*, 11th ed., 2003, p. 244). And – to return to our criterion number 1 above – were all these notations necessary for the reader's understanding of the analysis of the passage? And finally, were the notations actually used in the published analysis of the passage? O'Connell and Kowal (1995a, p. 99) have observed that Jefferson (1984) has nowhere explained her notation for laughter and has made no further use of this example in any systematic way.

3 *The internal integrity of words should not be interrupted by any supernumerary symbols.*

In addition to Jefferson's *flo:r* in *Example 4.2*, the following three syllables of the word *grandmother* and three laugh syllables (transcribed as @ in Du Bois, 1991, p. 87) manifest the difficulty occasioned by failure to follow this criterion:

Example 4.3
gra@ndmo@the@r

The original pronunciation of this word is not transparently clear. We might add that the phonetic realization of the laugh syllables themselves in the original utterance as *ha*, *he*, *hu*, or perhaps some other articulation is also not clear. And even though the laughter is transcribed here as three additional syllables, it is not at all clear whether they are such in the original utterance; laughter overlaid on words does not necessarily yield additional "laugh syllables" (see also our Chapter 17).

4 *Subjective perceptions and/or categorizations of the transcriber should not be recorded as objective measurements.*

An example from Levinson (1983, p. 328; see Schegloff, 1979, p. 37; see also our *Example 16.1* in Chapter 16) was presented as three turns on the grounds that "a significant pause," i.e., a pause long enough in duration to signal to the second speaker that a transition relevant place had been missed, and that a "repair" had occurred on the part of the first speaker. It should be noted that the "significant pause" was not measured, but was estimated perceptually as 0.2 s in duration. Pauses of that duration are not only not "significant," they are not even recorded by many researchers. Use of such procedures for transcription on the part of conversation-analytic researchers constitutes a completely subjective element in their presentation and interpretation of data.

5 *Symbols used in transcription systems should stand for only one feature of the spoken discourse, and no feature should be represented by more than one symbol.*

The dash, for example, in the Jeffersonian notation system, as presented by Atkinson and Heritage (1984, p. xii f.) – in addition to its conventional linguistic use "to indicate a break in the thought or structure of a sentence" (*Merriam-Webster's collegiate dictionary*, 11th ed., 2003, p. 316) – can stand for a short untimed pause within an utterance or for "a halting abrupt cutoff" or for "a stammering quality" or for a timed pause with reference to gaze; similarly, double parentheses may indicate a pause or "some phenomenon the transcriptionist does not want to wrestle with" or "other details of the conversational scene" or "various characterizations of the talk." O'Connell and Kowal (1990b, p. 454) have enumerated 14 ways in which a silent pause of 0.375 s has actually been listed in notation systems. All these examples violate the criterion of parsimonious use of notations.

6 *Descriptions, explanations, commentaries and interpretations should be clearly distinguishable from the transcription of phonological features of spoken discourse.*

The incorporation within parentheses of prosodic, paralinguistic, extralinguistic, and/or interpretative elements in a transcript diminishes the clarity and legibility substantially (see, e.g., Dorval, 1990, p. 152 f.).

7 *The transcriber, considered as a language user, is 'often quite unreliable' (MacWhinney & Snow, 1990, p. 457).*

The problem is decidedly *not* that transcribers do their work in a slovenly manner. Quite the contrary, the source of the problem is that transcribers suffer from the same perceptual deficits and biases as all language users, and are intent upon searching for meaning. Thus they alter spoken discourse by deleting words and phrases that were primarily spoken discourse markers, and substituting words and phrases close in meaning to the intended word or phrase. Another source of unreliability in transcribers is the deliberate decision to leave out of transcripts passages that are "not very interesting" or are "too extensive" (Gutfleisch-Rieck, Klein, Speck, & Spranz-Fogasy, 1989, p. 5; our translation). This decision to eliminate some spoken passages is reminiscent of Goldman-Eisler's (1961, p. 167) equally arbitrary decision to leave out all "irrelevant vocal productions, i.e., noise, such as repetitions of the same words or other obvious forms of marking time vocally."

Transcription as Theory

Even a partial solution to all these problems of transcription involves training, experience, meticulousness, and mutual agreement of multiple transcribers or monitors; there is no easy way to obtain good transcripts. Bruce's (1992, p. 145) suggestion that "a transcription system should be easy to write, easy to read, easy to learn, and easy to search" is not entirely realistic.

Spontaneous spoken discourse can be investigated systematically only with the help of reliable transcripts. Three decades ago, Ochs (1979, p. 44) published a chapter entitled *Transcription as theory*, in which she stated: "The transcript should reflect the particular interests . . . of the researcher." Her statement is quite relevant; and the corollary of it should be that transcripts must be selective in the sense that they provide the response measures proper to the finality of the specific research as defined by "the particular interests . . . of the researcher." In any event, the spontaneous spoken discourse itself always remains the database. This methodological generalization seems to be accepted universally by the research community. Hence, Linell's (2005, p. 118) claim that, in the written language bias tradition, transcripts are "simply taken as *the data*" appears to be exaggerated. On the same topic, Cook (1990, p. 1), in his article *Transcribing infinity*, has emphasized both the impossibility of comprehensively transcribing everything relevant to a given corpus of spoken discourse and the correlative necessity for selecting a transcription methodology based on agreed-upon theoretical principles.

Chapter 5
Rhetoric

As a medium, writing is a million times weaker than speech. It's a hieroglyph competing with a symphony (Menand, 2004, November 8, p. 104).

Chapter Prospectus

Chapter 5, *Rhetoric*, engages the spoken discourse of experienced speakers and their ability to simulate spontaneity even in their reading aloud. Such discourse is communicatively meaningful precisely because it is consequential for their lives rather than a task performed under laboratory conditions. Our own research gradually has led us to the foundations of spontaneity through the examination of these skillful simulations. Similar devices are used rhetorically in reading aloud and in spontaneous spoken discourse. Considered psychologically, current rhetoric, including religious, political, and artistic forms, pursues an ideal delivery which is not so much syntactic in its focus as communicative, cogent, persuasive, pro-social, and conversational. It is the listener's or audience's expected reaction that must dictate the speaker's strategies, precisely because the purpose of the speaker must always be effective communication. The chapter ends with a discussion of the overarching importance of literacy and orality as twin foundations for a contemporary rhetoric of spoken discourse in both the public domain and in everyday social interaction.

What Rhetoric Is All About

Somewhat more than a decade ago, Gill and Whedbee (1997, p. 157) began their chapter on rhetoric by stating, "There is little consensus as to the meaning of the word *rhetoric*." A look at *Merriam-Webster's collegiate dictionary* (11th ed., 2003, p. 1069) confirms their statement insofar as it provides a variety of usages of the word:

> **1 :** the art of speaking or writing effectively: as **a :** the study of principles and rules of composition formulated by critics of ancient time; **b :** the study of writing or speaking as

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_5, © Springer Science + Business Media, LLC 2008

> a means of communication or persuasion **2 a :** skill in the effective use of speech **b :** a type or mode of language or speech; *also*: insincere or grandiloquent language **3 :** verbal communication: DISCOURSE

In other words, one must be very careful to specify what one means in using the word *rhetoric*, precisely because it can mean so many different things. Some would wish to begin dismissively with the *also* in the quotation above: Rhetoric is unworthy of upright people and is instead characteristic of sly, unprincipled politicians. That is hardly the place we wish to begin. In fact, it was partly the didactic teachings of traditional rhetoricians regarding the use of such devices as pauses of longer duration and repetition that led us to recognize the potential of such means for either effective or ineffective oral discourse. However, the bulk of the teachings in ancient Greco-Roman rhetoric have consistently been on the preparation of the text itself, typically written first and then delivered as if produced spontaneously (Quintilian, ca. 95/1958). Nonetheless, the *actio* and *pronuntiatio*, i.e., the nonverbal and verbal enactment of the presentation, have always been acknowledged as important.

Our own primary preoccupation with regard to rhetoric in the following is in accord with the dictionary definition **2 a**: "skill in the effective use of speech." The reader should note that the skill in question here is a dialogical or communicative skill on the part of the speaker, but with a view to the listener or audience. It is ultimately the listener who is the arbiter of the effectiveness of the speaker. Still, we are not in the business of evaluating speech, but of characterizing the use of rhetorical devices in a variety of settings. Our preoccupation with public and particularly, political spoken discourse is not intended to exclude other spontaneous spoken discourse in more private everyday settings from the realm of rhetoric: The youngster pleading with a parent for an addition to his or her weekly financial allowance and the geriatric patient arguing for the extension of his or her driver's license will both choose to make use of whatever rhetorical devices they can muster to accomplish their communicative goal. In other words, it is our assumption that rhetorical devices play a role in every utterance.

To return once again to political rhetoric, more than 20 years ago, Atkinson (1984a) noted the burgeoning conversational style characteristic of modern political oratory – a consequence of the new setting of TV oratory, the informality of the family living room. In this respect, he was following Ong's (1982) emphasis on secondary orality as the consequence of modern technology. In a similar vein, Jamieson (1988) has entitled a book on modern rhetoric *Eloquence in an electronic age* in acknowledgement of the fact that TV has indeed changed the way both speakers and audiences approach rhetoric. Both Atkinson and Jamieson have referred to former U.S. president Ronald Reagan as the ideal representative of this style of political oratory. According to Atkinson (1984a, p. 167), Alistair Cooke has described Reagan's first inaugural speech "as the first 'conversational inaugural' in American history." In our own view, Reagan's second inaugural has provided an even better example of the shift

to a conversational style of political rhetoric. His second inaugural contained a passage, certified by his speech writer Peggy Noonan (1990) as spontaneously substituted for the prepared script during the speech itself, in which he masterfully and humorously referred to the last-minute change of location:

> *Example 5.1*
> We stand together again at the steps of this symbol of our democracy – or we would have been standing at the steps if it hadn't gotten so cold. Now we are standing inside this symbol of our democracy. (Reagan, 1985, p. 374)

Another official transcript (Reagan, 1988, p. 58) has provided an even more informal version of this passage:

> *Example 5.2*
> We stand again at the steps of this symbol of our democracy – well, we would have been standing at the steps if it hadn't gotten so cold. [*Laughter*] Now we're standing inside this symbol of our democracy.

Example 5.3 is from our own transcript of the original recording; pauses have been noted in parentheses in seconds:

> *Example 5.3*
> we stand again at the steps (.75) of this symbol of our democracy (.36) or we would've been standing at the steps if it hadn't gotten so cold (.47) and now we're standing inside this symbol of our democracy

A comparison of *Example 5.3* with *Example 5.1* shows that our transcript has eliminated "together," has contracted "have" and "are," and has introduced "and"; and compared with *Example 5.2*, it has changed "well" to "or," has contracted "have," and has introduced "and." We may also note that the laughter of the audience indicated in brackets was far more diffused throughout the passage than is indicated in *Example 5.2*, that the pauses are essential to both the conceptual orality and to the initiation of laughter, and that the articulation rate before Reagan's departure from his written text was approximately 1 syl/s faster than after it ($5.9 > 4.9$ syl/s). But the departure from his text constitutes the most basic component of this shift toward relatively more conceptual orality. Finally, one may note that laughter on the part of an inaugural audience is most extraordinary – one more indication of Reagan's dialogicality with his listeners (for more detail on laughter, see our Chapter 17).

Jamieson (1988, p. 164) has summarized Reagan's oratory as follows:

> Even the most formal of Reagan's speeches are written in a conversational style and delivered in a conversational voice. Where his predecessors brought a formal style and oratorical delivery to their inaugurals and State of the Union messages and a more casual style and delivery to press conferences and extemporaneous speeches, Reagan is consistently conversational in both environments.

The general trend toward conversational style as an ideal in modern political oratory reflects also the shift from a formal conceptual literacy to more conceptual orality, whether it must be thought of as actually simulated or genuine.

Some Typographical Helps to Rhetoric

Even in silent reading, there are many ways in which emphasis, segregation, coherence, and import can be indicated in the written text, and all of these textual devices can influence in turn the manner of reading aloud. In a sense, they can be considered a sort of adjunct or extension to the role of basic punctuation marks. Thus, variation in font, size, distribution, and even color of print can be used.

One should note that the proper use of graphics is involved even in these simple variations. The field of graphics has become an artistic specialty in recent years, largely as a consequence of the electronic revolution. Entire books can be rated from optimal to worst-case scenarios on the basis of their graphical presentation. A skillful use of graphics can enhance communication even to the silent reader, while a clumsy use thereof can make for a busy, incoherent, poorly organized presentation. An excellent example of this is the Power Point Presentation, which has become almost the universal mode of presentation at many scientific conferences (for a critical discussion, see Tufte, 2003). Bullets and pointers, however, cannot change a presentation's basic intelligibility and credibility or lack thereof. Such presentations can unfortunately become merely an ever more sophisticated version of garbage-in-garbage-out. Nonetheless, when the right words, phrases, data, or formulae are highlighted, these typographical means can be most useful. Still, some presenters seem to be lulled into thinking that the effectiveness of Power Point communication is almost automatic; in any event, an inordinate number of such presenters have fallen into the practice of mumbling into the microphone, almost as if their verbalizations were totally redundant in light of the graphics. Another example of the multiplication of useless graphics is the dumbing-down seen in many introductory college textbooks through the use of inset boxes – with the mandatory flow charts, bullets, and sense lines – for many of the concepts already presented in the text. Signage for the direction of street traffic is another special case of the practical importance of lucid graphics: Position, shape, color, size, letter and number font, and the ever present danger of spatial clutter all become relevant in the rhetoric of signage.

Some Prosodic Principles

Prosody is one of the most important tools of oral rhetoric. In ancient rhetoric, it was subsumed under *pronuntiatio*. In the cases of both reading aloud and spontaneous spoken discourse, the implicit generalization that louder, bigger, and more frequent are all better must be avoided. For example, TV commercials are often televised with a louder volume and at a more rapid articulation rate than is regular programming; and it is not uncommon for viewers to make use of the mute button in these instances. Listeners generally do not wish to be browbeaten. The subtle, infrequent use of special prosodic means avoid this pitfall. In recent years, former President Bill Clinton has provided two dramatic

negative instances that bear a certain analogy to "The lady doth protest too much, methinks" (Shakespeare, *Hamlet*, III, ii, 239). The inordinate increment in loudness, accompanied by a number of nonprosodic devices such as the use of emblematic finger pointing, leaning forward, and scowling, has pinpointed both his lying (Upchurch & O'Connell, 2000) and his defensiveness (Fox News, September 24, 2006, interview with Chris Wallace). On the other hand, the very selective use of long pauses and of extremely slow articulation rate as part of the local organization in two of the most famous citations from the inaugural speeches of U.S. presidents manifests how rhetorically effective such devices can be (see Kowal, O'Connell, Forbush, Higgins, Clarke, & D'Anna, 1997, p. 25 f.). The first of these two citations is from Franklin Delano Roosevelt's first inaugural (with measured silent pauses indicated in parentheses in seconds):

Example 5.4
The only thing we have to fear is (1.18) *fear itself* (1.30).

The second citation is from John F. Kennedy's inaugural:

Example 5.5
And so, my fellow Americans: (1.12) *ask* (0.24) not (1.04) what your country can do for you.

Such a concentration of pauses considerably longer than a second in duration is most unusual. And the articulation rates with which the italicized words were uttered (2.78 and 2.50 syl/s, respectively) were extraordinarily slow. Once again, we are not dealing with mean rates in these instances; we are dealing with outliers, devices deliberately chosen for rhetorical effect. It is no accident that both these passages could be heard several times a week as recently as 2007 on the David Letterman Show as contrasts with the stumblebum absence of rhetoric in short passages of spontaneous speech by President George W. Bush. Obviously, these passages by President Bush were deliberately sampled for humorous impact on Letterman's TV audience, not for an objective presentation of Bush's style.

Some Other Relevant Measures of Rhetorical Performance

We have already mentioned two such measures above, duration of silent pauses and articulation rate, measured in syllables per second (syl/s); both are typically used rhetorically to slow down the speech rate, i.e., the overall number of syllables spoken within a period of time. In these cases, we are not interested in means and standard deviations, but in exceptionally extreme outliers. Still, one must know the statistical context in dealing with the outliers. With regard to the *Example 5.4* and *5.5* given above from inaugural speeches (Kowal et al., 1997), across all the inaugurals for which audio recordings were available at the time of this research, from F. D. Roosevelt's first inaugural (March 4, 1933), to Bill Clinton's first inaugural (January 20, 1993), the mean duration of all silent pauses was 0.97 s (SD = 0.15) and the mean overall articulation rate was 4.37 syl/s (SD = 0.39). In

other words, the silent pauses in the crucial positions of the rhetorically successful passages by Roosevelt and Kennedy were all about one standard deviation longer in duration than the overall mean. And the articulation rates of "*fear itself*" and "*ask*" were several standard deviations slower than the mean.

But rhetorical devices can also be specific to various genres. Kowal et al. (1997, p. 14) have compared additionally the overall means of inaugural addresses with those of spontaneous spoken discourse, both in German and in American English. The articulation rate of the inaugurals was strikingly slower than that of either German or American spontaneous spoken discourse ($4.37 < 6.09 < 6.55$ syl/s), and the mean silent pause duration in the inaugurals was similarly longer than the German or the American means ($0.97 > 0.55 > 0.38$ s).

The point to be made from these data is that rhetorical devices can be applied quite differentially from genre to genre and from one setting, context, or purpose to another. The discovery of appropriate units of measurement and indeed of performance measures themselves is an essential stage in developing a methodology for the investigation of any corpus of spoken discourse.

To return for a moment to rhetorical devices used in written political speeches, Kowal et al. (1997, p. 10) have found that even the rhetoric of the written text changed over time in presidential inaugurals. In their comparisons of the written text of the inaugurals before F. D. Roosevelt with those from F. D. Roosevelt on, the following striking differences were to be noted in measures involving words (W) and syllables (S): W/paragraph ($155 > 51$), W/sentence ($35 > 20$), W/discourse marker ($920 > 199$) (as defined by Schiffrin, 1987); S/paragraph ($252 > 77$), S/sentence ($57 > 30$), S/word ($1.60 > 1.50$), and S/discourse marker ($1481 > 303$). In other words, one trend was found to be in the direction of simplification: shorter paragraphs, shorter sentences, and shorter words. Another trend was found to be in the direction of the increased use of discourse markers. These particles (e.g., *oh, well, but,* y'know) are operationally defined by Schiffrin (1987, p. 31) as "**sequentially dependent** elements which bracket units of talk." Both these overall trends reflect what has been termed conversational style in modern political rhetoric (see Atkinson, 1984a, p. 166 ff.). The reader should note that all these differences are statistically highly significant and of large magnitude. As a corollary, these *written* differences necessitate different rhetorical devices for reading aloud. Unfortunately, the absence of acoustic recording in the earlier period makes an actual empirical study of such differences impossible. What is surprising in these data is the constancy of punctuation across the two chronological periods: W/punctuation ($13 < 14$) and S/punctuation ($22 = 22$). Another comparison involves only first-person pronominals: The percentage of singular/total first-person pronominals decreases over time ($43.4\% > 14.8\%$), whereas the overall use of first-person pronominals (first-person pronominals/1,000 W) increases ($39 < 64.8$). In other words, *I* yields dramatically to *we* – which one may well construe as an indication of the use of conversational style. All these measures must be taken into account if one is to understand the baseline of the written text as foundation for possible rhetorical devices of reading aloud. Finally, one

should note that these measurements constitute a quantitative approach to rhetoric rather than a qualitative one.

The measurements exemplified above must not be taken as the sole mandatory measurements relevant to rhetoric in written texts. For example, in poetry the W/line may index much better what rhetorical devices are possible or appropriate in reading aloud. In dialogue, the W/turn or S/turn may characterize a corpus better than any of the measures discussed above for the inaugural corpora. In dialogue too, the type of transition at turns can be extremely important as a rhetorical characterization of a corpus; the percentages of successful and unsuccessful interruptions, overlaps, and smooth transitions (with or without silent pauses) can shift such a characterization very notably. In short, the rhetorical methodology applied to written text and its reading aloud as well as to spontaneous spoken discourse must be flexibly adapted to the corpus that is to be characterized rhetorically. Such methodologies are necessarily complex, even as rhetoric itself is complex.

A final example of how a competent journalist can enlist a variety of hesitations in the service of rhetoric is to be found in a long interview question asked of Ronald Reagan by Walter Cronkite, a prominent American TV journalist of the late twentieth century. The efficiency of his questioning method can be appreciated from the fact that Reagan answered willingly, even jovially, while laughing aloud, although the gist of the question might well have been considered impertinent (in parentheses, duration of pauses in seconds):

Example 5.6

What what really philosophically is different (.8) from (.3) our (.43) going down to help a a (.3) democratic government uh (.37) sustain itself against guerilla activity (.27) promoted from the outside (.43) uh Soviet and Cuban uh uh aid as we believe it to be (.27) or as your your (.33) administration says it is (.7) and (.43) Afghanistan (.27) uh the El Salvador is in our sort of geo (.4) political sphere of influence Afghanistan on the border of the Soviet Union is certainly in their geo (.43) political sphere of influence (.47) uh they went in with troops uh to uh uh to support a Marxist government friendly to them (.63) what what's where where's the where why isn't that a parallel situation. (Kowal, Bassett, & O'Connell, 1985, p. 15)

On paper, the "what what's where where's the where" sequence appears to be simply outrageous. As spoken, it constitutes the core of the buffer function served by Cronkite's long, hesitant run-up to his simple, bold question: "Why isn't that a parallel situation?" In our Chapter 23, we will return to Cronkite's question as an example that instantiates all the theoretical principles that we emphasize throughout this book and develop in detail in Part IV.

Literacy and Orality

Walter Ong's (1982) *Orality and literacy* has served to introduce a wide audience of his readership to some important differences between a literate and an oral culture. But there is also a sense in which literacy and orality become relevant

within an almost entirely literate culture, a theme that has been taken up in turn by Koch and Oesterreicher (1994, p. 587; our translation) from the point of view of linguistics. By limiting themselves to what Ong (1982) has referred to as secondary orality, an orality derived from the use of modern technologies such as telephone, radio, and television within a literate culture, they have distinguished two types of orality and literacy: *medial* and *conceptual*. Medial orality and medial literacy have to do with the modality in which communication is delivered – phonic or graphic – and always constitute dichotomous variables. Conceptual orality and conceptual literacy have to do with aspects of linguistic variation, "referred to in research vaguely as 'colloquial language/literary language', 'informal/formal', 'levels of elaborateness' etc." Koch and Oesterreicher have emphasized that conceptual literacy and orality constitute a continuum, and in a similar vein, Raible (2002) has noted that any clear-cut distinction between conceptual orality and conceptual literacy is blurred by the very fact that they are in a continuum. According to the anthropologists Scollon and Scollon (1995, p. 20), this overlap between the two terms *orality* and *literacy* "has made both of them all but obsolete" and at the same time has made the usefulness of the term *orality* "for the characterization of speech events and cultures" questionable. But the availability of Koch and Oesterreicher's (1994) distinction between medial and conceptual orality has made it possible for us to use orality in a far narrower and precise sense for purposes of empirical research.

The historical roots of literacy and orality have been traced in a much broader, largely anthropological context by Khosrow Jahandarie (1999, p. 1 f.) in his *Spoken and written discourse: A multi-disciplinary perspective*. The importance of this book lies in the author's effort to present a comprehensive overview of the history of "Oral Theory" as it has developed in classical studies, history, media studies, literary criticism, anthropology, and psychology:

> This volume presents a systematic, reasonably exhaustive, and critical review of the scientific literature on the differences between speech and writing and, particularly, the cognitive and cultural implications of these differences. It is unique in its multidisciplinary scope and analytical depth as it brings together, for the first time, this multiplicity of theory and evidence from varied disciplines.

And indeed, Jahandarie has brought together a significant discussion of 1455 references for this purpose. It should be noted, however, that all of his references are in the English language and also exclude most of the authors that we repeatedly advert to in this book (e.g., Graumann, Hörmann, Koch & Oesterreicher, and Linell).

Jahandarie has listed as "the six [actually seven] theorists most closely associated with the orality–literacy contrast": Milman Parry (1971) and Albert Lord (e.g., 1991), "the principal formulators of the 'Oral Theory' "; Eric Havelock (e.g., 1986), Harold A. Innis (e.g., 1951), Marshall McLuhan (1962), Walter Ong (e.g., 1982), Jack Goody (e.g., 1987), and David R. Olson (1994). The works of Boas (e.g., 1940), of Claude Lévi-Strauss (e.g., 1964–1971), and of Edward Sapir (e.g., 1921) have been mentioned only in passing by Jahandarie

(1999, p. 280 f.; but see Bringhurst, 1999). His claim that "probably the biggest blindspot in the oral–literate theories has been their almost total disregard for the findings of cognitive psychologists over the past three decades" (p. 151) puts him squarely in the corner of mainstream psycholinguistics. And yet, he has acknowledged that "it is obvious that we are still largely in the dark about the exact processes and mechanisms that determine the cognitive handling of spoken and written discourse" (p. 196).

A much broader, anthropological concept of *orality* has been critically discussed by Scollon and Scollon (1995, p. 27) in terms of a number of negative implications in the use of this concept. These would include: (1) a view of orality as an obstacle that has to be overcome; (2) a certain phonocentrism, i.e., the limitation to sound and the correlative exclusion of such sensory modalities as touch and vision; and (3) a logocentrism, i.e., the inordinate emphasis on words alone, to the exclusion of nonverbal material. As a possible substitute for, or at least a supplement to the concept of *orality*, they have suggested "somatic communication" in order "to make reference to the human body as the foundation of communication." We have found this notion of communication quite compatible with our own thinking and have accordingly chosen in Chapter 23 the term *somaticization of syntax* to characterize the structures of spontaneous spoken discourse that transcend sentential syntax. Nonetheless, we wish to retain the term *orality*, but without the negative implications listed by Scollon and Scollon.

Orality as a Rationale for Our Research

As the examples given above indicate, our research has always involved both medial and conceptual literacy and orality. Our first project was accordingly a comparison of reading narratives aloud with the subsequent retelling of the narrative (O'Connell, Kowal, & Hörmann, 1969, 1970). The context of literacy and orality has been continued in the comparison of political speeches with interviews of the same politicians (Kowal, 1991), and in the comparison of a more formal with a more conversational style in older and more recent inaugural speeches by presidents of the United States (Kowal et al., 1997). In addition, we have compared the speaking of interviewers with that of interviewees in radio and TV (e.g., O'Connell & Kowal, 1998). Chronologically, the projects have increasingly involved a relativization of conceptual literacy and conceptual orality, based on the fact that they are not discrete, but continuous variables. And the comparisons themselves have instantiated a gradual transition from literacy to orality.

The literacy of TV interviewers can be assessed in two ways: from the well-formedness of their sentential structures and from the fact that they can be seen to hold written notes in their hands or place them on a desk. One of the most illustrious of the post World War II German media journalists has for many

years been Günter Gaus, considered already in Chapter 3 in regard to his interview with Hannah Arendt. He has broadcasted or televised over 200 interviews in the German language. In his memoirs (Gaus, 2004, p. 189; our translation), he has characterized his preparation for an interview as follows:

> In a first draft I noted 60 or more questions, which I then reduced in the second and third pass to approximately 20 to 25.
>
> I not only sketched the questions, but formulated them meticulously word for word.

Paradoxically, although "the most interesting forms of broadcast talk have a feel of spontaneity" (Tolson, 2006, p. 11), Gaus has accomplished this goal precisely by careful scripting (see also Hilton, 1953). It is obvious from video recordings that he took notes with him into his interviews. The importance of this sort of preparation can easily be related to the fact that he received numerous awards for the excellence of his interviews. The well-formedness of his interviews can be assessed by a comparison of the changes needed in the translation of the audio transcript into a published version (Arendt, 1996) of his famous interview with Hannah Arendt: Only 20% of the necessary changes were in his own contributions, whereas 80% were in Arendt's contributions (O'Connell & Kowal, 1998, p. 550).

A Rhetorical Perspective for Everyday Talk

Our emphasis thus far has been largely on public rhetoric. We now turn our attention to rhetoric in everyday talk as emphasized in the research of Karen Tracy (2002, p. 26 ff.). Her "rhetorical perspective" regarding everyday talk has three important dimensions: (1) Above all, people are active in *choosing* means to express themselves for effective communication in various settings. (2) The fact that choices are involved makes people morally responsible as agents. (3) Such a rhetorical orientation is centered on problem solving insofar as rhetoric must always encounter conflicts of interest in the everyday setting. These three dimensions of everyday talk converge upon "building and reflecting identities," as Tracy has expressed it in the subtitle of her book. Such an emphasis, arising from the viewpoint of communication science, we have found to be quite compatible with our own psychological orientation regarding rhetoric. Tracy's first dimension appropriately orients the choice of expressive means to the needs of the listener. The second dimension is quite in accord with what we designate in Chapter 22 as verbal integrity – the moral responsibilities that interlocutors have toward one another and to society. The third dimension localizes personal growth in its interactive arena and in the sharing of personal perspective with the perspectives of one another – "interactional tension ... is part of the business of talking."

Chapter 6
Intentionality

This first step in planning an utterance is the conception of a communicative intention. In view of this end, appropriate means will have to be marshaled (Levelt, 1989, p. 4).

Chapter Prospectus

Chapter 6, *Intentionality*, engages the rationale for initiating speech, continuing to speak, and ceasing to speak. Hence, we are here concerned with the psychological meaning of intentionality rather than with the philosophical meaning as presented, for example, by Austin (e.g., 1962) and by Searle (e.g., 1983). Spontaneous spoken discourse is not a continuous or constant activity of human beings. It is a chosen activity; it must be initiated, and this initiation requires on the part of a speaker a reason or reasons to begin to speak and to continue speaking or not. The same is correlatively true of the listener: Listening must be engaged initially and then sustained; it is not automatic. The basic motivation on the part of both speaker and listener is a search for intelligibility and coherence that cannot be satisfied by nonverbal means only, but requires words. Even more fundamentally, these considerations are based on the fact that we are not dealing with a sort of *homo linguisticus*, one whose nature is to be constantly engaged in speaking or listening. Instead, we are dealing with the occasional speaker and listener.

Starting, Stopping, and Continuing

Both speaking and listening are motivated social interactions. There must, therefore, be a reason for a speaker to begin to speak: some need or desire that he or she feels can be fulfilled by verbal interaction rather than by nonverbal means. For example, if one enters a living room and spies a large amorphous mass of wooden furniture, one might ask one's host about it: "What is it? What does it do for a living? Is it an objet d'art?" Such a setting seems to fulfill ideally *Merriam-Webster's collegiate dictionary*'s (11th ed.,

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_6, © Springer Science + Business Media, LLC 2008

2003, p. 273) definition of a *conversation piece*: "Something (as a novel or unusual object) that stimulates conversation." One could, of course, snoop about alone, but this would be taken as impolite in the presence of a host. After all, as a guest, one enters a living room in order to interact with the people therein. It would be an arch insult to enter a home and not address one's host. These are simple expectations that we live by, and violations are immediately perceived as at least odd, if not insulting or pathological. We are all socialized into such expectations along with the acquisition of our native language.

Chris Raschka (1993; German translation, 1997) has written a lovely children's story that has been singled out as a Caldecott Honor Book. The story itself may be used to exemplify very eloquently the concept of intentionality in spontaneous spoken discourse. We wish to use it here to further explicate the requirement of intention in speaking and listening. Note that we are, in this instance, using a written example to explicate intentionality in spoken discourse.

First of all, the words of the book fulfill, even more eloquently than those of books written primarily for adults: Ong's (1982, p. 75) description of them as merely "marks on a surface." A children's book must be brought to life in a specific setting: an adult reading aloud beside a child – who perhaps very authoritatively assumes the prerogative of turning pages. In other words, the interaction that brings the words to life is quite essentially dependent upon the simultaneous seeing of pictures and hearing of words. Anyone who thinks of such a reading aloud as a monological role has never read a storybook to a child: The concomitant commentaries, pointings, exclamations, questions, and control of the sequencing by page turning are all very much dialogical and are vivid realizations of the intimate dialectical relationship between the written and the spoken in everyday life.

For our purposes here, however, the *narrative* itself is the object of our interest insofar as it has to do with the initiation and sustaining of both speaking and listening on the part of the characters therein. Two summaries of the story can be applied to our purposes: (1) On the page after the title page of the English original, the publisher has inserted: "Two lonely characters, one black and one white, meet on the street and become friends" (Raschka, 1993). (2) A newspaper announcement provides a somewhat longer description of the German translation (B., 1997, August 3, p. 4; our translation):

> MONOSYLLABIC. The two exchange few words in their "conversation." The boy in the baseball shoes is direct and forward, when he sees the other boy shyly and sadly looking away. He begins the dialogue with "Yo!", whereupon the other replies only half-heartedly "Yes?" The shy boy finds himself lured out of his reserve by the insistent questions of the black boy. It lasts for only a few pages, but a conversation begins in which at most two words are spoken at a time. In short, the shy boy has no friends, and the other boy volunteers himself. Why not?

The storyline is important for our purposes insofar as it prominently displays a major problem in initiating both speaking and listening: The black boy has no reason to be optimistic about the white boy's listening; and the white boy's shyness prevents him from engaging the black boy except with a hesitant "Yes?"

In fact, he leans away from the black boy. The story tells us how the difficulties are overcome.

The story accomplishes this goal in a quite charming manner – and certainly not through the use of syntactically well-formed sentences. The pictures carry the story, and the voice of the adult reader brings the pictures to life with only a very few words, mostly one word at a time. Table 6.1 summarizes the words along with the accompanying terminal punctuation of both the English and the German versions of the story: These constitute the total verbal content of the narrative. The respective needs that are invoked here are, on the part of the energetic black boy, the necessity to share his exuberance, and, on the part of the sulking white boy, the very joy of spirit the other boy stands for. The first boy must coax the other into interaction – first into acknowledgement of his presence, then acknowledgement of the possibility of interaction, then, quite specifically, acknowledgement of the offer of his friendship, and finally acceptance of the offer. All this is brought about for each boy with only the vertical string of words with their punctuation as given in Table 6.1. Externally, very little happens, but a touching story of new friendship is told.

The burden of communication in this instance seems to be on the energetic black youngster. He is the one who is motivated to speak, certainly not in the first instance the other boy. But the need for fun and friendship gradually and reluctantly emerges in the second boy's words, and his motivation both to listen to the message of the black boy and to respond to it in words comes to life. It is interesting that, typographically, the two boys end up in the last two spreads

Table 6.1 Words, Along with their Terminal Punctuations, in the English- (E; Raschka, 1993) and German-language (G; Raschka, 1997) Versions of *Yo! Yes?* (G: *Hey! Ja?*) On the Left (L) and Right (R) of Sequentially Numbered Spreads of Adjacent Pages

Spread	Words, Along with their Terminal Punctuations			
	E		G	
	L	R	L	R
1.	Yo!	Yes?	He!	Hä?
2.	Hey!	Who?	Hey!	Ja?
3.	You!	Me?	Du!	Ich?
4.	Yes, you.	Oh.	Ja, du.	Oh.
5.	What's up?	Not much.	Was läuft?	Nicht viel.
6.	Why?	No fun.	Wieso?	Kein Spass.
7.	Oh?	No friends.	Oh?	Keine Freunde.
8.	Oh!	Yes.	Oh!	Ja.
9.	Look!	Hmmm.	Hier!	Hmmm.
10.	Me!	You?	Ich!	Du?
11.	Yes, me!	You!	Ja, ich!	Du!
12.	Well?	Well.	Na?	Ja. . .
13.	?	Yes!	?	Ja.
14.	Yo!	Yes!	Hey!	Ja!
15.	Yow!		Ye-a-h!	

literally on the same page, whereas before the black boy who initiates the conversation always appears on the left, the shy white boy on the right page; one should note that the left–right sequence is used here by the author to indicate initiation and response, respectively. And so, the necessary intersubjectivity of the two boys, if their communication is to continue – that is, their mutual and reciprocal consciousness of one another – is symbolized by their co-presence on the first page of the last two spreads. Hence, in an extraordinarily simple interaction of two youngsters, the accompanying presence of intention both to speak and to listen is eloquently expressed – for the most part without well-formed sentences and without many words. The specific words used by the boys are themselves of interest: There are eight interjections in the English and nine in the German version; there are seven turns involving *you* and *me* in both the English and the German versions; and there are six *yes*es in the English and seven *ja*s in the German version.

In passing, some peculiarities of the words and punctuation might well be noted. In English, the question and answer are the same, "well," but in German different ("Na," "Ja") in spread number 12. In spread number 13, the question is entirely wordless, signified graphically by a question mark over the head of the black boy. Both examples illustrate the relative unimportance of the specific words to communicate the story and the importance of the question mark as a symbol of the momentary intentionality of the black boy. In general, one may note that the dramatic punctuation serves as a cue for the adult who reads the story aloud and suggests both the intentionalities of the youngsters and the appropriate prosody for the reader. In fact, in Table 6.1, the terminal punctuation listed under R for the white boy's responses in both the English and the German versions proceeds from question marks at the beginning to periods in the middle, and finally to dramatic exclamation marks at the end (for further considerations of punctuation see our Chapter 9).

The story thus manifests how reading a picture book aloud can truly simulate many of the properties of spontaneous spoken discourse, even the very special ones of a bedtime story told by an adult to a preliterate child.

A Historical Note on Intentionality

Intentionality undoubtedly has to do with consciousness. As such, it was completely incompatible with the Zeitgeist of the twentieth century in psychology. The myopia of behaviorism held the profession pretty much in thrall. And yet, in their everyday engagement of the social environment, human beings address one another with consciousness and purpose. Not that the purpose is always entirely transparent to the consciousness; even our simplest utterances can be complex, devious, indirect, and convoluted. But they are engaged with purpose.

For all its claim to be at the very core of modern psychology, cognitive psychology is still dealing with behavior as mechanistic, automatic, determined; cognitive

psychologists continue to "treat people as machines" (Costall, 1991, p. 163; cited in Linell, 1998, p. 58). We are convinced that psychology cannot effectively engage the occurrence of speaking and listening from such a mechanistic point of view. Until recently, the most extensive and definitive treatments of speaking and listening have been, respectively, suggested by Levelt (1989, p. 1) and by Handel (1989, p. 547). Levelt has considered the speaker as "a highly complex information processor," and Handel has acknowledged that his own emphasis on "the psychophysics of listening" leaves as unengaged "the role of the listener's knowledge and experience and of the listener's goals and intentions in representing the world."

It is precisely this level of "goals and intentions" that we wish to emphasize in our treatment of speaking and listening. Short of such a level of engagement, spontaneous spoken discourse ceases to be spontaneous and ceases to be discourse, defined appropriately, although incompletely, as: "**2 :** verbal interchange of ideas; *esp*: CONVERSATION" (*Merriam-Webster's collegiate dictionary*, 11th ed., 2003, p. 357). As we shall see, far more than "ideas" is spoken of and listened to in spontaneous spoken discourse. We might add here that the acoustic–auditory mode is nearly always supplemented in spontaneous spoken discourse by the optical–visual mode. Only telephone conversations are generally excluded from this supplement.

Chapter 7
From Monologism to Dialogicality

Speaking did not exist at all, until someone had been spoken to; speaking could devolve into monologue only after dialogue had been broken off or shattered (Buber, 1967, p. 13; our translation).

Chapter Prospectus

Chapter 7, *From Monologism to Dialogicality*, engages the fact that spontaneous spoken discourse is essentially dialogical and is correspondingly most clearly recognizable through the observable phenomenon of turn-taking. Language use is always a social and cultural engagement rather than some sort of solipsistic behavioral phenomenon, as it would appear to be in the reductionistic treatments and artificial experimentation of mainstream psycholinguistics. Empirical research dare not neglect this essential component of spontaneous spoken discourse.

Where Are We?

In our search for a concrete way to contrast monologism and dialogicality, we encountered an appropriate example in a presentation at a recent Psychonomic Society convention. A paper by Swets, Ferreira, and Altmann (2006, November 18, p. 29) manifests the differences more clearly than any example we might have concocted. The paper was entitled "*Where was I?*": *A psycholinguistic investigation of interruptions to language production*. We include here their entire abstract:

> When people communicate in a dialog, the speech stream of one speaker is sometimes interrupted by the speech stream of another. In such cases, it is often difficult for the interrupted interlocutor to return to the point where he or she left off. Hence, interruptions present an interesting problem concerning language production: How do speakers keep track of where they were before being interrupted? We report four experiments that investigate this unexplored issue. Experiment 1 used a seminatural dialog in which a confederate interrupted at predetermined narrative junctures.

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_7, © Springer Science + Business Media, LLC 2008

> Measures of resumption difficulty reveal that interruptions with conflicting conversational goals are particularly disruptive, as are interruptions requiring long verbal responses before resumption. Experiments 2 through 4 investigated similar processes for sentence production. Results demonstrate that verbal and nonverbal interruptions early in sentence production are more disruptive than later interruptions. We discuss the implication of these results for theories of language production.

The empirical logic as evident in the experimental design is actually of more interest to us than the findings themselves. The "return to the point where he or she left off" is the first formulation of interest. The "point where" is a spatial metaphor, whereas the dialogue in question takes place in time. It is, then, really a question of the "time when," not the "place where," insofar as the post-interruption dialogue has moved forward in time. And, unless the interruption is totally trivial, irrelevant, or empty, it brings the dialogue necessarily to a new interactive moment of resumption. In short, at that moment, the post-interruption dialogical situation cannot be properly conceptualized as a regression to a previous moment in time. To consider the speaker's goal at this moment to be exclusively the retrieval in memory of what he or she had been saying *before* the interruption took place is therefore a thoroughgoing misunderstanding of the nature of dialogue. The history of a dialogue is indeed relevant to continuation, but it is dialogically relevant only as part of the momentary situation to which the dialogue has now advanced in time. Swets et al. have insisted that the "return to the point where he or she left off" is in accord with Levelt's conceptualization of the resumption process post-interruption. But, their conclusion that "interruptions with conflicting conversational goals are particularly disruptive" simply confirms the ordinary finality of dialogue: To try to resolve the conflict so that the dialogue can advance in a coherent direction. In other words, conflicting goals move the dialogue in multiple directions at the same time and must be somehow redirected. Of course, this does not constitute genuine disruption at all, but simply points up the need for ongoing clarification of goals because the dialogue is moving forward in time. The moment-to-moment need for clarification is intrinsic to any dialogue. Or, as Lueken (1996, p. 88 f.; our translation) has put it:

> My comprehension of what another person wants me to understand develops only in the course of the dialogical process, sometimes aided and abetted by the correction of misunderstandings.

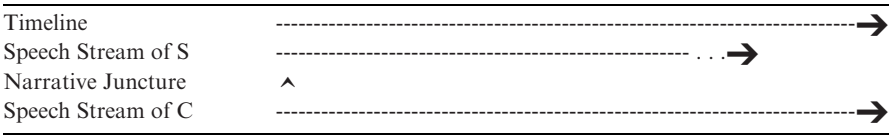
But Swets et al.'s experiment conceptualizes interruption as synonymous with disruption, not as forward dialogical movement in time. Setting up an artificial situation in which "a confederate interrupted at predetermined narrative junctures" sets the speaker up as a monological agent and does indeed make the interruption trivial, irrelevant, and empty. The resumption needed in a post-interruption situation does not answer the question *Where was I?* or even *Where am I?*, but must instead always reflect *Where are we now?* Or to put it in another way – without the intrusion of the spatial metaphor, – the question for the researcher at any moment in a dialogue is: What is the dynamic movement at

this moment in time, and how are the interlocutors to proceed from this moment on? Actually, for an interlocutor to revert to the "*Where was I?*" question constitutes a narcissistic withdrawal from dialogue; this would not only impede the forward movement of the dialogue in time, it could itself be highly disruptive. This is not intended as an assertion on our part that the interrupted speaker's first psychological inclination in time might not indeed be to stay with the "*Where was I?*" or – later – to revert to it. The point to be made, however, is that, by perseverating in or reverting to that preoccupation, the interrupted speaker would be essentially neglecting the ongoing dialogue by failing to integrate the interruption into the ongoing discourse as a legitimate and indeed essential part thereof.

And so, the Swets et al. experiment has attempted to generalize to dialogue from a very unnatural situation in which the confederate has no role other than that of the stooge who interferes with the experimental subject, and the speaker in this "seminatural" setting may have no investment in the topic or procedure that the experimenters refer to as a dialogue. The experimental subject is not engaged in a genuine dialogue at all, but in a monologue which is to be disrupted at the behest of the experimenters. One should also note that their abstract begins with a concern about people communicating "in a dialog," but ends with "the implication of these results for theories of language production" – a phrase that in itself carries monological overtones.

In Figure 7.1, we have sketched the temporal relationships of the utterances of the speaker (S) and the confederate (C) and the interruption in the Swets et al. experiment. The horizontal axis is a timeline. The "speech stream" of S must be interrupted by the "speech stream" of C, in accord with the operationalization of the experiment: While S continues to speak "seminaturally," C awaits a "narrative juncture" at which to articulate a preplanned interruption. Note that "narrative juncture" here does not imply a pause; the "speech stream" of S continues into the interruption. The abstract does not indicate that there is any necessary connection or semantic relationship between what S has said and what C says; apparently, they are both just following instructions. To label any of this as dialogue seems quite inappropriate. It is, in fact, monological insofar as the interest of the experimenters is only in what S says; the role of C is literally to affect what S says. What is abundantly clear in this setting is that the demand characteristics set up by the experimenters dictate what happens. There is certainly no spontaneous spoken discourse here insofar as a basic characteristic of spontaneous spoken discourse is absent – open-endedness (see our Chapter 21). The interruptions do not come from C's paying attention to the substantive message of S. Note also that we have inserted a triple ellipse at the termination of the speech stream of S in Figure 7.1. This indicates that S is not able to complete her or his utterance precisely because of the interruption. Were this not to be the case, we would have a simple overlap, but no genuine interruption. The only operationalization of interruptions given by Swets et al. (2006, November 18, p. 29) in their abstract is that "the speech stream of one speaker is sometimes interrupted by the speech stream of another." In other words, as diagrammed in Figure 7.1, C begins

Figure 7.1 Interruption of a Speaker (S) by a Confederate (C) at a Narrative Juncture



to speak while S is still speaking. This criterion for interruptions is actually neither necessary nor sufficient to define an interruption (for further details regarding interruptions, see our Chapter 16).

In summary, then, Swets et al. (p. 29) have claimed to be dealing with "a dialogue," but have only extracted some information that seems quite irrelevant to dialogue insofar as it pertains only to an individual speaker from "a semi-natural dialogue." In a dialogue engaged under genuinely natural conditions, the question for speakers at the moment of resumption after an interruption is generally not "where they were before being interrupted," but how the interactive dynamic has changed the post-interruption moment and how it is to be dealt with *now* – in real time. To ask the question solely about the pre-interruption setting is to treat the whole transaction of an interruption not as a transaction at all, but as a solipsistic behavior on the part of a speaker. An interruption is not something that goes wrong in a speaker's own little world, not something necessarily "disruptive," but rather one way of furthering the inevitable moving forward of a dialogue in real time. This inevitable moving forward involves typically an emotional reaction on the part of the speaker at being brought up short by the interruption; the psychological setting may be radically changed thereby for the speaker from what it had been a moment before.

Mainstream Psycholinguistics and Monologism

O'Connell and Kowal (2003, p. 191) have contended that mainstream psycholinguistics has been predominantly monologistic in its orientation ever since its inception in mid-twentieth century, i.e., that it "is concerned only with the person in whom cognition takes place and from whom communication proceeds." This is essentially an asocial methodology or what Clark (1985, p. 179), as we have noted above, called "the *individualist* view of language use."

It is hardly surprising to note that a psychology that historically specialized in laboratory experimentation and originally in introspection should view language use from the vantage point of the individual speaker. The approaches during mid-twentieth century and the advent of modern-day psycholinguistics were, after all, closely guarded by the twin powers of behaviorism and positivism. Neither had a penchant for the cultural or social and neither was modest enough to acknowledge competitors for the domain of language use. Hence, the early days of mainstream psycholinguistics were not

exactly oriented toward the listener. After all, listeners could hardly be characterized as overtly and observably active in any verbally relevant way.

There had indeed been some nineteenth century signs that the dialogical was not to be entirely disregarded and neglected. But Lazarus's (1879/1986) emphasis on the investigation of conversation went unheeded (see also Käsermann & Foppa, 2003, p. 767), and even Wundt's (1900–1920) *Völkerpsychologie* was not able to arouse an interest in the socio-cultural nature of language use. Similarly, the social issues raised by World War II and the concomitant rise of clinical psychology failed to jar the psychology of language use loose from its individualist moorings. Instead, the formulation of information theory and the birth of generative grammar took central stage and won over psychologists interested in language use to a mainstream psycholinguistics which felt at home in the laboratory and comfortable with concocted and unrealistic strings of words, phrases, and sentences.

Even so, the monologistic orientation of mainstream psycholinguistics was astoundingly strong. The influence exercised by linguists such as Noam Chomsky and by psychologists such as George A. Miller was very powerful through the first decades of the new psycholinguistics. The prestige of both MIT and Harvard Universities also bolstered the new orientation to a psychology of language use. Linell (1998, p. xii f.) has summarized the spirit of those times among mainstream psycholinguists as the following theoretical orientation:

> Thus the paradigm of dialogism must be understood in contrast to something else, namely 'monologism'. The latter is the dominant theoretical framework in the language sciences. The term alludes to the tendency to identify the speaker alone as the origin of the utterance. Basically, such a framework adopts some version or other of the following theories; cognition as individually-based information processing, communication as information transfer, and language as a code.

Note that all of these preoccupations are about the dealings of the individual speaker, without any advertence to the listener or to any other dialogical or socio-cultural consideration, and they do indeed constitute a veritable 'monologism.'

The phrase "cognition as individually-based information processing" contains most succinctly the monologistic notion. Cognition inheres within an individual; only individual human beings know, and even matters known by many individuals are not eo ipso dialogically known, i.e., known by dissemination from individual to individual. The terms *cognitive* and *cognition* have, in the course of time, taken on almost mystical importance. And cognitive psychology has become the home base of mainstream psycholinguistics. It excludes reference to dialogical and socio-cultural variables, and insofar as it engages only individual knowledge, it also excludes the domains of human intentionality, motivation, affect, intersubjectivity, and volition. By the beginning of the twenty-first century, the term *cognitive* had been reduced to a meaningless good-old-boy designation, a sort of *Good Housekeeping* or an Underwriters' Laboratory (UL) stamp of approval – and nothing more.

The implication in a cognitive and monologistic approach to language use that communication is simply "information transfer" has been most clearly expressed

by Norman and Rumelhart (1975, p. 4): "People use language to convey information." Such a conduit theory of human communication "distorts the act of communication beyond recognition" (Ong, 1982, p. 176). It disregards the fact that understanding is a creative act: "The utterance in itself does not convey any information to the hearer: it only guides the hearer to creating the information for himself" (Hörmann, 1981, p. 308). "Information transfer" is clearly not an adequate basis for a psychology of language use.

Finally, language is more than a code; word meanings remain mere potentialities for meaning and are not automatically decoded without further ado, i.e., without the intervention of thinking and understanding human beings. The meaning of the code itself is neither self-sufficient nor automatic.

Is All Human Speech in Principle Dialogical?

In a section entitled *Monological speech and thought*, Linell (1998, p. 267) has taken up the obvious problem that some speech and thought appear isolated and hence monological. He has argued that dialogism "is supposed to be a theoretical framework valid for monological speech and thought as well" and that "thinking is largely arguing with other dimensions of one's self.... The thinker is, according to the dialogistic theory, not a Cartesian ego, but a profoundly social being." Furthermore, "monological speech is thus intrapersonally dialogical, though interpersonally it exhibits only limited dialogicality."

In short, a sort of distancing from oneself is necessary in order to dialogue effectively within oneself, and this reflexive dialogicality incorporates the monological within its theoretical ambit. It is undoubtedly true that even our most private thinking is contextualized both epistemologically and metaphysically by an ambient reality which is eminently socio-cultural. The ontogenesis of language use must also be said to be pre-eminently dialogical; it is doubtful that a child could learn language and language use from the sole presence of TV, without the intervention of speaking adults who interact with the child. We might even consider for a moment the extreme case of a letter written to a correspondent in the privacy of one's room. This is truly an extreme case that appears to be legitimately monological. But the ultimate rationale even of such isolated, solitary composition is not at all monological, but dialogical: We write in answer to letters and in turn request answers to our letters. The very use of the terms *correspondence* in English and *Briefwechsel* (literally letter-exchange) in German are eloquent witness to the socio-cultural embeddedness of letter writing (and receiving).

Hence, dialogism is to be embraced not as a supplement or appendage to mainstream psycholinguistics, but as a radical departure from its monologism. We would accordingly be in complete agreement with Linell (1998, p. 23) that mainstream psycholinguistics is "strongly misleading if presented as a full theory of communication through spoken interaction."

Chapter 8
Listening

But his speaking is rather a voiced, attentive listening that loosens the tongue of his interlocutor (Schäfer, 2007, February 26, p. 24; our translation).

Chapter Prospectus

Chapter 8, *Listening*, is complementary to our treatment of dialogicality in Chapter 7. The present chapter explicitly emphasizes the active participation of the silent partner or partners in spontaneous spoken discourse at any moment in time. If one assumes that the average number of interlocutors across the board is $n > 2$, then listening predominates over speaking in amount, and often in power. We deliberately refer to this active role of *listening* rather than to the more passive concept of *hearing*. The manner in which one listens can often determine not only the content of discourse, but the direction in which the discourse is moved, or even the very possibility of communication. Indeed, it is no exaggeration to say that listening can change people's lives. Mainstream psycholinguistics has reduced the role of the listener to perception or at best comprehension, both of which remain minimalist with regard to active participation in dialogue.

Listening vs. Hearing

We begin our treatment of listening by returning to the role of the transcriber as listener. We have already insisted in Chapter 4 that the scientific investigation of spoken discourse must proceed by means of analysis of written transcripts, even though the audio recordings remain the primary accessible database; there is no other way to transcend the ephemeral nature of spoken data. But the process of listening to spoken input and writing it down is a very complex one indeed. Listeners with deeply embedded language habits may, in fact, not be accurate processors of acoustic signals when asked to write down what they hear. They have learned to process acoustic input with a view to understanding possible

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_8, © Springer Scienc+Business Media, LLC 2008

message content from their earliest years and from an ambient socio-cultural community that constantly impinges upon them. Hörmann (1986, p. 262 f.) has discussed this phenomenon under the psychological concept of "sense constancy":

> When understanding is "making sense (of something) by placing it in a context," it acquires a constructive aspect: it is more than mere reception. The listener construes a sensible context from what the utterance stimulates and makes possible, from his knowledge of the situation, from his knowledge of the world, and from his motivation.

Thus, language users may find it impossible to prescind from their language habits in their attempt to transcribe exactly what is auditorily presented to their ears.

The first author has opened many a discussion of listening with a brief, informal classroom demonstration, in which students were asked to listen to a short sentence of only 11 syllables distinctly and loudly articulated by him. The sentence was uttered only once, but with the warning that there were no tricks or gimmicks involved and that there was no ulterior purpose beyond finding out what the students perceived and wrote down. The students were asked to write down exactly what they heard, a task quite comparable to that of the transcribers mentioned in Chapters 4 and 10, although without the possibility of repeated hearing. The sentence itself was a variant of a familiar nursery rhyme:

Example 8.1
Peter Piper picked a peck of pickled peppers.

What the professor actually said was:

Example 8.2
Peder Pider peddled priddy *uh* pencils.

What students almost universally wrote down was:

Example 8.3
Peter Piper peddled pretty pencils.

In general, the vowels were transcribed quite correctly, but the reduced or altered consonants were almost entirely missed, and the filler syllable was not transcribed. There is no reason to believe that there was any failure to hear on the part of the students who participated in this demonstration on numerous occasions. But this simple exercise demonstrates the perceptual search for meaning and closure, i.e., Hörmann's "sense constancy," even in such a seemingly trivial task. People did not perform a simple acoustic task in this demonstration, even though they were instructed to do so. They could not avoid actively seeking for an integrated understanding of the sentence uttered by the professor. And that understanding was derived from their everyday usage of language and quite likely even from their early experiences with nursery rhymes. Assuming the role of active listeners rather than passive hearers, the students transcribed the reduced consonants in their unreduced forms: Peder → Peter, Pider → Piper, priddy → pretty; and *uh* was simply omitted. Somehow, what was fed back to the professor

was what the students assumed was *intended* by him, what he meant to say, and therefore what he must have expected in their transcripts. In any event, what they wrote was definitely *not what they heard*. The demonstration incorporates all the elements included by Harley (2001) as recognition, understanding, and comprehension of speech; interestingly enough, he has mentioned listening only in the context of an experimental paradigm that involves listening for mispronunciation of words.

The Listener's Disappearing Act

Throughout the history of modern psychology, the listener has been an unimportant player in the study of discourse, conversation, speaking, communication, and language use in general. An extreme case of this neglect of the listener can be noted in the German textbook on *Sprachpsychologie* by Langenmayr (1997): Chapter 5 (pp. 251–540), entitled *Der Sprecher* (The Speaker), is 290 pages long, whereas the following Chapter 6 (pp. 541–576), entitled *Der Angesprochene – Sprachwahrnehmung und -verarbeitung* (The One Spoken to – Speech Perception and Processing), is only 36 pages long; and his use of the term *Der Angesprochene* defines the role of the listener precisely in passive terms: The One Spoken to. But his emphasis on "the active role of the hearer" (p. 546; our translation), which he refers largely to Hörmann, seems quite paradoxical, given his chapter title and the brevity of his treatment.

In fact, listening is a clear instance of language use, but it has only rarely been investigated. Instead, the speaker has been made to define the research situation almost uniquely. The tradition of turn-taking has typically defined turns only in terms of the speaker. In order to avoid such "a speaker-oriented bias," Linell (1995, p. 208, Note 3) has substituted for S (speaker) and L (listener) the designations "A for the interactant whose utterance is in focus at the moment" and "B for A's interlocutor." Transcripts have made minimal allowances for the role of the listener, apart from acknowledging the necessity of having an addressee and transcribing back-channeling on the part of the listener. For example, in his running text, Levinson (1983, p. 327 f.; see our Chapter 16, *Example 16.1*) explicitly ascribed a silence of 0.2 s to a listener ("R's delay"), but in the excerpt of the transcript, no allowance for this ascription was made. What happens during dialogue has instead been measured only in terms of what has been said by the interlocutor whose turn it is. Brown (1995, p. 39 f.) has emphasized the methodological problem entailed in transcriptions that feature the speaker, while making the listener appear quite passive. The paradox seems to lie in the fact that the good listener is characterized by an ability or virtue of refraining from speaking – silence.

Historically, psychologists have latched onto overt behaviors that have shown promise as explanations. The behavior of a listener is neither overtly verbal, nor easily codifiable in some nonverbal dimensionality. The net result

has been that the role of the listener has been demoted and considered unimportant for the developing course of a dialogue – to the advantage of monologism. Psychology, including especially mainstream psycholinguistics, has persistently investigated what was the easiest to investigate: the words, phrases, and sentences overtly uttered by speakers. The result has been the reduction of research on dialogue to research on the successive monological parts of discourse. To put it another way, such monologically inspired research may indeed be capable of getting at how B responds to A subsequently to A's speaking turn, but it is incapable of taking into account what might have been going on very importantly on B's part while she or he was listening. The implicit principle in play here is the following: If our verbal measures cannot get at these covert or at least nonverbal activities on the part of B (and perhaps of C, D, E, F, or an entire audience), then they are unimportant for the psycholinguistic analysis of spontaneous spoken discourse.

This is at best a very dangerous generalization of methodology. For, even apart from the sometimes almost momentary interventions called back-channeling, in which a listener actually turns into the speaker for that moment, there are many ways in which the listener can be extraordinarily active *as listener*. These activities have not been considered of importance, however, precisely *because* they are not linguistic activities and are methodologically difficult to access. It is now high time that these nonlinguistic activities of the listener be acknowledged as an essential and important component of the very structure of the overall speaking/listening situation. This sort of nonlinguistic activity, we will claim, determines a sort of nonsentential syntax, or sequential structuring, of the total speaking/listening situation in important ways. To argue – as the purists do who wish to maintain the autonomy as well as the hegemony of linguistic factors in the structuring and determination of these eminently human activities – that nonlinguistic factors cannot enter into the molding of meaning and understanding, is unacceptable. It negates the very intelligibility that speakers and listeners garner from the nonlinguistic setting in all spontaneous spoken discourse.

There have indeed been efforts – largely outside of mainstream psycholinguistics – to investigate various nonverbal activities of listeners. Duncan and Fiske (1977, p. 43 ff.) have taken into account a number of nonverbal behaviors of the listener as well as of the speaker (e.g., smiles, laughs, gazing, self-adaptors, and postural shifts), and Papaioannou (2003) has investigated applause on the part of audiences at U.S. presidential inaugurals and state-of-the-union addresses (see our Chapter 18 for further details). Some conversation-analytic researchers have included gaze direction (e.g., Goodwin, 1979, p. 106 f.), applause (e.g., Atkinson & Heritage, 1984, p. xv f.), and laughter (e.g., Jefferson, 1984, p. 346 f.; Hutchby & Wooffitt, 1998, p. 84) in transcripts. Our own research on laughter, including the laughter of the listener, is reported in Chapter 17. Finally, Bohle (2007) has emphasized the importance of gestures on the part of listeners for the organization of turn-taking.

Agonistic Listeners?

But the neglect of listeners goes even further. Not only are their back-channelings as well as nonlinguistic activities, such as eye movements, gestures, bodily movements, and stance, often considered as noncontributory to the meaning and understanding of the speaker/listener situation, but they are also considered to be essentially agonistic or antagonistic. For example, Ong (1967, p. 112) has maintained that "sound signals the present use of power," and Jahandarie (1999, p. 70) has in turn interpreted Ong's statement as the basis for the tendency of cultures "to treat words as weapons." In a similar vein, the biased stereotype of the speaker as the more important party can lead to the corresponding stereotype of the listener as in fact not listening, but waiting impatiently while preparing to speak. If one begins from the fact that every utterance is generically motivated by the perspective of the speaker – as we have explicitly done ourselves in this book – then one must be careful to avoid the Scylla and Charybdis of an impossible neutralism on the part of the listener on the one hand and an allegedly unavoidable adversarial role for the listener on the other hand. Neither is inevitable for the listener. Perspective *is* inevitable for the listener, even as it is for the speaker, but perspective does not lead inevitably to either the avoidance behavior of neutralism or the embracing of its opposite in an adversarial role on the part of the listener.

The Transcendence of Listening

Hörmann (1981, p. 302) has written about the listener in terms of transcendence and what he has referred to as "linguistic transparency": "The act by which the hearer 'sees through' the phonemes, syllables, words, and sentences to identify what the speaker 'means.' " The linguistic elements serve to focus the consciousness of the speaker. But it is "the process of understanding going on in the hearer" that shows us "how and to what end this focus is directed." Hörmann's (p. 308) summary of this concept of transcendence has been incorporated into the penultimate paragraph of his *To mean – To understand*:

> Thus the utterance in itself does not convey any information to the hearer; it only guides the hearer in creating the information for himself. That the hearer knows how to do this, and that he is able to follow the instructions built into the utterance, is the outcome of his incessant striving to *make* the world and all events around him fully intelligible. The criteria which the hearer sets for himself in terms of the explicitness and precision of his action depend on the task he is facing.

Such a doctrine of understanding departs radically from mainstream psycholinguistics. It defines the role of the listener as literally transcendent relative to the linguistic elements of the speaker's utterance. The listener is on his own, so to speak, and must determine for himself or herself, in keeping with the *instructions* given by the speaker through the linguistic elements, what he or

she, precisely as listener, wishes to take from the intended message of the speaker. This is clearly an active and creative role, not merely the unpacking of information provided by a speaker. Paradoxically, only such an active and creative listener can possibly *misunderstand* the input of a speaker.

The Ideal Listener

In Chapter 22, the concept of *verbal integrity* is developed as the underlying attitude proper to speakers, listeners, and overhearers alike. For the present, and specifically with regard to listeners themselves, one can note extremes to be avoided and attitudes to be cultivated by listeners. In one sense, it should suffice to insist that the listener simply listen. The listener who repeatedly consults his or her wristwatch, gazes over the shoulder of the speaker to other clusters of conversation, rolls the eyes, raises eyebrows, or eavesdrops on neighboring conversations while listening is hardly to be described as the ideal listener. The ideal listener is open to the message intended by the speaker, searches for the intelligibility of that message, takes the better part, and is basically respectful toward the person who is speaking. Verhaar (1963, p. 111) has summed up these ideals in the term

> sympathetic listener . . . who listens without saying so much as a syllable; he will look at his partner, he may nod now and then, smile or give similar signs of attention, showing that nothing in the world could interest him more than what his partner has to say.

The fact that mainstream psycholinguistics has been so intent upon developing the concept of the ideal speaker, and has not in any similar fashion entertained a notion of ideal listener perhaps tells us a great deal about the biases with which it has been working. Paradoxically, Chomsky (1965, p. 3) himself has mentioned the "ideal speaker/listener," but has not further developed the concept.

Types of Listeners

The notion that listeners constitute a homogeneous body that can be lumped together for research purposes must be rejected. There happens to be a great variety of listeners in real life, depending upon the immediacy of contact, the role and expectations of the listener, the responsibility and possibility of determinative response and action, and the ways in which those responses can be executed (e.g., by applause or booing; see our Chapter 18 for further details). Clark (1996, p. 14) has illustrated some of the listener's basic roles with a sort of Venn diagram, with the eavesdropper at the outermost limit, then the bystander, the side participant, and finally the speaker and the addressee. Of course, any and all of these position holders can be pluralized, including even the speaker (e.g., a choral or indeed any simultaneous utterance on the part of more than one speaker). But even more variation is possible in listeners' roles.

Accordingly, Brown (1995, p. 4) has pointed out a number of ways in which a listener's performance varies dependently upon the number of interlocutors, the newness of the information, and the level of mutual belief.

Active Silence

From a philosophical point of view, Dauenhauer (1980, p. 3 ff.) has developed the concept of silence as active performance. By silence, he intended the silence that is "always connected somehow with discourse" (p. vii). It excludes, therefore, the silence of someone listening to a concert or of someone involved in a baseball game or in manual work. The active silence associated with discourse is always engaged with the ongoing speaking. It is "a phenomenon which is at least primordial with utterance" (p. 5). It can be an *intervening* silence, insofar as it punctuates, defines melodically and rhythmically, and marks "a sequence of sound phrases as 'mine,' 'anyone's,' etc." (p. 8). Then there is *fore-* and *after*-silence, depending on whether it introduces or terminates an utterance (p. 10). And finally, Dauenhauer has claimed that *deep* silence "is at play in utterances of every sort" (p. 16). It includes "the silence of intimates, liturgical silence, and the silence of the to-be-said. This last mode is a kind of 'normative' silence" (p. 16 f.). Deep silence has been presented in a metaphysical, transcendent, even theological frame as reflected in Dauenhauer's (p. 20) citation of Gadamer (1975, p. 17):

> To pass over something does not mean to avert the gaze from something, but to watch it in such a way that rather than knock against it, one slips by it. Thus tact helps one to preserve distance, it avoids the offensive, ... the violation of the intimate sphere of the person.

Dauenhauer's (p. 24) summary has listed the characteristics held in common by all these types of silence:

> Silence is an active human performance which always appears in connection with an utterance, ... silence is never an act of unmitigated autonomy. Rather, ... silence involves a yielding following upon an awareness of finitude and awe. The yielding involved in silence is peculiar inasmuch as ... it is a yielding which binds and joins.

Dauenhauer has not claimed that these characteristics are all logically independent of one another. However phenomenological they may sound, they constitute a healthy antidote to the mechanistic treatment of silences as simply blocks of downtime. In short, "silence is a complex, positive phenomenon. It is not the mere absence of something else" (p. vii): "Without utterance there can be no silence" (p. 4). According to Pope Paul VI (January 5, 1964), the "lesson of silence" is one that every human must learn; it constitutes a sort of thoughtful forbearance while another speaks and the listener stays with him or her actively (for a detailed empirical discussion of silence, see our Chapter 11).

By way of summary, we find that research engagement of the listener remains today much as it was described more than a quarter of a century ago by Goffman (1981, p. 133): "The relation(s) among speaker, addressed recipient and unaddressed recipient(s) are complicated, significant, and not much explored."

Chapter 9
Punctuation

A king who had to sign the following verdict: "Clemency ill advised, continuation of imprisonment!" felt that this verdict was too harsh as he thought of the future of the prisoner and his family and children. And so, he formulated a pardon by simply moving the comma: "Clemency, ill advised continuation of imprisonment!" (U. Hampel, personal communication, 2006; our translation).

Chapter Prospectus

Chapter 9, *Punctuation*, emphasizes both the historic and current importance of punctuation for understanding written discourse in silent reading and for the reading aloud thereof. What was originally used to signal breath breaks in reading aloud has in the course of time come to signal for the most part division into structural units in both silent reading and reading aloud, thus serving to disambiguate meaning and provide clues for coherence of the written. Punctuation in the written is not an adequate substitute for prosody in spontaneous spoken discourse; punctuation has, therefore, only a remote relationship to prosody. However, in the case of reading aloud, the relationship is much closer. And even in the case of written texts themselves, punctuation can be used rhetorically to simulate spontaneous spoken discourse (e.g., "What the hell!"). As we have discussed in Chapter 4, conventional punctuation marks are frequently used in notation systems for transcription of spoken discourse to indicate phenomena other than in written texts. Such usage may pose problems for readers of transcripts.

Benign Neglect

Over the years, both linguists and psycholinguists have neglected the topic of punctuation. Apparently, very few researchers have considered it important enough to engage empirically. Kainz (1969, p. 218; our translation), an Austrian psychologist of language use, referred to punctuation as "a background phenomenon . . . at the periphery of the language system and not the focus of

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_9, © Springer Science+Business Media, LLC 2008

research." Later on, Waller (1980, p. 247) claimed: "It is hard to find one modern general linguistics textbook that even mentions punctuation." Smith (1982, p. 154) has further contended that "punctuation does a rather poor job of representing how 'speech actually sounds'." O'Connell (1988, p. 134) has in turn suggested that the reason why punctuation does such a poor job may be that it is not the function or job of punctuation to represent "how speech actually sounds." In other words, punctuation does indeed dictate in some fashion pause placement in reading aloud, but to a far lesser degree other aspects of prosodic variation. It is certainly true that periods and question marks signal for someone who is reading aloud the appropriateness of a falling and a rising intonation, respectively, but they do not specify these prosodic elements in detail. The research described in the following will engage this question further.

It should be noted that we will not enter into the research on punctuating text in the process of writing itself. Thorndike (1948, p. 222) had attempted to engage this issue from a psychological perspective and had considered punctuation to be largely a matter of the writer's personal style and "changes in fashion." More recently, Weingarten (n.d.) has studied punctuation in the process of writing in an attempt to get at the cognitive processes involved in text production and has emphasized the syntactic function of punctuation. As a result of his analyses of pause distribution before and after the most frequent punctuation signs, periods and commas, he has concluded that (a) syntactic features of punctuation are primarily responsible for pause distribution; (b) the planning of punctuation occurs simultaneously with verbal planning; and (c) punctuation may indicate a turning point between text revision and text planning.

Syntax, Rhetoric, or Both?

Historically, there have been a number of different viewpoints regarding the functions of punctuation. Honan (1960) has studied the history of punctuation theory in eighteenth and nineteenth century English. He found a massive shift from an elocutionary to a syntactic emphasis at mid-nineteenth century. A possible explanation for this shift has been pointed out by McLuhan (1960, p. 126; cited in Baron, 1981, p. 180). He has contended that editors of Shakespeare added grammatical punctuation in the nineteenth century "to bring out, or hold down his meaning . . . But in Shakespeare's time, punctuation was mainly rhetorical and auditory rather than grammatical." By contrast, Bruthiaux (1993, p. 32), in his historical account of punctuation from medieval times to the twentieth century, has pointed out evidence "of the enduring hold of the rhetorical view of punctuation in both British and American public life" during the nineteenth century. And according to his summary of twentieth century punctuation, it appears that both the rhetorical and the syntactic functions of punctuation still survive. But Quirk, Greenbaum, Leech, and Svartvik (1985, p. 1611), in their *Comprehensive grammar of the English language*, have

stated: "Punctuation practice is governed primarily by grammatical considerations . . . traditional attempts to relate punctuation directly to (in particular) pauses are misguided." For the French language, Védénina (1973) has noted that orthographers emphasize today, much as did Diderot in the eighteenth century, the elocutionary function of punctuation (indicating intonation and pauses). For the German language, Gallmann (1996, p. 1456) has remarked that punctuation functions to segment and/or classify written text. With respect to the comma, however, he has reported that, for various languages, it has repeatedly been proposed to introduce the so-called "Pausenkommatierung" (p. 1641), i.e., the placement of commas where there would be a pause in oral reading. He has mentioned Danish as a language in which this rule is in place, but has remained skeptical as to such an attempt to make spoken and written language more parallel. Clearly, the matter is quite complex and in need of extensive empirical research. Our own research with the English and German versions of Raschka's (1993, 1997; see our Chapter 6) children's story has illustrated the use of punctuation across both languages alike to indicate the expressive dynamic of the story. The syntactic function of the punctuation is minimal in this case simply because there is very little syntactic structure in the text.

What Can Punctuation Tell Us About Reading Aloud?

The custom in classical antiquity was always to read aloud. St. Augustine (ca. 397–400/1960, p. 136; see John K. Ryan's *Notes to Book* 6, Chapter 3, Note 1., p. 385) has given us the *locus classicus* in this regard in his *Confessions*. His description of silent reading on the part of St. Ambrose makes it clear that such reading was most unusual at that time:

> When he read, his eyes moved down the pages and his heart sought out their meaning, while his voice and tongue remained silent. Often when we were present – for no one was forbidden to entry, and it was not his custom to have whoever came announced to him – we saw him reading to himself, and never otherwise. After sitting for a long time in silence – who would dare to annoy a man so occupied? – we would go away. We thought that in that short time which he obtained for refreshing his mind, free from the din of other men's problems, he did not want to be summoned to some other matter. We thought too that perhaps he was afraid, if the author he was reading had expressed things in an obscure manner, then it would be necessary to explain it for some perplexed but eager listener, or to discuss some more difficult questions, and if his time were used up in such tasks, he would be able to read fewer books than he wished to. However, need to save his voice, which easily grew hoarse, was perhaps the more correct reason why he read to himself. But with whatever intention he did it, that man did it for a good purpose. [1]

[1] It was the ancient custom to read out loud and in company with others. Augustine has already given an instance of this when he tells how he and Faustus read together. The present passage is reported to be one of the few descriptions of silent reading in ancient literature. The detail with which Augustine describes St. Ambrose's custom indicates how unusual silent reading must have been.

Originally, then, punctuation was interspersed as notation for breathing. Accordingly, one might well expect that capitalization and punctuation, even now, reliably indicate at least where (breathing) pauses should occur in reading a text aloud. In other words, such a function might be expected to remain part of the basic function of punctuation. Both Akinnaso (1982, p. 105) and Carroll (1986, p. 270), respectively, have asserted this to be a fact: "Commas *signal* pauses" and "Pauses are the oral equivalent of commas."

Currently, punctuation is defined by *Merriam-Webster's collegiate dictionary* (11th ed., 2003, p. 1009) as "the act or practice of inserting standardized marks or signs in written matter to clarify the meaning and separate structural units." There are additional typographical devices that also serve these purposes, e.g., paragraphs, stanzas, the use of white space, blocking, insets, capitalization, and various fonts and sizes of print. In general, all these means are intended to clarify message structure – both syntactically and rhetorically – for a reader and, in reading aloud, for a listener.

But does that mean that punctuation is a good predictor of performance, e.g., regarding the placement and duration of pauses in reading aloud? O'Connell and Kowal (1986, p. 94) have traced such hypotheses over the past 300 years for English, French, German, and Hungarian and found that:

> The only common element in all these is the hypothesis that pauses at periods are longer than pauses at commas; otherwise there is not agreement, either across languages or within a given language.

O'Connell (1988, p. 138 f.) has summarized the research of both O'Connell and Kowal (1986) and Van De Water and O'Connell (1986) in which radio homilies that had been delivered in the German and English languages, respectively, were compared with the manuscripts from which they had been read aloud. The hypotheses regarding durations of pauses at various types of punctuation, along with their actual mean durations as measured instrumentally (in s) in both German and American English are presented in Table 9.1.

The only noteworthy exception in this decreasing sequence of pause durations according to punctuation is a very long duration for the dash (–) in the English homilies. This mean can be largely accounted for by one homilist who clearly used in his written text an unusual number of dashes between sentences as a substitute for periods. As a consequence, this mean is almost identical with both the German and English means for periods ($0.89 < 0.98 < 0.99$ s). Such idiosyncratic use of punctuation is to be expected across writers, as Thorndike

Table 9.1 Hypothesized Relative Magnitudes of Pause Durations at Various Types of Punctuation along with the Actual Measured Mean Durations in Seconds (s) in German and American English Radio Homilies; **null** Refers to Positions without Punctuation

Hypotheses	¶ >	?>	!>	.>	:>	;>	–>	,>	**null**
German	1.65	1.14	1.09	0.98	0.72	0.65	0.59	0.47	0.36
English	1.68	1.31	1.18	0.99	0.72	0.63	0.89	0.55	0.54

(1948) had long ago claimed. Where such variability occurs, the predictability of pause durations in various locations is seriously limited.

Nor are rules for the use of punctuation quite the same across languages. Punctuated locations in German accounted for 91% of the pause time, whereas the same locations accounted for only 76% of the pause time in English in these studies. The reason is straightforward: In German, a comma is required between all clauses within a sentence (e.g., "Ich meine, dass. . ."), but such is not the case in English (e.g., "I think that. . .").

The proportion of pauses at positions marked and unmarked by commas is also of interest. In both the German and the American English, commas were responsible for a high percentage of punctuated locations where a pause did *not* occur (95% and 82%, respectively; see O'Connell, 1988, p. 139). These data should suffice to demonstrate that pauses are not to be thought of as "the oral equivalent of commas" (Carroll, 1986, p. 270). Nor do commas "*signal*" pauses (Akinnaso, 1982, p. 105). In fact, were this equivalence of pause and comma the case, there could logically be no pauses in the **null** positions (i.e., positions at which there was *no* punctuation) of the data given above. But empirically, 14% (205/1433) and 36% (955/2647) of all pauses occurred in **null** positions in German and English, respectively (see O'Connell & Kowal, 1986, and Van De Water & O'Connell, 1986).

The further generalization from research on reading aloud (e.g., Brown & Miron, 1971; Butcher, 1981; Goldman-Eisler, 1972; and Grosjean & Collins, 1979) that use of pauses of various durations indicates an understanding of the sentence structure on the part of the reader is an excellent example of one variety of William James's (1891/1981, p. 195) psychologist's fallacy: "*the assumption that the mental state studied must be conscious of itself as the psychologist is conscious of it*." For the reader himself or herself, the subjective use of punctuation suffices to nuance pausing; no further knowledge of sentential structure is required for this use of pauses. Knobloch (1984, p. 97) has pointed out the same fallacy: The researchers actually infer from their own knowledge of the syntax that the readers who are subjects in their research have the same knowledge as they do. But it does not follow that a reader who consistently pauses at a period understands that he or she has just finished a syntactic unit. The fact that the researcher himself or herself understands that a unit has herewith been completed remains irrelevant to an interpretation of the data with respect to what a reader understands, unless additional empirical evidence regarding his or her inner processing becomes available.

A number of scholars (e.g., Dillon, 1976; Hartmann, 1980; Johnson, 1986; Lehiste, 1984; and Turner & Pöppel, 1983) have hypothesized that pauses are mandatory at the end of the poetic line in reading poetry aloud. O'Connell (1988, p. 151 f.) has reviewed 14 studies in which 289 readings (aloud) by the poets themselves, by dramatic artists, or by ordinary people, were investigated. In this entire corpus, only one poem was read aloud with every poetic line terminated in a pause, and that one case happened to be a poem with only two lines (Nemerov's, 1977 reading of his own *Power to the people*). In fact, across

the board, pauses occurred at only 44% of unpunctuated line-end positions in poetic readings, and the mean duration of these pauses was also less than half the duration of punctuated line-end pauses (0.57 < 1.12 s). Hence, one can hardly refer legitimately to the poetic line as the performance unit of poetry. As a literary genre, poetry is far too complex for such oversimplified generalizations regarding the relationship of the written structure of the poem to the oral performance.

The Written Without Punctuation

We know that the written was originally without punctuation; such a script is formally designated as *scriptio continua* (e.g., Parkes, 1993, p. 9 ff.). What would written text be like today without punctuation? The question is not really that far from reality. William James (1891/1981, p. 233) had introduced the concept of the stream of consciousness in the late nineteenth century:

> Consciousness, then, does not appear to itself chopped up in bits. Such words as "chain" or "train" do not describe it fitly as it presents itself in the first instance. It is nothing jointed; it flows. A "river" or a "stream" are the metaphors by which it is most naturally described. *In talking of it hereafter, let us call it the stream of thought, of consciousness, or of subjective life*.

The last chapter of James Joyce's (1922/1960) *Ulysses*, Molly Bloom's soliloquy, is written almost entirely without paragraphs, capital lettering (except for proper names), and punctuation. These typographical conventions were intended by Joyce to correspond to her stream of consciousness as she reflects on the events of her life. The text thereby becomes extremely difficult to read, whether silently or aloud, since the customary landmarks used "to clarify meaning and separate structural units" (*Merriam-Webster's collegiate dictionary,* 11th ed., 2003, p. 1009) are absent. Joyce's experiment with Molly Bloom also clarifies why James applied the concept of the stream of consciousness to thought, but *not* to either the written or spoken forms of language use: It is precisely the thought process that James has claimed to be a stream "*of consciousness*," not the communicative expression thereof in either the written text or in spoken discourse.

The following example, taken from the *St. Louis Post Dispatch* sports section (February 10, 2007, p. B 7), shows how easily a combination of ellipsis and the omission of a single space in printing can monumentally disrupt the meaning and understanding of a phrase. As printed, the phrase was presented as "the 6-foot-3,220-pound bruiser." The erroneous reading that is almost inevitable diminishes the poor man's height by 3 inches and multiplies his weight by 15 times. The minimal correction needed is, of course, a space after the "3,". But, even with the space, it remains elliptical and could well be further changed to an unambiguous "the 6-feet, 3-inches, 220-pounds bruiser." Were the original phrase to be spoken, the "6-foot-3" portion would inevitably carry its own

prosodic intonational envelope and would be effectively segregated from a following prosodic unit for the "220-pound" portion. In lieu of prosody, the written must carefully supply the proper punctuation. The reader may well note that this discussion assumes that even spacing is an element of punctuation in writing, just as pausing must also be considered an element of prosody in speaking (see Oliveira, 2002).

Other Concepts of Punctuation

Merriam-Webster's collegiate dictionary (11th ed., 2003, p. 1009) has provided a second meaning of "punctuate": "to break into or interrupt at intervals <the steady click of her needles *punctuated* the silence – Edith Wharton>." This is the sense in which Provine (1993, p. 296) has asserted that laughter punctuates speech:

> Laughter punctuates speech. Laughter of both speaker and audience occurred at the end of phrases or sentences. This punctuation effect was extremely strong; a speaker's laughter interrupted speech in less than 1% of laugh episodes. The prominence of this effect is so striking that it may be confirmed by cursory observations of social conversation.

As we shall see in Chapter 17, this punctuation effect is not nearly as strong as Provine has stated above. But his usage points out the possibility of other intrusions as well into spoken discourse, e.g., as in a statement such as "The church bell just outside the window punctuated the pauses in their conversation." For the most part, however, punctuation remains part of the silent world of the written. The only exception to this that we know of is the extraordinarily idiosyncratic system of oral punctuation for spoken discourse devised by the comedian/pianist Victor Borge (1990), his so-called "Phonetic Punctuation." His facetious explanation was that his system would prevent the frequent misunderstandings in oral conversations. He used brief vocalized sounds as intrusions into the speech stream for each of the types of punctuation as he read aloud. The effect was a cacophonous and unusually humorous chain of sounds that truly intruded upon the stream of spoken discourse and hacked it into small pieces. The extraordinary redundancy had the effect of reducing the message itself to background noise – for the sake of the humorous. And in the course of time, this presentation has become one of Borge's most popular routines.

The Problem of Reading Quotation Marks Aloud

According to Gallmann (1996, p. 1464), quotation marks belong to those punctuation signs which do not signal syntactic boundaries but have a classificatory function: They classify a specific part within a text as a text from another author or source ("*Fremdtext*"). When reading quoted material aloud (rather

than speaking spontaneously), the problem arises as to how to indicate prosodically the quoted part of the text because there are no conventional means to do so, except to indicate in words that a segment of text is a quotation (e.g., "quote," "unquote"). Klewitz and Couper-Kuhlen (1999, p. 482 f.) have analyzed spoken conversational discourse in order to ascertain the role of prosody in identifying parts of utterances as reported speech. They found that the following prosodic devices were most frequently used for this purpose: "global pitch (register) and loudness shift, global changes in speech rate and shifts to isochronous timing." As Klewitz and Couper-Kuhlen have concluded from their own research, "prosodic marking . . . is a stylistic device rather than a norm."

In recent years, there have been two very high profile cases in which accusations have been publicly leveled against someone who read his own speech aloud, in which the words of someone else were quoted. In both these cases, the listeners were not individual, but were instead grouped into large audiences. The first instance took place in the German Bundestag on November 10, 1988, the 50th anniversary of the pogrom against the Jews in Germany. The then speaker of the Bundestag, Phillipp Jenninger, used anti-Semitic terminology in his speech. This terminology was clearly noted with quotation marks in his written text. Kerstan and Honsel (1992, p. 248; our translation) have summarized the incident as follows:

> For the reader of the manuscript, it was therefore recognizable that Jenninger was quoting. However, since one cannot hear the quotation marks during the reading of the speech, . . . a misunderstanding on the part of his audience in the Bundestag could arise, namely that Jenninger was presenting this terminology without criticism.

Shortly thereafter, and precisely because of this incident, he was forced to resign his position.

More recently, Pope Benedict XVI gave a speech in German at the University of Regensburg on September 12, 2006 and quoted a medieval text of the Byzantine Emperor Manuel II Paleologos (Vatican translation, p. 2; www.Zenit.com) regarding the ubiquity of violence in Islam:

> He [the emperor] addresses his interlocutor with a startling brusqueness, a brusqueness that we find unacceptable, on the central question about the relationship between religion and violence in general, saying: "Show me just what Mohammed brought that was new, and there you will find things only evil and inhuman, such as his command to spread by the sword the faith he preached."

There was an instant outcry throughout much of the Islamic world and instances of the very violence that had been Benedict's theme.

In both cases, the reactions of many people made it clear that there was a readiness in the audiences that heard either the original (or reports of the original) to interpret the citations as the expression of a personal opinion on the part of the speaker. This readiness itself was perhaps a sign of bad will and/or of editorial and political irresponsibility on the part of those who interpreted the speeches in this fashion. Nonetheless, both instances point out how crucial it is for one who is reading aloud to be as clear – and emphatic – as possible that a

quotation is not to be understood as the opinion of the reader himself or herself. Both these instances occurred in the use of the German language, where it is customary to begin a citation with the verbalization of *Zitat* (quote) and end it with *Zitatende* (end of quote). In the present instances, the written punctuation was quite adequate; but the oral expression left room – at least for the witch hunter – for misinterpretation. Whether it is possible, through the implementation of appropriate prosodic means, to avoid such embarrassments when there is political eagerness on the part of listeners to catch someone in an infelicitous utterance is quite doubtful, and certainly in need of further research. The subtle power of punctuation and its necessary use "to clarify the meaning and separate structural units" can indeed be earth shaking. The reader may well note that, in spontaneous spoken discourse, ambiguity and misunderstanding are frequently avoided in such cases by the use of an expression such as "and then she said." In fact, in the instance noted above on the part of Pope Benedict XVI, he explicitly introduced the citation of Emperor Manuel II Paleologos by saying "He. . ., saying." In the interval represented here by the triple ellipse, he had clearly stated his disapproval of the emperor's words, but this did not suffice to prevent an outcry on the part of extremists who had an axe to grind.

At this point, a number of considerations are in order that take us back to what we have written regarding intentionality in Chapter 6 and about listening in Chapter 8. The matter of intentionality concerns both the speakers and the listeners. One might well ask whether the inclusion of the respective parts of the speeches on the part of Jenninger and Pope Benedict was a prudent act or carelessness in view of foreseeable reactions. On the part of the audience, the creativity we have described in Chapter 8 now shows itself capable of abuse – the possibility of deliberate and exaggerated misinterpretation that furthers the listener's own personal goals at the cost of integrity. We might note too that our very discussion here moves the putatively monological situations of both cases into a dialogical setting. The reactions to both speeches can be looked upon as delayed audience reactions that carry the speeches beyond their temporal and spatial limitations throughout a veritably world audience – for good or ill.

Thus, our four foundational Chapters (5–8) are shown in these two cases to be quite relevant to the rhetorical, intentional, monological/dialogical, and listening aspects of our empirical data. The present chapter itself, largely dedicated as it is to reading aloud rather than to spontaneous spoken discourse, might seem to be quite far removed from such considerations, but as we have found, it is not at all. Public speeches cannot remain within a monological ambit; they are instant dialogical fodder for everyone from theologians to paparazzi.

Chapter 10
Transcription

To describe every Particular, and to relate the whole Conversation of the ensuing Scene, is not within my Power, unless I had forty Pens, and could, at once, write with them all together, as the Company now spoke (Fielding, 1749/1974, p. 377; cited in Meise, 1996, p. 28).

Chapter Prospectus

Chapter 10, *Transcription*, is dedicated to the production and use of written transcripts for the research analysis of spontaneous and reproductive (i.e., reading aloud) spoken discourse. More specifically, we wish herein to consider the preparation, the use, and the reproduction of transcripts, all as types of language use in their own right. Our own research regarding these various forms of language use, particularly with regard to specific problems and biases of transcribers, the question of standardization of notation systems for transcription, and the subsequent reproduction of transcripts in research publications is reviewed. The need for tailoring notation systems to specific research goals is emphasized once again.

The Transcriber as Language User

In Chapter 4, *The Written*, we have outlined a set of principles that we consider fundamental to the design of notation systems for the transcription of spoken discourse for research purposes. We now turn to the transcriber as the change agent involved in this process of transforming spoken discourse into written text. Transcribing is thus to be considered a type of language use on the part of the transcriber.

The production of a transcript from recorded speech depends upon the intentions, abilities, and attention of the transcriber. He or she can produce a transcript that is in accord with the utterances spoken in a given corpus or a transcript in which – deliberately or involuntarily – utterances or parts of

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_10, © Springer Science + Business Media, LLC 2008

utterances are deleted, added, substituted, and/or relocated. Since these decisions are not always a matter of error, we have chosen to speak of changes (specifically, *deletions, additions*, *substitutions*, and *relocations*) rather than errors on the part of the transcriber. It should be noted that, in our own research, we have concentrated on the transcription of verbal and temporal components of utterances; other prosodic and nonverbal components have not been taken into account. Changes are frequently incorporated into transcripts deliberately or at least out of some specific, though often implicit bias. As noted above, the influence of such biases on the part of the transcriber has led Ochs (1979, p. 71) to the assertion that "transcription is theory."

Some Transcriber Difficulties and Biases

O'Connell and Kowal (1994) have analyzed six heterogeneous corpora of spoken discourse in the German language by comparing the original audio recordings with their respective transcripts. In other words, we did not request the production of a transcript by experimental subjects as part of the research in this instance, but rather analyzed transcripts made for other purposes, on other occasions, and by other researchers. These transcripts were compared with a set of master transcripts prepared by ourselves from the original audio recordings. And since we are subject to the same limitations and biases as are all transcribers, the master transcripts were prepared as follows: Both authors listened to the spoken discourse separately. The procedure was off-line in the sense that we listened to a passage again and again until both of us were certain as to how to transcribe it. Sometimes this required that both of us eventually had to listen together to a passage before a final decision was made. In indecipherable cases, the doubtful syllables were entered into the transcript only as a parenthesis marked with a number of syllables, e.g., (4 syl).

The first challenge to be met by the transcriber is the type of spoken corpus to be transcribed. The simplest task is the preparation of a transcript of reproductive spoken discourse, i.e., of the reading aloud of a text. The baseline is obviously the text that is read aloud. O'Connell and Kowal (1994) did not include such an extreme case, in which the number of expected changes is always relatively small (although a third-grade youngster might have a huge number for a simple text). We began instead with parliamentary transcription in which a perfectly well-formed, archival transcript was the desired product. It should be added that, in this instance, we did not make the master transcripts, but used the published ones. Such spoken corpora are easy to transcribe insofar as they are typically produced by very articulate speakers in a setting in which rhetoric is very important. However, the process of transcribing can still be made quite difficult by an – either antagonistically or approvingly – intrusive audience with their interruptions and brief commentaries. At the other extreme of our heterogeneous corpora was a rapid-fire conversation (i.e., an articulation rate of 6.16

syl/s) in colloquial German engaged by two college students. This corpus also included overlapping passages, laughter, and extraneous noise. Needless to say, the former corpus should be much easier to transcribe than the latter. One very useful index for these corpora turns out to be mean number of syllables per change (syl/change). For example, if a speaker actually said, "In the *uh* four years before the *uh* reunification, several things happened," and the transcript read "In the four years before the reunification, several things happened," then syl/change = 20/2 = 10. In other words, a change was made on the average every 10 syllables. For these two corpora, respectively, the mean number of syllables per change in the original transcripts in comparison with the master transcripts was 13 < 17. It should be noted that the lower index of syl/change actually reflects more changes than a higher index. In the present instance, the 13 syl/change reflects a higher number of changes due to the transcribers' goal of obtaining well-formed sentences for the publication of the parliamentary record; the 17 syl/change reflects the fact that the students who transcribed their own audio recorded conversation were intent upon transcribing as accurately as possible. In this instance, our finding pinpoints the salient importance of the transcribers' motivation and specific purpose in comparison with the complexity of the audio source to be transcribed.

The broadest general conclusion to be derived from our research on the six German corpora is that "*transcribers introduce verbal changes in corpora of spoken discourse*" (p. 139). Across the board, the numerousness of the various types of changes in these corpora involving a total of 1558 changes overall was as follows: deletions > additions > substitutions > relocations (655/1558 [42%] > 534/1558 [34%] > 282/1558 [18%] > 87/1558 [6%]). The percentage of originally spoken syllables actually transcribed varied from 82% to 100% (M = 93%); the lowest percentage of transcribed syllables was also that of the transcriber with the largest percentage of deletions (71%), who indulged in the self-instruction to correct the spoken corpus by omitting erroneous German expressions and hesitations. The most common deletions across the board were *und* and *auch* > *äh* > *also* (161 > 144 > 36); the most common additions were *is(t)* > *nich(t)* > *(ei)n(e-)* > *und* and *auch* (88 > 53 > 46 > 20). All six spoken corpora included fillers (*äh*), but the transcribers whose goal was a transcript of well-formed sentences transcribed none of them. The elision *is'* was nearly always transcribed as *ist*, but *ist* was never transcribed in the elided form. Only the college students, who had been specifically instructed to produce an exact transcript, transcribed *is'* in all cases as it had been spoken. Only one corpus was transcribed in accordance with a formal system of notation; this transcript and that of the college students were the only ones without relocations. In summary, even this first project has eloquently manifested that transcription is an extraordinarily complex instance of language use that depends on many different factors, including the intention and ability of the transcribers, the speech genre, and the quality of the audio source.

The reader may note that individual changes involved for the most part short function words. The danger exists that the numerousness of these hide the sometimes quite substantive changes made in content words as well. The latter

changes were usually occasioned by characteristics of the audio source: the presence of extraneous noise, unclear pronunciation on the part of the speaker, or poor acoustic quality of the recordings. In medical, legal, and emergency settings, such changes can alter the meaning of a transcript so as to do great harm. Walker (1986, p. 209) has mentioned such a case from a court transcript in which the spoken designation "male in extremis" was changed in the transcript to "male, an extremist." Suffice it to say that the legal consequences for a gentleman at the point of death are most likely nonexistent, but those to be exacted by the court against an extremist might well involve years of imprisonment.

Slips of the Ear

Ferber (1991) has argued that there is no way of validating most of the collections in the archival literature of slips of the tongue, insofar as they have been collected mostly from memory, without the assistance of audio recordings. Accordingly, she set out to ascertain empirically whether slips of the tongue are not really slips of the ear, i.e., "incorrect transcription" (p. 106). For example, students who hear an isolated "oth" as in "other," nearly always transcribe "of". Ferber found that "no slip was recorded by all four [of her] listeners" (p. 119), and she concluded that "the only way of collecting spontaneous slips would seem to be by means of tape recordings, which should be listened to repeatedly, preferably by more than one person" (p. 120): The on-line listeners "*recorded only about one-third as many slips as were detected by repeated listening, and, even so, about half the items noted as slips proved erroneous*" (p. 105). In this context, on-line refers to an uninterrupted playing of the recorded speech, whereas off-line refers to the opportunity to playback any portion of the recorded speech at will.

Taking their cue from Ferber and from Lindsay (1988), Lindsay and O'Connell (1995, p. 101) had four undergraduate volunteers transcribe an audio-taped interview of former president Ronald Reagan with Dan Rather. Their instructions were simply to transcribe the tape-recorded interview from a single playing; stopping the taped recording was allowed, but no repetition or replay. Thereafter, two of the experimental subjects repeated the transcription on-line, and the two others repeated it off-line. Lindsay and O'Connell have summarized their results as follows:

> *None produced a verbatim transcription, but all preserved semantic content quite well. Still, deletions were numerous, particularly of discourse markers and hesitation phenomena, both of which characterize spoken, not written discourse. Significantly more deletions in the on-line than in the off-line condition indicated the difficulty of audiotape processing without off-line replay.*

The differences occasioned by an on-line vs. off-line method of transcription are clearly of considerable magnitude: The on-line method cannot be recommended as an appropriate research methodology for transcribing audio recordings.

The cumulative evidence does indeed appear to indicate that much of what had been presented as slips of the tongue really constitute slips of the ear, i.e., errors made by transcribers. Hence, Bock (1996, p. 405) has referred to the identification of slips of the tongue in the literature as "abysmal," even though they have largely been detected by "trained listeners."

Some Limitations of Transcripts

Brown (1995, p. 39 f.) has pinpointed the considerable loss of information about the behavior of listeners when a conversation is transcribed, because the transcript does not contain the interlocutors' reading of the face and movements of listeners:

> The very nature of transcription conventions concentrates on the speaker and what the speaker is doing while uttering, leading us readily to a view of the active speaker and a listener who is quite passive during the speaker's turn. But collaborative conversation does not consist of a series of discrete stages, as the physical nature of the transcription suggests, with a participant either being actively on-stage or passively off-stage. From each participant's point of view, that participant is constantly on-stage but playing different roles, which overlap and merge into each other.

Our present discussion, then, goes beyond the limitations of abilities and purposes on the part of the transcriber, and even beyond the complexity of the acoustic signal and its setting. *Transcription itself is a limited and defective device*. Even the simplest of spoken discourse involves an unlimited richness of analyzable facets. There is no notation system that is in principle capable of embracing altogether this virtually infinite richness. Abercrombie (1967, p. 114) has expressed this virtual infinity quite bluntly: "It is impossible to give a truly complete description of a segment." Furthermore, the rote addition of elements in a transcript simply leads to a cumbersome transcript that is itself not analyzable or even legible in any practical way: The seen/read simply cannot adequately depict the spoken/heard. An extreme example of this outcome is Pike's (1943, p. 155) 88-character description of [o].

In transcribing, more is not necessarily better. One can pick up at random a current issue of a journal in the language sciences and find there transcripts bristling with various notations: idiosyncratic orthography, diacritical marks, conventional punctuation marks used in some idiosyncratic way, multiplication of graphemes to indicate a variety of phenomena, along with a multitude of other symbols. Many such notations neither serve the user-friendly function of allowing the journal reader to process the passage intelligibly nor do they enter into any kind of analysis of the passage. In other words, they seem to be made for show; they make the presentation appear more technical, more scientific. This is not science. The most extreme example of this sort of over-transcription that we have found to date is a 356-page book by Dorval (1990) of which 40% is dedicated to transcripts and transcript notations – without any inferential

argumentation whatsoever. His appendix of 75 pages consists entirely of transcripts, with the instruction to the reader that "they should be used for illustrative purposes only" (p. 276).

Since transcripts are tools for analysis and intelligibility, not cosmetic devices, they should include only what is relevant for a given research project. Hence, the call for a standardized notation system for transcribing (e.g., Edwards's, 1989, 1993, p. 141 ff., "field-wide standard" and MacWhinney's, 1995, p. 1, "sharing of data"; see also Selting, Auer, Barden, Bergmann, Couper-Kuhlen, Günthner, Meier, Quasthoff, Schlobinski, & Uhmann, 1998, p. 91) must be challenged. MacWhinney (1995) has been the most explicit regarding the necessity for "a standardized system for data transcription and analysis" (p. 2). Indeed, we do need guidelines to maximize compatibility and comparability from one project to another. But for this purpose, a single, standardized notation system is neither practical nor scientifically heuristic. Sinclair (1995) has put it nicely. We do not need "parading in front of us these incomprehensible stretches of mumbo jumbo" (p. 107), but some common sense: "Avoid interfering with the plain text" (p. 109).

In summary, one might readily agree that simple phoneme/grapheme correspondence is an acceptable form of standardization in transcript notation, but the effort to standardize the entire notation system is ultimately inappropriate, even impossible. Transcribing the virtually infinite richness of even a simple spoken corpus is pie-in-the-sky science.

Reproduction of Transcripts for Research Purposes

One application of transcription research that exemplifies yet another form of language use – reproduction of transcripts for research purposes – manifests very clearly many of the problematic aspects of this domain. Specifically, excerpts of transcripts are frequently reproduced in publications subsequent to the original publication, both to contribute to further research endeavors and to instruct colleagues in the research applications of such transcripts. In both cases, the importance of accuracy is paramount. An indication of how frequently this sort of reproduction occurs can be found in Levinson (1983, pp. 284–370), where, in a single chapter on "Conversational Structure," 124 such excerpts have been reproduced.

Discrepancies between the original and the reproduced transcript – in terms of our standard set of changes, including deletions, additions, substitutions, and relocations – are indicative that something is amiss in this application of a notation system. It was precisely the discovery of these discrepancies that led O'Connell and Kowal (2000) to a more systematic investigation of such reproductions in order to discover empirically whether the incidence of discrepancies was inordinately high. In order to assemble not only a representative corpus but one that exemplified the highest quality, we chose 10 excerpts from prominent textbooks (Duranti, 1997; Garman, 1990; Whitney, 1998), 10 excerpts from Levinson (1983), and six versions of a single German transcript from Keppler (1987, p. 291).

No reproduced excerpt that we examined was found to be without at least one change – in a feature relevant to the notation system – by comparison to the originally published excerpt. In terms of numerousness of changes across the board, the 308 changes were distributed according to the following frequencies: format > prosodic > verbal > extralinguistic > paralinguistic: (131/308 [42%] > 91/308 [30%] > 77/308 [25%] > 9/308 [3%] > 0/308 [0%]). And in terms of types of change, frequencies were distributed similarly to the distributions found for original transcription in O'Connell and Kowal (1994), except that substitutions were proportionately more frequent than additions: deletions > substitutions > additions > relocations (again, of the 308 changes, 113/308 [37%] > 111/308 [36%] > 72/308 [23%] > 12/308 [4%]). In summary terms, "the overall rate of change is 6.6 syllables per change (2032/308) across 41 comparisons" (O'Connell & Kowal, 2000, p. 247) of originally published excerpts with reproduced excerpts, i.e., some change was made roughly every seven syllables in this corpus.

At the risk of presenting even more errors of transcript reproduction through the process of printing this book, we offer the following comparison of an original excerpt of a transcript from Schegloff (1979, p. 52) and the reproduced excerpt as it appeared in Levinson (1983, p. 344):

Example 10.1

The Original Transcript	*The Reproduced Transcript*
I: Hello:,	R: Hello:,
→B: H'llo Ilse ?	→C: Hello *Il*se?
→I: Yes. <u>B</u>e:tty.	R: Yes. *Be* :tty.

Without the inclusion of changes involving the name initials, the arrows, and the underlining in the original, there are still five changes from the original 8-syllable excerpt to the reproduced excerpt: (1) H'llo → Hello; (2) Ilse → *Il*se; (3) Ilse ? → *Il*se?; (4) <u>B</u>e→ *Be*; and (5) <u>B</u>e:tty. → *Be* :tty. Changes (1), (2), and (4) introduce prosodic changes in the reproduced excerpt; changes (3) and (5) introduce changes in spacing in the reproduced excerpt. Change (4) requires some further explanation. The change from <u>B</u> to *B* is not considered a change insofar as underlining was previously the common notation for italics; however, the change from e to *e* involves a prosodically meaningful shift in the conversation-analytic transcript notation system.

Levinson has provided no explanation or justification for any of these changes. In addition, the changes in spacing have not been explained in the appendix to his Chapter 6 (p. 369 f.) where the details of the notation system are listed. The lack of commentary seems paradoxical in view of Levinson's own claim that in conversation-analytic research "heavy reliance inevitably comes to be placed on transcriptions" (p. 295).

Clayman and Heritage's (2002) reproductions of excerpts from transcripts constitute a special case: In this instance, the authors have reproduced their own original excerpts in the same volume. O'Connell and Kowal (2006b, p. 160) have summarized the evidence regarding Clayman and Heritage's reproductions:

> Overall, of the 55 identical or partially overlapping excerpts used by Clayman and Heritage for empirical argumentation, 31 (56.4%) involved one or more (sometimes numerous) erroneous changes.

O'Connell and Kowal have suggested that the same excerpt had perhaps been transcribed in these instances by different assistants without any effort to compare the variant versions.

Conversation-analytic researchers have insisted that "the transcript plays a central role in research on spoken discourse" (Edwards, 1993, p. 3; see also Psathas & Anderson, 1990, p. 76 f.), but our empirical analyses have indicated that both the validity and the reliability of reproduced transcripts may be quite low.

The Diagnosis

It is hardly overdrawn to refer to the high rate of changes in reproduced excerpts of transcripts as disconcerting. The usefulness of such defective reproductions is thereby considerably reduced. Kitzinger (1998), who carried out similar research, has ascribed the phenomenon to simple carelessness on the part of the scholars in question. We find this diagnosis too harsh for a number of reasons. The materials themselves constitute a formidable challenge. They are dense, unfamiliar, and remote; their reproduction is a task that violates many of the language habits and expectations of a native speaker who is dealing with the written reproduction of an already published transcript. For example, the presence of the item *stors* in a transcript excerpt could be a misspelling of *stores* or a correctly transcribed mispronunciation (characteristic of the St. Louis, MO region) of *stars*. Many such minute instances add up to a complexity that overloads the human processor. But the identification of the specific human processor responsible for defective reproductions of transcripts is extraordinarily difficult because any given instance of a reproduced transcript goes through a very complex series of stages: A scholar prepares a manuscript, which is then typeset, edited, proofread, and finally printed. Where in this sequence the changes are inserted is itself an empirical question. We ourselves have found in the process of publishing journal articles that page proofs not infrequently contain many changes in excerpts from transcripts. And in those instances in which authors do not receive page proofs, there is no recourse short of the subsequent publication of an erratum in a later issue of the journal in question. In other words, the problem should be acknowledged as an important and real one that is not entirely traceable. Accordingly, extreme caution is needed in the use of reproduced transcripts.

Currently, there are no notation systems for the transcription of spoken discourse that are truly user-friendly and efficient. Schenkein's (1978, p. xi) goal of producing "a reader's transcript – one that will look to the eye how it sounds to the ear," has not been realized in the intervening three decades, simply because it is not possible. The habits associated with the learning of our native

language do not include the reading of complexly notated transcripts. The evidence presented by O'Connell and Kowal (2000, p. 266) seriously challenges "the practical usability of current notation systems" in research publications. Their suggestion warrants the reader's attention. It is that:

> Henceforth researchers transcribe spoken discourse with *only* those notations which are to be used for analyses in keeping with the purposes of the research. The resulting transcripts will be less dense and hence easier to reproduce – and an appropriate level of parsimony will be preserved.

In summary, our research on the preparation of transcripts of spontaneous spoken discourse has shown that it is a very complex type of language use: The transcriber's own rhetorical habits, his or her intentions as regard the specific task of transcribing, immersion in the dialogical, and ability to listen carefully all influence the product in important ways. The same complexity applies as well to the reproducer and the reader of transcript excerpts.

Chapter 11
Pauses

We kept quiet for quite a while like a pair of old friends who aren't troubled by silence falling between them, feeling it might be empty; on the contrary, we considered the silence, at least as far as I was concerned, the most eloquent form of conversation (Pamuk, 1997, p. 226; transl. G. Gün).

Chapter Prospectus

Chapter 11, *Pauses*, introduces a long tradition of research on (silent or unfilled) pauses, much of it based on the concept of hesitation mentioned in Chapter 3. In such a tradition, the multi-determination of pauses, including their physiological necessity for breathing as well as their pragmatic and rhetorical functions for the communicative purposes of spoken discourse, has often been neglected. For the most part, these omissions have left only a cognitive interpretation of pauses viable in the psycholinguistic tradition; by contrast, conversation-analytic research has emphasized the communicative impact of pauses in turn-taking. There are also problems with the assessment of pauses, since perceptual rather than instrumental methods have often been used to identify them, and in many cases in which instrumental measurements have been used, a variety of cut-off points for minimal duration have been adopted. These inconsistencies in the assessment and operational definitions of pauses significantly reduce the comparability of data from various empirical investigations. Hence, the literature is difficult to interpret and much of it has yielded a sort of mythology regarding the temporal organization of spontaneous spoken discourse. Although some researchers have wished to distinguish silences from pauses, in the context of spontaneous spoken discourse, we consider them synonymous. We close the chapter with an overview of our own research on pauses.

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_11, © Springer Science + Business Media, LLC 2008

On Again, Off Again

On the face of it, nothing seems quite as simple as the segregation of on-time from off-time in speaking; at any given time, one is either speaking or not speaking, and the intervals should not be difficult to identify. But nothing could be further from the facts. Among the components of *on-time* that have been categorized as *off-time* in pause research are the following:

(1) Goldman-Eisler (1961, p. 167) has classified "irrelevant vocal productions, i.e., noise, such as repetitions of the same words or other obvious forms of marking time vocally" as off-time.
(2) Maclay and Osgood (1959, p. 24) have included syllabic prolongations or "non-phonemic lengthening of phonemes" in the category of off-time.
(3) Without explicitly relegating pauses to off-time, but rather leaving the matter unspecified, Perfetti and Papi (1985, p. 345) have included under pauses any interruption of the informative process, thus including not only silences, but also hesitations, repetitions, corrections, and paraphrases which, they have assumed, have almost no informative value.
(4) Clark (1996, p. 259) has introduced the concept of "the point of suspension," where "speakers cease their presentation." Such a hiatus or cessation or suspension of speaking includes "{}, {*uh*}, {*I mean*}, and {*uh you know*}" (Clark & Fox Tree, 2002, p. 82, [2']). But in none of these examples is there any silence involved; there is no pausing, no cessation of speaking. Even Clark and Fox Tree's empty brackets do not indicate an actual pause.

We are left with a research tradition that confuses the difference between speaking and silence.

The confounding of on-time components with off-time components, along with the assumption that the latter are without informative value, is most unusual. It is as if certain verbal elements, typically those importantly characteristic of spontaneous spoken discourse, e.g., the discourse markers singled out by Schiffrin (1987), have been considered sometimes verbal and sometimes not – to our way of thinking, one more consequence of the written language bias. A more unbiased analysis of verbal components of spoken discourse must start with the assumption that *all of them* are uttered for some reason and contribute to the overall interaction. The inconsistent operationalizations of *pause* are the foundation for non-comparable datasets and must be taken into account in reviewing the literature in terms of the overall temporal organization of speaking.

Basically, the organization is as follows: on-time + off-time = total time. *Speech rate* is calculated in terms of syllables per second (syl/s) of total time, whereas *articulation rate* is calculated in terms of syl/s of on-time only. It is clear that any transfer of on-time to off-time affects the calculation of articulation rate. The following fictitious example of a transcript will demonstrate this change:

Example 11.1
he he he w- went uh immediately homeward after the game

Assume that the supernumerary syllables (*he he w-* and *uh*) sum to 1 s in duration and that the total duration of the spoken sentence (13 syl) is 3 s. Then the transfer of these four syllables to off-time changes the articulation rate from 13/3 = 4.33 syl/s to 13/2 = 6.5 syl/s or, excluding the four syllables, 9/2 = 4.5 syl/s. The first of these two calculations shifts the articulation rate from moderately low to quite high and renders the comparability among datasets with respect to this variable thoroughly void. And even for this fictitious example which includes no pauses, the transfer to off-time inflates the percentage of off-time within the sentence from zero to 33%. It was precisely through addition of these pseudo off-times to genuine off-time that has led to the perpetuation of Goldman-Eisler's mythology of 1961 in the twenty-first century: "50 per cent of a person's speaking time is made up of silence" (Erard, 2004, January 2, p. A 13; see also Zuo, 2002, p. 2). The dimensions of on-time and off-time discussed in this paragraph, then, constitute the basic units to be investigated.

The History of Pause Research: Pausology

The word *pausology* was introduced by Tosi (1965). It was adopted by O'Connell and Kowal (1983) for their review of the research literature regarding silent pauses from the beginnings of modern times up to 1980. They began their review by a preliminary acknowledgement of the complex multi-determination of pauses:

> Pauses are determined by breathing, embarrassment, weariness, anxiety, confusion, anger, interruption, pain, syntactic complexity, mendacity, availability of lexical items, emphasis, boredom, and a host of other situational, organismic, intersubjective, linguistic, and conventional factors. (p. 222)

The array appears quite daunting for anyone who wishes to investigate the functions of silent pauses in human communication. The fact that pauses in and of themselves are the absence of speaking makes their functional interpretation completely dependent upon the context in which they occur.

Quintilian (ca. 95/1958) and, much later, many elocutionists, e.g., Sheridan (1787), have emphasized the rhetorical power of pauses. Another elocutionist, Walker (1811), has even listed places where one should pause. A well-known literary reference is that of von Kleist (1806/undated, p. 976; our translation) in the context of "tricks to extend the discourse in order to win the requisite time to construct my ideas in the workshop of my mind."

The first empirical study in modern times was that of Cattell (1885, 1886). He found that the speech rate of reading words aloud as rapidly as possible was an inverse function of the familiarity of the speaker (Cattell himself; $n = 1$) with the language of the words; i.e., the time required for reading such a list of presumably single-syllable words aloud was shorter in milliseconds (ms) for more familiar languages: "English 138, French 167, German 250, Italian 327,

Latin 434, and Greek 484" (1886, p. 65). He did not measure speech rate in units of syl/s and did not segregate off-time from on-time at all. But he did measure actual duration in ms. The experiment has been described in detail precisely to indicate to the reader how complex such a simple finding is when it is dissected in terms of methodology. Cattell (1886, p. 63) himself has put his finger on two chronic problems in such research:

> The physical apparatus used seldom produces the stimulus in a satisfactory manner or measures the time with entire accuracy, and must be so delicate and complicated that it requires the greatest care to operate with it and keep it in order. The other difficulty lies in the fact that the times measured are artificial, not corresponding to the times taken up by mental processes in our ordinary life.

It is gratifying to read Cattell's insight regarding the artificiality involved in using isolated words. One might add another recurrent problem in such research: Experimenters seem to have inordinate difficulty with the conceptualization of short temporal intervals, about which they have no intuitive practical experience. For example, Crystal and Quirk (1964, p. 49), while noting that "length of silent pause is its relevant gradient characteristic," have couched their discussion in terms of a unit never realistically considered relevant to pause measurement, namely the microsecond, i.e., one millionth of a second. Similarly, Tannen (1986, p. 48 f.) has claimed "enormous impact" for "tiny little things like microseconds of pause"; Gee (1986, p. 418), by simply misplacing a decimal, has made much of a pause of less than one-thousandth of a second: "children (0.960 ms) shouting"; and Glukhov (1975; cited in O'Connell, 1988, p. 111) has also systematically misplaced his decimal points.

Throughout the first third of the twentieth century, efforts to improve instrumental measurement of silent pauses were engaged in more than a dozen studies. Newman and Mather (1938) were the first to use subjective judgments to identify pauses. Hahn (1949, p. 338) made such subjective judgments her norm: "To judge the vocal and articulatory aspects of speech, one cannot set up objective measures." By mid-twentieth century, there was still no sign of an integrative theory of silent pause usage.

Because of its appearance in Osgood and Sebeok's (1954) *Psycholinguistics*, Lounsbury's (1954) views have assumed an inordinate importance. More specifically, he has introduced considerable confusion and controversy because of his distinction of juncture and hesitation pauses on the basis of syntax, speaker and/or hearer purposes, and duration (see O'Connell & Kowal, 1983, p. 233). Maclay and Osgood (1959) took their cue from Lounsbury. Their study had far-reaching influence by stimulating subsequent research, but its reliance on a subjective method of identifying silent pauses and its very small and selective corpus made the results minimally significant. Goldman-Eisler's research has also been very influential over the years. Her own summary of her major studies was published in 1968, and has served as a starting point for our own research on the temporal organization of spoken discourse. It should be noted, however, that her research, along with a number of others' (e.g., Jaffe & Feldstein, 1970;

Mahl, 1956), began with a preoccupation with pathological speech and with clinical descriptions of the mentally ill. This preoccupation has encouraged the development of conceptualizations of pauses in terms of hesitation, disfluency, and inefficiency. Given a boost from this preoccupation with pathological speech, it was a small step, aided and abetted by the WLB, to ascribe deficiency to *all* use of hesitations in normal speech.

O'Connell and Kowal (1983, p. 272 ff.) have summarized the entire period from the beginnings in modern times up to 1980 as a problematic phase of pausology. Some of the necessary factors that were widely neglected in that tradition of research have included multi-determination of silent pauses and their careful instrumental measurement. There has also been an inordinate amount of a-historicity and an a-theoretical approach that leaves much to be desired. In any event, the research endeavors encompassed by the term pausology seem now to be a Zeitgeist that is past, a research tradition that – in view of our concern with spontaneous spoken discourse – we must fault particularly for

> the use of meaningless stimulation materials (e.g., numbers, nonsense syllables, isolated words, isolated sentences, and atypical, fabricated paragraphs), simulation of affect, and overdependence on linguistics. (p. 275 f.)

In the meantime, a more socio-culturally oriented approach to silent pauses and other temporal factors of spoken discourse has become the order of the day. The carefully controlled experiment has had to make room for "more representative designs, naturalistic observation, realistic complex situations, dialectology, and socio-economic levels" (p. 276). Macro-analyses of field-observational data have already proven to be all important. In 1980, our conclusion was that "pausology has [had] a moderately unimpressive past and [can have] a promising future" (p. 276).

Off-time as a Research Problem

The misnomer *off-time* has been very influential in framing attitudes toward research on silent pauses. Off-time connotes passivity on the part of the listener, a time in which he or she is doing nothing, the opposite of on-time activity. For the speaker, off-time seems to allow only covert, internal activity such as cognitive planning and makes no provision for nonverbal activities he or she may be engaged in. The terms *filled* and *silent pauses* are also misnomers: There is no such thing as a filled pause in contrast to an unfilled or silent pause. The so-called fillers do not fill a void; they are not pauses or down-time in any sense of the term (for a more detailed discussion of fillers see our Chapter 13).

The literature on pause research has proven one thing beyond the shadow of a doubt: There is not one single function that pauses fulfill; they are indeed complex in their multi-determination and multiple functions. And all these functions must be taken into account in any explanation of the temporal organization of spoken discourse in real time, including reading aloud as well

as spontaneous speaking. Pauses are lawful components of spoken discourse and are in need of explanation just as much as any of the so-called on-time components. The question for empirical research is: What are the functions that can be fulfilled – for both speaker and listener – by pauses in various locations and of various durations in spoken discourse? Pauses also serve to throw light on verbal phenomena which they precede or follow. For example, the hypothesis of Ameka (1992b, 1994) was not per se about pauses, but emphasized the temporal isolation of interjections by pauses in parallel to their syntactic isolation. In other words, one very important research application of pauses is that their patterning throws light on other phenomena of interest. In the following chapters, many such instances will be discussed, including rhetoric, style, fillers, interjections, turn-taking, laughter, and applause.

More must be added regarding what time we are really concerned within research when we speak of pauses. The tenacity with which the opinion has prevailed that pauses as defined by the perception of a hearer are the phenomena of research interest is noteworthy. Almost 40 years ago, Gülich (1970, p. 277; our translation) expressed it quite forthrightly:

> What the ear of the hearer perceives is therefore critical for our observations, not what instrumentation records. It would be inappropriate to want to correct the subjective auditory perception by means of an objective instrumental recording.

Note that Gülich had in mind not perceptual research on the conditions of such judgments, but the definition of pauses as components of spoken discourse and for the purposes of research on speech production. A vast amount of pause research has been contaminated, indeed invalidated, by convictions like those of Gülich. As we have mentioned above, her subjective definition of pauses had been adopted much earlier by both Newman and Mather (1938) and by Hahn (1949). Perhaps the most well-known and also one of the most extensive corpora of English spoken discourse, the London-Lund corpus (Svartvik & Quirk, 1980), suffers from this limitation but has nonetheless been used for the purpose of arguing on the basis of the pause durations therein (e.g., Clark & Fox Tree, 2002; for a more detailed discussion, see our Chapter 13).

A Benevolent Take-over by Conversation-analytic Researchers

A great deal of the research in the past quarter century that has entailed the analysis of pauses has been sociolinguistic and pragmatic rather than psycholinguistic – unfortunately with little noticeable mutual acknowledgement of one another. And the definition of pause has remained the earlier subjective one (e.g., Hahn, 1949; Gülich, 1970) according to which "relative differences. . . are more important than any such notion as 'absolute' or 'clock-time' differences" (Psathas & Anderson, 1990, p. 87):

> Transcriptionists strive for a rendering that is as close as possible to the experience of those actually participating in the interaction. The transcriptionists' close and repeated listening to the interaction enables her/him to perceive the relative differences in the spaces (pauses, gaps, silence) that occur.

The empirical evidence has been summed up recently (Kowal & O'Connell, 2000, p. 358 and p. 372 ff.; see also Spinos, O'Connell, & Kowal, 2002): Perceptual estimates of duration are neither valid nor reliable. In addition, they cannot reflect "the experience of those actually participating in the interaction," because the repeated listening makes the transcriptionists' experience of the corpus quite idiosyncratic, but in no way closer "to the experience of those actually participating in the interaction." In fact, it is characteristic of participants' experience that it must generally rely on a single pass in real time, not a closely analytic inspection of time. Moreover, the preoccupation of the participating speakers and listeners in dialogue is precisely with the intelligibility of the spoken input. Whereas researchers use the transcript for analytic understanding, interlocutors synthetically pursue meaning and understanding (see also our Chapter 8).

The insistence upon perceptual assessment of pauses can be found also in a recent attempt on the part of Meise (1996, p. 36; our translation) to formulate a typology of silences in dialogue, including various types of pauses. However, her rationale for declining to operate "with exact physical time units" is somewhat different from the traditional argument: She considers it impossible to determine exact physical time units *characteristic of various functional types of pauses*. But, such a correlation of pause functions with pause durations has never been the rationale for insisting on the instrumental measurement of pause time.

The positive contributions of a sociolinguistic orientation to pauses have included a dramatic shift to field-observational research, use of informal, conversational corpora, and emphasis on local management of spontaneous spoken discourse. In conversation-analytic research, the use of qualitative analyses has often been pitted against the inferential statistical tradition of mainstream psycholinguistics. It is certainly true that, in its extreme forms, qualitative analysis can devolve into unabashed anecdote. And at the other extreme, mainstream psycholinguistics has sometimes emphasized inferential significance to the point where the most trivial of results are presented as highly significant. But there is no reason why the two methods cannot be amicably combined in a moderate way, as we have endeavored to do in our research with Suleiman (herself a sociolinguist; e.g., Suleiman, O'Connell, & Kowal, 2002).

The Proper Temporal Dimensionality of Pauses: Measurement

We have already mentioned that microseconds and misplaced decimal points put pauses outside the pale of both human experience and research. The vast majority of pauses in spoken discourse are <1 s in measured duration. In this domain,

the >1.0 s outliers become very important for emphatic, rhetorical, and other special purposes. Any accurate count of numerousness of pauses must also rely on a cut-off point for the minimal duration of measured (and measurable) pauses. Pauses of < 0.1 s in duration are not perceived accurately and reliably by either interlocutors or researchers and cannot be measured accurately and reliably in most corpora and with most extant instrumental methods. On the other hand, they occur relatively frequently and, if included operationally among measured pauses, assume an inordinate importance in the database.

Some pause researchers have used cut-off points much longer than 0.1 s. Goldman-Eisler's (1968, p. 12) original cut-off point was 0.25 s, but some have used cut-off points even longer than 2.0 s (e.g., Siegman, 1979). Even in current research, Yakovleva (2004, p. 134) has mentioned no use of a cut-off point and has distinguished, without further justification, short pauses (up to 3 s), medium pauses (from 3 to 8 s), and long pauses (from 8 s on). Large cut-off points inflate the mean duration of pauses and deflate their frequency of occurrence (see Hieke, Kowal, & O'Connell, 1983). A physically defined cut-off point presumes instrumental measurement. Until the advent of the www.PRAAT.com software, such measurement was cumbersome and expensive, but Boersma's programs (e.g., Boersma & Weenink, 2005) have been a godsend to this area of research. There is no longer any excuse for variations in methods of identifying and measuring pauses.

Pauses are one of the sequential means that speakers have at their disposal to accomplish communication. Just as their on-time cousins such as words, articulation rate, syntax, prosody, and even nonverbal devices that accompany speech, pauses are multi-determined and therefore multi-functional means to reach the goal of effective communication among human beings. And they are also useful to help us understand their on-time cousins as well.

We conclude this section with another very simple example: The young woman who has just received a proposal of marriage – and pauses before answering – may be emotionally overwhelmed, undecided, surprised, shocked, outraged, or even amused. Eventually, she and the young man in question had better make very sure which it is!

Our Own Research on Pauses

Our research on pauses reflects a gradual shift over the years from carefully controlled laboratory experimentation to field observation, which has provided us with access to genuinely communicative language use. And so, the emphasis on cognitive functions has moved gradually to a preoccupation with communication. Meanwhile, the pause methodology has come to be applied as well to a variety of on-time verbal and nonverbal phenomena (e.g., interjections and laughter).

Our engagement of pauses began, as we reported in our Preface to the book, with research in Berlin with Hans Hörmann (O'Connell, Kowal, & Hörmann,

1969, 1970). The emphasis in these studies came from a combination of both disbelief in the Chomskyan position regarding the primacy of syntax that had been adopted by many mainstream psycholinguists, and of our own interest in pauses as a response measure to discover psychological processes in the speaker. This application of pauses had been pioneered by Goldman-Eisler (1968). We did indeed find that, with syntax held constant, semantic factors significantly affected pause usage in both reading a story aloud and in retelling it. Shortly thereafter, we extended the same methodology to a developmental study: O'Connell and Kowal (1972, p. 161) found that, for teenagers as well, "An unusual turn of events noted in the reading of the story literally gives people pause." Both frequency of occurrence and duration of pauses were increased by the unusual turn of events, and adult subjects were more sensitive to this semantic variation than were adolescents.

The developmental preoccupation continued in a study by Kowal, O'Connell, and Sabin (1975), in which youngsters from 5 to 18 years of age, grouped into seven age levels told a cartoon story aloud. The consistent increase in speech rate with increasing age was found to be the result of decrease in the duration of pauses on the part of the older subjects, but not to be the result of a decrease in the frequency of pause occurrence.

Still, it was 1977 by the time our first genuinely field-observational study was undertaken with spontaneous spoken discourse. Szawara and O'Connell (1977) compared a number of formal radio homilies with homilies delivered live to students in a university chapel. The more spontaneous homilies manifested pauses of longer duration, which in turn resulted in a slower speech rate. Thus, homilies proved to be the initiation of our studies on the topic of rhetorical readings: Funkhouser and O'Connell (1978) investigated poetry readings; Clemmer, O'Connell, and Loui (1979) continued with church lectors and drama students who read a passage of the New Testament aloud; and O'Connell (1980, 1982) pursued the rhetorical function of pauses in the reading of poetry. All these studies have shown the importance of relatively longer pauses for the purposes of rhetorical expression.

Meanwhile, we were learning that there were many methodological problems attendant upon pause research. A serious one was the use of various cut-off points for the minimum duration of pauses. Goldman-Eisler (1968, p. 12) had defended her use of a cut-off point of 0.25 s as a means of separating "hesitation pauses from phonetic stoppages." However, Hieke et al. (1983, p. 212), in a phonotactic study of several corpora of readings, have suggested "a minimum pause duration of somewhat over 0.10 s," because they found that most pauses between 0.13 and 0.25 s were psychologically functional, e.g., were used rhetorically to express emphasis. Currently, the availability of the PRAAT software allows us to use a cut-off point of 0.10 s. But this availability is of recent vintage. We began with rather cumbersome off-line equipment: first a Schwarzer Physioscript recording machine, then a Brüel and Kjaer audio frequency spectrometer and level recorder, and finally a Siemens

Oscillomink L along with an F-J fundamental frequency meter, before the PRAAT equipment became available.

Our insistence upon instrumental measurement somehow left us with the burden of proof that perceptual assessment of pause occurrence and duration was not acceptable because it yielded unreliable data. And so, we engaged empirical research on pause reports, in which measured and perceptually assessed pauses were compared. We enlisted a variety of types of experimental subjects to make such assessments. Stuckenberg and O'Connell (1988, p. 19; see also Carpenter & O'Connell, 1988) have concluded from their research with student subjects that because of over- and under-estimates of pause duration:

> *Pause reports of this kind diverge from objectively measured pause data as a function of a number of independent variables and are therefore not to be trusted as objective estimates of either pause occurrence or pause duration.*

Kowal and O'Connell (2000, p. 353; see also Spinos, O'Connell, & Kowal, 2002) have examined pause notations in published transcripts produced by transcribers trained in the use of three commonly used notation systems. These pause notations were then compared with the pauses in the original audio recordings as measured instrumentally by ourselves:

> This comparison revealed that long ($\geq$1.0 s) and medium ($>$0.30 to $>$1.0 s) pauses were transcribed relatively accurately, but short pauses ($\geq$0.12 to $\leq$0.30 s) were not; both false alarms and misses occurred. Scaling of pause duration was found to be ordinal, but not in accord with the interval and ratio scales specified within the SEU [*Survey of English Usage*] and GAT [*Gesprächsanalytisches Transkriptionssystem*] notation systems, respectively.

Both the *SEU* system (e.g., Svartvik & Quirk, 1980) and the *GAT* system (Selting et al., 1998) as well as the *Halbinterpretative Arbeitstranskriptionen* (*HIAT*) system (e.g., Ehlich, 1993) have thus been found to overburden even trained transcribers.

Political rhetoric as a venue for our research has been a relatively late addition. But the need for audio recordings of good quality, for speakers with rhetorical skills and good articulation, and for the discussion of serious topics, along with our own growing interest in truly dialogical speech, has led us in that direction. Kowal (1991) has found that the distribution of pauses is different in both German and American political speeches compared to interviews: Pauses between sentences were less frequent in the interviews than in the formal speeches; pauses within syntactic phrases were almost five times as numerous in the interviews as in the speeches; and the phrase-internal pauses in the interviews served both hesitation and rhetorical functions, insofar as they were completely out of line with the syntactic structure of the phrases. Not only does this sort of pause usage in the interviews indicate that they are genuinely spontaneous in comparison with the formal speeches, but also that they reflect a higher level of conceptual orality compared to the relative literacy of the formal speeches.

Silence

The concept of *silence* has attracted the attention of researchers from various disciplines, including sociology (e.g., Bellebaum, 1992), philosophy (e.g., Dauenhauer, 1980; see our Chapter 8), and linguistics (e.g., Jaworski, 1993; Kurzon, 1997; Zimmermann, 1983). The term itself has from time to time embraced a wide variety of meanings. In a narrow sense, we have used the term *silence* in this book as synonymous with pause time; but we must also inspect the recent history of its usage in a broader research context.

In the introduction to their edited book on *Perspectives on silence*, Tannen and Saville-Troike (1985, p. xvii) have listed five forms of silence: The shortest is the cessation that Goldman-Eisler (1968, p. 12) had claimed to be "part of articulation"; next comes the pause perceived (if at all) as hesitation; then there is the interactional pause termed by Goffman (1967, p. 17) "a momentary lull," which, in his own example, an interlocutor might use as an occasion to withdraw from a conversation without offending anyone; then there is the complete absence of speaking on the part of an interlocutor; and finally, the broadest form of silence is the background silence that epistemologically provides the contrast to speaking, a form of silence used ritualistically in some religious settings.

The *locus classicus* for a distinction between pauses and silences in spontaneous spoken discourse is to be found in Sacks, Schegloff, and Jefferson (1974, p. 696). They have distinguished among pauses, gaps, and lapses, depending upon whether a silence occurs within or between turns:

> Intra-turn silence (not at a transition-relevance place) is a 'pause', and initially not to be talked in by others; silence after a possible completion point is, initially, a gap, and to be minimized; extended silences at transition-relevance places may become lapses. (p. 715, footnote 26)

It is evident that, for Sacks et al., at least the "extended silences" between turns indicate a problem of communication on the part of the interlocutors. Such a negative evaluation of silences that are essential parts of spontaneous spoken discourse is not at all uncommon. Thus, Grabher and Jessner (1996, p. XI) have stated: "In Western culture silence is, by its very definition, negatively connoted"; and they have added that "silence in company is experienced as threatening, creating discomfort, doubt, and irritation." They themselves, however, have emphasized that silence "contributes to communication." One possible reason for a negative view of silence may be that it has been conceptualized as a background phenomenon against the positively evaluated figure of talk; it becomes noticeable only where the expectation of speech is disappointed. Interestingly enough, Saville-Troike (1985, p. 17), in her classification of various categories of silence, has listed only examples of negative emotions under the subcategory of silences that are related to the personal psychology of a speaker: "timidity, embarrassment, fear, neuroses."

Zuo (2002) has critically discussed the concepts and theories of silence in both mainstream psycholinguistics and the conversation-analytic research tradition in modern times. Unaccountably, however, his review of the literature stops abruptly in the mid 1990s of the twentieth century. Hence, it is currently outdated by more than a decade.

Baldauf (2002, p. 55; our translation) has claimed that there are two different types of silence:

> When silence is described in the scientific literature, "marked silence" is generally meant: Silence has multiple meanings, but it always *means* something, whether agreement or disagreement, making contact or breaking it off, being reserved or forward. This, however, disregards the fact that silence can also be unmarked, unnoticed, that it need not be perceived – either positively or negatively – as silence.

She has also cited Meise (1996, p. 55; our translation), who has used the term "unmarked nothing" for the time that occurs before an oral communication begins and after it has ended, periods of time that, it is asserted, are often difficult to distinguish from marked silence. However, in point of fact, it is not plausible to refer to "unmarked nothing" as silence: A prerequisite for identifying any down-time as silence relevant to speaking and listening is *intersubjectivity*, i.e., a mutual and reciprocal awareness of one another on the part of the interlocutors (see also our Chapter 19). The mere acoustic absence of sound before or after someone speaks or listens does not suffice to constitute silence. Were this the case, I would be silent with respect to an old friend (whom I am *not* expecting to see when the elevator door opens) during my trip from the first to the sixth floor where he is standing in front of the elevator; but such down-time is simply irrelevant to speaking and listening. The same refers to the down-time after speaking and listening. The time immediately following the batter's shout "heads up" to the fans in the nearby stands, as a baseball comes whizzing by their heads, is not silence; the interaction is all over and the ball game goes on. There is no residual interaction left, and certainly no intersubjectivity between the batter and fans at that time.

In summary, we wish to emphasize the objective measurement of pauses (both within and between turns) and their positive potential for communication. It was to emphasize the latter point that we began this chapter with a quotation from the recent (2006) Nobel Prize winner for literature Orhan Pamuk (1997, p. 225 f.) – a passage from his novel *The new life* that provides a fitting example of the communicative value of silence between a pair of old friends. Such silence transcends the extant categories of the various disciplines that concern themselves with language use. It is a mutual subjective stance that incarnates an unusual level of intersubjectivity and accord. It is an open, receptive, and respectful, but not at all a passive or inactive silence. Nor is it in any sense of the term an uncomfortable silence.

But not all approaches to face-to-face interaction leave room for such silence. As a generalization, the following description by Linell (2005, p. 20) seems far too restrictive:

> The speaker must produce his utterances quickly and readily, and the listener must respond just as rapidly, under the pressure of the emotive and social atmosphere of the face-to-face interaction.

Pamuk's (1997) example above would appear to emphasize that both "the emotive and social atmosphere" may definitely work together to foster long silences.

Chapter 12
Prosody

What is most comprehensible in speech is not the word itself, but rather the tone, intensity, modulation, tempo, with which a sequence of words is spoken – in short, the music behind the words, the passion behind this music, the person behind this passion: in other words, everything that cannot be written (Nietzsche, 1980, p. 89, Fragment Nr. 296; cited in Blank, 1991, p. 9; our translation).

Chapter Prospectus

Chapter 12, *Prosody*, emphasizes on-time in spoken discourse for the most part. Variations in the on-time variables of loudness, articulation rate, and intonation, and in the off-time variable of pauses are precisely the elements absent in written texts. Such deletion in the written relative to the spoken makes the written barren in itself and lacking in potential for expressiveness; only the prosodic skill of an expressive reader can bring a written text to life by reading it aloud. The continuous variability of prosody defies transcription into discrete units, even though its variability in spoken discourse constitutes an important determinant of meaning. Researchers are gradually coming to terms with the necessity of considering all these factors of prosodic variability as simultaneously operative, if indeed they are to adequately assess the contribution of prosody to meaning in spontaneous spoken discourse.

The Concept

Merriam-Webster's collegiate dictionary (11th ed., 2003, p. 998) has offered the third – and for our purposes, relevant – meaning of *prosody* as "the rhythmic and intonational aspect of language." The *APA dictionary of psychology* (American Psychological Association, 2007, p. 742) is more detailed: "A phonological feature of speech, such as stress, intonation, intensity, or duration, that pertains to a sequence of PHONEMES rather than to an individual SEGMENT" and refers the reader further to "PARALANGUAGE" and

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_12, © Springer Science + Business Media, LLC

"SUPRASEGMENTAL." Note that the APA definition comes down on the side of phonology (over phonetics) and lists stress independently of intensity. Couper-Kuhlen and Selting (1996b, p. 11) have conceptualized prosody as comprising "the 'musical' aspects of speech": "Auditory effects such as melody, dynamics, rhythm, tempo and pause." A phonetic definition of prosody has been provided by Kohler (1995, p. 13 f.; our translation), including "variations in pitch, intensity, levels of emphasis, tempo, register, general voice quality."

In the previous chapter, we have already analyzed the role of pauses in spoken discourse. Here it needs only to be further emphasized that the patterning of on-time and off-time by the frequency and duration of pauses constitutes an important contribution to prosody. In this respect, the generalization – normative rather than empirical – on the part of Duranti (1991, p. 137) and others that silence is out of place in conversation is an unwarranted oversimplification:

> In certain kinds of verbal exchanges – what conversation analysts call "conversations" – silence is to be avoided and gaps between turns are to stay as short as possible (cf. Sacks, Schegloff & Jefferson, 1974; Schegloff & Sacks, 1973).

In the context of prosody, the term tempo is also popular. Couper-Kuhlen and Selting (1996b, p. 11; see previous paragraph) have clearly associated tempo with on-time and have distinguished it from pauses. Hence, tempo is here synonymous with articulation rate as traditionally measured in terms of syllables per second (syl/s) of on-time, and it presumes assessment (perceptual or instrumental) of off-time. Just as with the off-time or pause components of prosody, the on-time elements of prosody also work in consort with one another and with pauses to construct the rhythmic and intonation patterns of spoken discourse.

Paradoxically, definitional problems arise far more with respect to adjunct concepts than with respect to the basic physical variables of loudness, pitch, and time. Thus, *stress, emphasis*, *prominence*, and *salience* are all in need of clarification. They are used sometimes interchangeably and sometimes with specific meanings exclusive of one another. Perhaps the most opaque of all is *stress*. *Merriam-Webster's collegiate dictionary* (11th ed., 2003, p. 1235) offers the following subordinate meanings of *stress* relevant to prosody:

> **4 :** intensity of utterance given to a speech sound, syllable, or word producing relative loudness **5 a :** relative force or prominence of sound in verse **b :** a syllable having relative force or prominence.

Note that definition number **4** limits *stress* to loudness produced by intensity, whereas definition number **5 b** refers generically to force of a syllable or prominence. Crystal (1997, p. 174) has very clearly limited stress to loudness in at least one context: "In English words, each syllable is pronounced with a certain level of loudness, or stress." Maynard, Houtkoop-Steenstra, Schaeffer, and van der Zouwen (2002, p. 494) have introduced the term *emphasis* and have made it broader than just loudness: "Emphasis is done with some combination of a changing pitch, rise in volume, stretch of a sound, or stress on a vowel." In

this case, stress is clearly distinguished from "rise in volume." We are faced with a confusing array of terminology. Sometimes, the definitions and their examples even get in one another's way, as in *Collins cobuild advanced learner's English dictionary* (2003, p. 1433), where *stress* has been defined in terms of a syllable that "sounds slightly louder." But the corresponding example thereof is precisely an example of equal stress: " *'Sit down,' she replied, stressing each word,*" so that one must ask, "Louder than what?"

Accent in turn is defined in *Merriam-Webster's collegiate dictionary* (11th ed., 2003, p. 7) as "**1 :** an articulative effort giving prominence to one syllable over adjacent syllables; *also* **:** the prominence thus given a syllable." It would seem reasonable to assume that the "prominence" intended here for both *stress* and *accent* can be contributed by a number of prosodic variations, including relative loudness, syllabic prolongation, surrounding pauses, and – paradoxically – diminution of loudness. Even an extra-linguistic variable, e.g., an accompanying gesture, bodily movement, or deixis, can contribute such prominence to a speech sound, syllable, or word.

Intonation too must be defined: "The rise and fall in pitch of the voice in speech" (*Merriam-Webster's collegiate dictionary,* 11th ed., 2003, p. 656). The richness of this concept is reflected in Crystal's (1997, p. 173) listing of the many functions of intonation: emotional, grammatical, informational, textual, psychological, and indexical.

A danger in the midst of all this variability is the possibility that the researcher may analyze the cycles per second (cps) of pitch, the decibels (db) of loudness, and/or the milliseconds (ms) of time as isolated physical phenomena, without acknowledgement that they are used by speakers together and in conjunction with extra-linguistic factors for an overall effect on meaning. In this respect, Quirk, Greenbaum, Leech, and Svartvik (1985, p. 1589) have warned us "against simple equations such as regarding stress as identical with loudness. . . .other factors are or can be involved – notably duration and pitch." Unfortunately, this very complexity has provided an occasion for some qualitative analysts to pooh–pooh quantitative analyses as purely physicalistic and irrelevant to the interactional situation. It is hardly our intent to join in this chorus, but rather simply to warn against the ever present danger of reductionism and oversimplification.

Prosody and Meaning

It is a basic fact that the *way* something is *said* may alter what is *meant*. What we will later (see our Chapter 23) refer to as a veritable *somaticization* of the syntax of spontaneous spoken discourse that sometimes transcends, modifies, supplements, or supplants the sentential syntax of well-formed grammatical units is saliently subject to the uses of these prosodic means. For example, in context, a very emphatically uttered masculine third-personal pronoun in the

assertion *he did it* can constitute a firm statement of the innocence of a female suspect for whom the assertion *she did it*, with a correspondingly emphatically uttered feminine third-personal pronoun, would have been appropriate. Her innocence is not so much a logical conclusion that must be inferred from the emphatic *he*. Instead, it is a shift in the very meaning of *he* in this context from simply *he* to *he, not she*. There are innumerable cases of irony, sarcasm, play on words, and other usages in everyday speech that display such *somaticized* syntax. Even if the traditional structural linguists may not be quite ready to acknowledge this determination of meaning in the interaction of spontaneous spoken discourse, it is nonetheless very important psychologically for the understanding on the part of the listener of the meaning intended by the speaker.

The Transcription of Prosody

Couper-Kuhlen and Selting (1996b, p. 11), in their edited volume *Prosody in conversation: Interactional studies* (1996a), which has recently been reprinted in paperback (2006), have deplored the vacuum of research on prosody: "It is surely no exaggeration to state that a large part of this field has been left untilled by modern structural linguistics." They have ascribed this neglect to the failure to allow "speech prosody and language-in-use" to cross-fertilize one another and have concluded that "it is doubtless the overwhelming influence of literacy on thinking about language which has been responsible for the neglect of prosody." Couper-Kuhlen and Selting have further alleged three sources for the neglect of prosodic phenomena: (1) They are not "segment-based, referential units"; (2) they are continuous rather than discrete units; and (3) they are not systematically codified in writing.

Related to this third source, there seems to be one additional reason for the neglect of research on prosody: conflicting and/or inefficient transcription notations. The *GAT* system (Selting et al., 1998), used by Couper-Kuhlen and Selting (1996b), reflects some of the problems of transcription notation systems. In examples of German utterances transcribed according to *GAT* in Kowal and O'Connell (2003a, p. 100), one finds emphasis notated as "!PIK!"; rising intonation as "gewesen?"; falling intonation as "nich."; syllabic prolongation as ":"; loudness as "<<f>wir>"; and quiet speech as "<< p>wir>". None of these notations can reflect either the *suprasegmental* nature of the variables or their *continuous* rather than discrete nature (see Selting, 2001, p. 1065 ff., on problems of prosodic transcription).

Herrmann and Grabowski (1994, pp. 32 f.; our translation) have insisted that there is frequent agreement in the notation conventions of various research groups with regard to "the verbal, the nonverbal, and the utterance-accompanying components" of spoken discourse. However, Kowal and O'Connell (2003a, p. 102) have found only 30% of a total of all the notation

symbols across five German and three English transcription systems in complete agreement with one another. For example, *GAT* (Selting et al., 1998, p. 114) uses uppercase lettering preceded and followed by an exclamation mark as one option for notating emphasis (in the original German, "extra starker Akzent"), whereas all the other systems use a different codification of symbols to notate emphasis. Another example of the problems arising from the diversity of transcription notation systems can be seen in the very first chapter of Couper-Kuhlen and Selting (1996b). They have presented a multitude of cited examples transcribed in accord with various notation systems. These examples make it abundantly clear that there is no unified system for presenting prosodic data. In addition, the diversity makes it very difficult for a reader to understand the transcripts, not to mention the difficulties of reproducing already published transcripts in further publications, as we have discussed in Chapter 10. Quite in accord with these observations, Crystal (1997, p. 172) has referred to the notation systems for transcribing intonation as "competing descriptive frameworks" that vary greatly precisely because they reflect different theoretical views (e.g., phonetic vs. phonological; auditory vs. acoustic).

Research

Most of the prosodic research that has been undertaken in recent years has been concerned with intonation. Many years ago, Abercrombie (1965, p. 6; as cited in Couper-Kuhlen & Selting, 1996b, p. 12) set the stage for such investigation:

> If you are reading aloud a piece of written prose, you infer from the text what intonations you ought to use, even if, as is almost always the case, you have a choice. The intonation, in other words, adds little information. But if you try to read aloud a piece of written conversation, you can't tell what the intonations should be – or rather what they actually were. Here the intonations contribute more independently to the meaning.

While crediting conversation-analytic research with an interest in intonation, Couper-Kuhlen and Selting (1996b, p. 13) have contended that the interest of this research stops largely at the level of the transcript and rarely figures "in the analyses which conversation analysts have so far offered." Gumperz (1982, p. 100; 1992) they have considered to be the exception, with his process of contextualization through the use of prosodic features.

Couper-Kuhlen and Selting (1996b) have found a number of problems with current intonation research. Criteria for the identification of intonation units have become controversial because they pit phonetic and phonological persuasions against one another. Couper-Kuhlen and Selting themselves have proposed to bypass the problem by going beyond traditional grammatical units and "taking a discourse perspective" (p. 16), in accordance with which

> the basic *prosodic phrase* in speech, when viewed interactively, is likely to be *not* the prosodic counterpart of a grammatical sentence or clause, but rather a unit defined with respect to the utterance as a turn-constructional unit, a 'phonetic chunk' which speakers use to constitute and articulate turns-at-talk.

Thus, they (p. 21) have linked intonation to interactional functions and goals and have referred to this as "*pragmatic* 'meaning,' " "situated, inference-based interpretation" rather than "semantic meanings of decontextualized linguistic forms." Once again, we seem to be dealing with a syntax for spontaneous spoken discourse that transcends and supplements traditional grammatical categories, and in the case of intonation analyses goes beyond a long historical tradition of such grammatical categories. For our own part, we would find Couper-Kuhlen and Selting's "*pragmatic* 'meaning' " quite compatible with our own understanding of the basic semantic meaning of an utterance from a psychological perspective.

It should be noted, however, that their "situated inference-based interpretation" may well go beyond the evidence. There may be some confusion between the immediate understanding of meaning on the part of interlocutors and the inferential processes of research analysts, in the sense that the researchers are indeed referring to their own (quite legitimate) research inferences. However, the interlocutors do not necessarily proceed by "uncoupling intonation from lexico-syntax" (p. 22); rarely does the interactional use of intonation to determine meaning require a throw-away of lexico-syntax. Quite the contrary, there is most commonly a co-determination on the part of intonation (and many other prosodic and contextual factors) and lexico-syntax. In fact, we would argue that Brazil, Coulthard, and Johns (1980, p. 18; cited in Couper-Kuhlen & Selting, 1996b, p. 22) have made an artificial dichotomy between "linguistic features of the message" and "the speaker's assessment" in the following:

> Tone choice, we have argued, is not dependent on linguistic features of the message, but rather on the speaker's assessment of the relationship between the message and the audience.

After all, it is the speaker who deliberately chooses the linguistic features – along with tone – precisely to aid and abet in the communication of the message to the audience. Intonation is decidedly *not* "primarily a symptom of how we feel about what we say" (Bolinger, 1989, p. 1; as cited in Couper-Kuhlen & Selting, 1996b, p. 23), it is a constituent determinant of what we say, part of what we say. Nonetheless, the distinction between syntactic and prosodic units remains extremely important (see Kern & Selting, 2006, p. 244).

The reader will notice that mainstream psycholinguistics has not played a prominent part in the intonation research detailed above. By and large, the psychologists have not been ready to go beyond the lexico-syntax as Couper-Kuhlen and Selting (1996b, p. 21) have done with their "*pragmatic* 'meaning.' " For example, the prosodic research of Grosjean and his colleagues (e.g., Grosjean, Grosjean, & Lane, 1979) was based on "isolated passages without

any communicative intent" (O'Connell, 1988, p. 162). More recent psycholinguistic research displays the same neglect. We have sampled a number of English- and German-language psycholinguistic texts with the following outcome: Clark (1996, p. 182) has limited himself to several paragraphs in which he seems to make "intonation or prosody" synonymous; Harley (2001, p. 106) has mentioned prosody only in the context of the language development of infants, but has not included intonation in his index; in Dietrich's index (2002), neither intonation nor prosody is to be found; and Rickheit, Sichelschmidt, and Strohner (2002, p. 52) have included only one short paragraph regarding a generic definition of prosody. Hence, it is gratifying to find in Carroll's (2007, p. 70 f.) recent textbook an extended treatment of both prosody and intonation.

Our Research on Articulation Rate

We wish to present here a set of our own research projects on one particular topic in the domain of prosody – articulation rate. Goldman-Eisler (1968, p. 25) had considered articulation rate to be "a personality constant of remarkable invariance" and had accordingly neglected to observe its variation across settings and genres. However, in a number of projects, spanning 1986 to 2004 – all with a cut-off point of 0.12 or 0.13 s for pauses – one of our clearest findings has been that there was no overlap whatsoever between mean articulation rates for rhetorical readings of poems (3.69, 4.20, and 4.72 syl/s; Sahar, Brenninkmeyer, & O'Connell, 1997, p. 453) and inaugural addresses (4.37 syl/s; Kowal et al., 1997, p. 14) on the one hand and TV interviewers and interviewees (means ranging from 5.04 to 6.14 syl/s; Kowal & O'Connell, 1997, p. 313; Kowal & O'Connell, 2004b, p. 91; O'Connell & Kowal, 1998, p. 549) on the other. Articulation rate seems to be far more complex than Goldman-Eisler's personality constant makes room for.

Futuristics

At the beginning of the twenty-first century, Couper-Kuhlen (2001, p. 16) set out to review research on the use of intonation in discourse. Essentially, her approach has been to record as historical the competitions of the past, to note the divisions of researchers into several schools in the present, and to emphasize the importance for all of dealing with the complexity of prosody in the future:

> Intonation – in the restricted sense of "pitch configuration" – rarely functions alone to cue an interpretive frame. The same frame may be cued by timing and volume as well. . . . in the contextualization-cue approach there has been a subtle shift away from the study of "intonation" to the study of *prosody* and discourse.

Her mention (p. 25) of "a second type of new territory in the field of interactional prosody" is even more futuristic. She has emphasized the universality of

the temporal dimension of spoken discourse and has predicted that "the focus here will be on *timing*." She has acknowledged the availability of objective measures of timing, but has striven to go beyond that level in search of "the metric which is behind participants' subjective judgment of time." Perhaps mainly as a consequence of our original training as experimental psychologists, our own research has always been couched in terms of objective – instrumentally measured – time. We have indeed confirmed in a series of studies (see our Chapter 11) that both experimental subjects' and trained transcribers' judgments of time are sometimes quite different from objectively measured time. Moreover, we are aware that the study of time as prosody is not about "participants' subjective *judgment* of time," but about participants' *use* of time. It is the research analyst who *judges* time; participants *use* time.

Chapter 13
Fillers

People around the world fill pauses in their own languages as naturally as watermelons have seeds. In Britain they say "uh" but spell it "er," . . . "uh" is the only word that's universal across languages (Erard, 2007, p. 55).

Chapter Prospectus

Fillers such as *uh* and *um* are themselves hesitations, and they sometimes announce new material and/or planning difficulty. They are most characteristic of spontaneous spoken discourse, but are seldom found in written discourse. Hence, to know about their form and function is important for any theory of spontaneous spoken discourse. For all their formal simplicity, they have occasioned much controversy and confusion. They may or may not be acknowledged by lexicographers as words in a given language. In American English, *uh* and *um* serve as the most common fillers, in British English *er*, and in German *äh*. Other syllables are eligible to serve as fillers (e.g., *hm*), but across the board, the type/token ratio of various types of fillers is not high.

The Remarkably Versatile Schwa

There is a very simple, weak sound that has had a most interesting career to date in the English language. Defined in *Merriam-Webster's collegiate dictionary* (11th ed., 2003, p. 1111) as "an unstressed mid-central vowel (as the usual sound of the first and last vowels of the English word *America*)," the schwa appears in printed form as either *uh* or *a*, and, with the addition of an *m*, becomes *um*. Although *um* serves as a somewhat longer hesitation particle than *uh*, it does not enter into our discussion at the moment. In spoken English, it is peculiar to the schwa that it can serve in four very different linguistic functions: as an interjection, as a filler, as a syllable of laughter, and as the indefinite article. No other simple sound has such marvelous versatility in the English language. When it serves in one of the first three of these functions, it is ordinarily spelled as *uh*; when it serves as the indefinite article, it is spelled as *a*.

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_13, © Springer Science+Business Media, LLC 2008

Methodologically, this versatility can and does lead to difficulties for the researcher. For example, if a speaker begins a phrase with an indefinite article, pauses, and then proceeds otherwise without articulating the noun that belongs to the indefinite article, it becomes unclear to the researcher whether the schwa is being used by the speaker as a filler or as the indefinite article. Or if a speaker breaks into laughter with *uh*, *hu*, *hu*, *hu*, it may not be conclusively clear whether he or she is beginning with a filler or an interjection followed by laughter or with an initial pulse of *uh* laughter. For the listener involved in a conversation, these are only momentary glitches, whereas for the researcher, they may remain permanently uncategorizable events. There are sometimes additional (e.g., prosodic or semantic) cues which can help to decide these doubts for the researcher, but not always. And so the transcription of the schwa can occasion some very subtle discernment – and may even have to be left indeterminate in some transcripts.

The Lexicographer's Written World of Fillers

In Table 13.1, a selection of words in printed form that are considered to be interjections in *Merriam-Webster's collegiate dictionary* (11th ed., 2003) is to be found. Along with these are a number of non-entries. We wish to discuss the rationale for selection of the entries and non-entries in order to shed some light on the nature of fillers. We will revisit Table 13.1 in Chapter 14 for the discussion of interjections themselves.

First of all, it should be noted that non-entry in a dictionary is not the definitive decree that a printed item is not a word. Fully 21% (7/34) of the entries listed in Table 13.1 are entries made for the first time within the twentieth century. Similarly, O'Connell and Kowal (2004b, p. 464) have listed fillers that are to be found in 20 different dictionaries between 1938 and 2003; *er* and *um* are entries in eight and *uh* in seven of these, mostly in the most recently published dictionaries (i.e., since 1990). Language usage is constantly undergoing change, and such changes can be reflected in dictionaries only with a certain time lag. Understandably also, dictionaries – as printed sources themselves – primarily reflect written usage and the literate traditions of a language. Hence, in our discussion of fillers and their role in spontaneous spoken discourse, lexicography can throw light on the conceptual literacy/orality involved. With respect to the non-entries in Table 13.1, they are items characteristic of spontaneous spoken discourse which are immediately recognizable by any native speaker of English as legitimate usage; they are decidedly not characteristic of written English. The orthography chosen for them in Table 13.1 thus has an element of the arbitrary, although we have attempted to abide by the conventions of written English. In this connection, it may be noted that the non-entry *huh-uh* occurs in all likelihood more frequently in spontaneous spoken discourse than the synonymous word *uh-uh* which is an entry.

Table 13.1 A Selection of Words (+) Listed as Interjections in *Merriam-Webster's collegiate dictionary* (11th ed., 2003), their Origin in Middle (ME) and Old English (OE), and French (Fr.), Along with the Century (C) of Origin, their Meaning, an Example of their Use, and a Selection of Common Items from English Spontaneous Spoken Discourse that are Non-entries (–)

Entry	Word	Function	Origin	Meaning	Example
a(a)rgh	–				
ah	+	interjection	13th C ME	used to express delight, relief, regret, or contempt	ah, home at last
aha	+	interjection	14th C ME	used to express surprise, triumph, or derision	aha, I caught you
ahem	+	interjection	1963	used to attract attention, or to express disapproval or embarrassment	Jack, ahem
ahoy	+	interjection	1847	used in hailing	ship ahoy
ay	+	interjection	14th C Fr.	usually used with following *me* to express sorrow or regret	ay me, all is lost
boo	+	interjection	15th C ME	used to express contempt or disapproval or to startle or frighten	boo, surprise
eek	+	interjection	1951	used to express surprise or dismay	eek, a mouse
eh	+	interjection	13th C ME	used to ask for confirmation or repetition or to express inquiry	eh?
er	+	interjection	1862	used to express hesitation	er, he's ok
gee	+	interjection	1884	used as introductory expletive or to express surprise or enthusiasm	gee, that's nice
ha(h)	+	interjection	before 12th C ME	used to express surprise, joy, or triumph	hah, I win
ha-ha	+	interjection	before 12th C ME	used to express amusement or derision	ha-ha, you tripped
heh	–				
he-he	–				
heigh-ho	+	interjection	1520	used to express boredom, weariness or sadness – or encouragement	heigh-ho, heigh-ho, off to work we go
hey	+	interjection	13th C ME	used to call attention, or to express interrogation, surprise, or exultation	hey, good job
hey presto	+	interjection	1731	British: suddenly, as if by magic	hey presto, gone
ho	+	interjection	15th C ME	used to attract attention to something specified	land ho
ho-ho	–				
ho hum	+	interjection	1924	used to express weariness, boredom, or disdain	ho hum, another election

Table 13.1 (continued)

Entry	Word	Function	Origin	Meaning	Example
huh	+	interjection	1608	used to express surprise, disbelief, or confusion, or as an inquiry inviting affirmative reply	a miracle? huh
huh-uh	–				
o(h)	+	interjection	13th C ME	used to express an emotion or in response to physical stimulus, used in direct address, used to express acknowledgement or understanding of a statement, used to introduce an example or approximation	oh that hurts
o-ho	–				
oops	+	interjection	1833	used to express mild apology, surprise or dismay	oops, I didn't see you coming
ouch	+	interjection	1838	used to express sudden pain	ouch, that hurts
ow	+	interjection	1865 ME	used to express sudden pain [originally surprise]	ow, that hurts
phew	+	interjection	1604	used to express relief or fatigue; used to express disgust at or as if at an unpleasant odor	phew, how awful
piyu	–				
tsk	+	interjection	1927	used to express disapproval	tsk, bad boy
ugh	+	interjection	1678	used to express the sound of a cough or grunt or to express disgust or horror	ugh, how horrid
uh	–				
uh-huh	+	interjection	1889	used to express affirmation, agreement, or gratification	uh-huh, I did help out
uh-oh	+	interjection	1971	used to express dismay or concern	uh-oh, the cops
uh-uh	+	interjection	1924	used to express negation	uh-uh, I didn't do it
um	+	interjection	1672	used to indicate hesitation	well, um, no
whew	+	interjection	1513	used to express amazement, discomfort, or relief [unvoiced]	whew, that was a close one
w(h)oops	+	interjection	1833	variants of oops (see above)	woops, I missed
wow	+	interjection	1513	used to express strong feeling (as pleasure or surprise)	wow, what a game
yow	–				
yu(c)k	+	interjection	1966	used to express rejection or disgust	yuck, I can't eat it
zounds	+	interjection	1592	used as a mild oath	zounds, I'm caught

The most interesting thing to be noted in the O'Connell and Kowal (2004b) list of fillers is that, despite the fact that the *Oxford English Dictionary* (*OED*, 1989) has listed *uh* as a word since 1961, *Merriam-Webster's collegiate dictionary* (11th ed., 2003) still does not acknowledge *uh* as a word; that *uh* is a non-entry in *Merriam Webster's* is also reflected in Table 13.1. We may note too that the most common filler in the German language (*äh*) is currently listed as a word in the *Duden deutsche Rechtschreibung* (2004, p. 143). Paradoxically, as Table 13.1 also reflects, the filler *um* is a word entry in *Merriam-Webster's collegiate dictionary* (11th ed., 2003, p. 1356), even though its fellow filler *uh* is used far more frequently than *um* and, in fact, more frequently than even a number of very common words and phrases (e.g., *yes* and *I mean*; see Biber, Conrad, & Leech, 2002, p. 449). Interestingly enough, both with respect to *uh* as a non-entry and *um* as an entry, *Collins cobuild advanced learner's English dictionary* (2003, p. 1569) exactly parallels *Merriam-Webster's collegiate dictionary*.

It is difficult to understand the animus against *uh*, which is the rule all the way from the prototypical schoolmarm to *Merriam-Webster's* itself except through recourse to the written language bias (Linell, 1982, 2005) and its parallel scriptism (Taylor, 1997; see our Chapter 2). Hesitation is a necessary and normal concomitant of speaking spontaneously, and the primary means for indicating such hesitation in the English language is through the simple articulation of the schwa – *uh*. There is every reason to expect it to be considered an English word, in view of its frequency as a normal element in speaking.

Fillers in Written Materials

Since novels simulate spontaneous spoken discourse in writing perhaps best of all written genres, O'Connell and Kowal (2004b, p. 461) have counted the frequency of occurrence of fillers in a set of modern American novels (Clancy, 2000; Grisham, 2000; Kidd, 2002; and Ludlum, 2000). We have found to our surprise that even dramatic hesitations are sometimes represented as perfectly well-formed sentences. For example, Kidd (2002, p. 272) has related the following scenario:

Example 13.1
When I took a deep breath, it stuttered as it came out. *My mother's belongings.*

Similarly, Ludlum (2000, p. 232) has described the following scenario in well-formed prose:

Example 13.2
Finneran hesitated, barely a second or two, but it was too long. 'I'm not sure I should say just yet —.'

It seems, then, that the bias in favor of well-formed sentences in the written mode leads to the substitution of *descriptive prose* rather than the simulation of

exact wording that might offend against written well-formedness. And when a filler is indeed used, it reads as a very intrusively disfluent and disruptive element, as in Grisham's (2000, p. 202):

Example 13.3
A long pause as the old man filled his ear. [description]
Then, "No, no, they're not from the Federal Reserve. They're, uh, they're lawyers from Des Moines. They represent the family of an old college buddy of mine. That's all."
A shorter pause. [description]
"Uh, Franklin Delaney, you wouldn't remember him. He died four months ago, without a will, a big mess. No, Dad, uh, it has nothing to do with the bank."

The descriptions of "A long pause" and "A shorter pause" fit the well-formed sentence structure quite well. The three instances of *uh* are properly set off by commas, and these commas are suggestive of abruptness in the articulation of the speaker. And so, they stand out – deliberately, one would hope, on the part of the novelist – as somehow foreign to the smoothness of the prose passage. Across the more than 1000 pages of the four novels, we were able to find only 10 fillers. This yields a rate of occurrence of 2.5 *uh*/book. We think it is quite safe to say that the filler is not at home on the printed page.

Some Filler History: The Carrier of Disfluency

Why fillers in spontaneous spoken discourse have had such an erratic research history while at the same time being on the receiving end of both the opprobrium and the neglect of the linguistic community remains a mystery. Let the chronicle speak for itself. It seems to have started with a misnomer in the first place. Maclay and Osgood (1959, p. 21) were the first to use the term "Filled Pause." And more recently, Erard (2007, p. 54) has inverted the order, but retained the notion in the phrase "pause filler." But fillers are neither pauses nor are they used necessarily where there would otherwise be a silence; they are not a sort of putty used to fill the cracks in window frames – to stuff something into a silence. They are simply legitimate hesitations. The matter had been complicated from the very beginning because there was no accepted way of notating a filler in writing. Although there are not many tokens of this type (i.e., fillers), several researchers succeeded in thoroughly confusing the matter. Mahl (1956, 1958, 1959; Kasl & Mahl, 1965) has insisted that he was analyzing the filler *ah*. O'Connell and Kowal (2004b, p. 464) have been unable to find a single dictionary that recognizes *ah* as a filler; it is consistently designated instead as an interjection "used to express delight, relief, regret, or contempt" (*Merriam-Webster's collegiate dictionary,* 11th ed., 2003, p. 26), as has already been mentioned in Table 13.1. Nor is there any evidence that "*Uh* is sometimes spelled *ah* in North American English" (Clark & Fox Tree, 2002, p. 75), apart from Mahl's own misspelling of it. But Maclay and Osgood (1959, p. 21) have insisted that Mahl's " 'ah' is equivalent to our Filled Pause," even while

phonetically representing the fillers they themselves analyzed as "[ϵ, æ, r, ə, m]". Their list also excluded the only fillers recognized by *Merriam-Webster's* as English words (*er* and *um*; [r] is phonetically not the same as *er*). Later researchers have continued to accept Mahl's *ah* along with *er*, *uh*, and *um* (e.g., Bortfeld, Leon, Bloom, Schober, & Brennan, 2001; Schachter, Christenfeld, Ravina, & Bilous, 1991). But more recently, Clark and Fox Tree (2002) have investigated only *uh* and *um*, and Erard (2004, January 2) has included only *er*, *uh*, and *um* in his discussions. Perhaps the most extreme designation of fillers was that of Goldman-Eisler (1961, p. 167). She included them, as we have discussed in Chapter 11, under the category of "irrelevant vocal productions" and accordingly relegated them, along with all the other "irrelevant vocal productions," to the category of pause time, i.e., off-time.

But the history of fillers is also replete with research that has categorically failed to distinguish fillers from other on- and off-time entities as well (O'Connell & Kowal, 2004b, p. 466 f.):

> From silent pauses (e.g., Boomer, 1965; Hawkins, 1971) from nonverbal vocalizations such as snorting and noisy exhalation (e.g., Helfrich & Dahme, 1974); from hesitant repeats (e.g., Boomer, 1963); from [verbal] fillers such as *you know* (e.g., MacWhinney & Osser, 1977; Rose, 1998); from stuttering, mispronunciations, and corrections (e.g., Knapp, 1980; Tusher, 1978); and from the interjections *eh* and *ah* (e.g., *Filled pauses* website, updated August 7, 2003).

Given all these vagaries of categorization, it is small wonder that considerable confusion has emerged in this research tradition. Fillers must be analyzed as a specific category distinct from other verbal and nonverbal categories. But, as we have seen, such clarification has not occurred to date.

Spontaneous spoken discourse just happens to be the only natural habitat of fillers; they can be found nowhere else in such abundance. The other side of this coin is the consistent finding that fillers are perhaps the most valid indicators of the genuine spontaneity of spoken discourse. The only occasion for fillers to be used in written discourse is the simulation of spontaneous spoken discourse in the written mode, as, for example, in the novels mentioned above.

What Are Fillers Really Good For?

First and foremost, fillers characterize spontaneous spoken discourse and are independent of the ambient sentential syntax. It is not the case that they are absolutely mandatory; but they are to be found in abundance in every genre of spontaneous spoken discourse and among the most renowned public speakers and politicians, e.g., former president Ronald Reagan, who came to be known as "the great communicator" (e.g., Ripper, 1998) precisely because his speeches so successfully simulated spontaneous spoken discourse. Individual speakers vary in the fillers they use, in their duration and

frequency. Fillers are typically monosyllabic, and the core vowel is most commonly the schwa. In general, they occur more frequently than false starts and repeats, but not as often as silent pauses. Finally, they appear to be used in all natural languages (see Clark & Fox Tree, 2002, p. 92; Erard, 2007, p. 55; O'Connell & Kowal, 2004b, p. 460).

Traditionally, fillers have been consistently presented as indicators of disfluency (see Kowal, 1991, p. 128) or even pathology. In fact, both Mahl and Goldman-Eisler were clinical psychologists in search of indicators of emotion and even of mental illness. And so, the conclusion of Mahl's (1958, p. 349) research was: "Individuals using 'ah' most frequently were weaned early, had strict parents, and have obsessive traits." Maclay and Osgood (1959, p. 41f.) were the first to emphasize the communicative function of fillers; they assumed that speakers may use a filler in reaction to their own silence in order to hold the floor, i.e., to keep their interlocutors from interrupting. In other words, they emphasized the empirical significance of longer pauses that *precede* fillers. As mentioned above, Goldman-Eisler's (1961, p. 167) generalization was to relegate fillers to the pool of "irrelevant vocal productions." Clark and Clark (1977, p. 262) have referred to fillers as "speech errors," "the most common disruption of the ideal delivery." Kowal (1991, p. 135; our translation) has summarized these various research traditions as follows: From the point of view of clinical psychology, fillers have been considered as symptoms of emotional disturbance; from the point of view of mainstream psycholinguistics, they have been considered symptoms of cognitive planning by speakers; and in accord with an emphasis on conversation, fillers have been thought of as "signals for the listener in the organization of turn-taking." Erard (2004, January 2, p. A 13) has called them "throwaways" and has quoted Fox Tree, who refers to them as the source of "the fragmented nature of ordinary conversation." Clark and Fox Tree (2002, p. 98) have asserted that speakers use fillers to announce that they are having "preparedness problems," and Clark and Wasow (1998, p. 201) have similarly mentioned "planning problems."

For our part, we would emphasize that fillers may also serve an important rhetorical function at the same time as all these other functions. That is to say that a speaker can use fillers more or less deliberately in an effort to maintain fluency; they need not be conceptualized as the principal destroyer of fluency. In accord with this notion, Allwood, Nivre, and Ahlsén (1990, p. 3) have asserted that fillers belong to the class of "*speech management*" phenomena, "whereby the speaker manages his or her linguistic contributions to the interaction and to the interactively focused informational content." In keeping with such an approach, Rose (1998, p. 49) has provided empirical evidence to the effect that "speakers use FPs [filled pauses] to increase their apparent fluency" and to "mitigate undesirable effects of the message." Kowal (1991, p. 146) has provided exactly such an example of former German chancellor Helmut Schmidt speaking English in a TV interview and cliticizing *uh* after several words in the

following (cliticized *uh* has been transcribed in all-capital lettering as joined to the preceding word):

Example 13.4
Just by the way Sir Jimmy Carter and I got along quite well person personally uh now as regards your question I do not see that there is a gap uh uh of ideology between continental Europe andUH America what I do see is aUH clear distinction between the economic concepts or economic policy mixes in continental Europe and the United States a field in whichUH theUH presentUH British uh government is a little bitUH more to theUH side ofUH of Washington. . .

The reader should note that the second, fourth, and seventh examples of cliticization become more plausible if printed as āUH, thēUH, and thēUH, respectively. In this instance, the cliticization of *uh* serves momentarily rather to hide hesitation than to exemplify it.

Research in second-language learning has also indicated that mastery of a foreign language entails fluency precisely in the use of such fillers (Frommer & Ishikawa, 1980; Schmidt, 2004). In a similar vein, Arnold, Fagnano, and Tanenhaus (2003, p. 25) have suggested in their very title, even while still referring to fillers as disfluencies, that "Disfluencies signal theee, um, new information." All these indications constitute a turn back toward the view expressed by Chafe (1980b, p. 170) that "the fundamental reason for hesitating is that speech production is an act of creation," and that "pauses, false starts, afterthoughts, and repetitions do not hinder that goal, but are steps on the way to achieving it." Clark (1996, p. 389) too has more recently written that what were once considered by him to be unsystematic noise and performance errors are now to be thought of as "systematic and essential to the successful use of language." One more positive function of the filler can be traced all the way back to Maclay and Osgood (1959, p. 41), who, as we have noted above, characterized filled pauses as an effort on the part of the speaker to hold on to his or her turn. Their idea still has merit, even though Kowal (1991, p. 147) found the opposite for former president Ronald Reagan, who would use *uh* toward the end of an interview turn to indicate that he was ready to give the turn away or at least to acknowledge his difficulty in further prolonging his response to the interviewer. These suggestions of Maclay and Osgood and of Kowal, although in opposition to one another, both emphasize the dialogical and multi-determined character of the functions of fillers.

Our Own Research on Fillers

Our very first empirical project together was undertaken under the influence of Goldman-Eisler's research on filled and unfilled pauses as an effort to uncover psychological processes within the speaker. O'Connell et al. (1969, 1970) have found that the occurrence of fillers in reading short narratives aloud was negligible. In the story retellings, however, fillers occurred frequently, and 73% of them were preceded by pauses that were on the average the longest

pauses in the experiment; pauses followed only 47% of the fillers. These basic patterns have held up in all our subsequent research. Another consistent result across studies has been that fillers were the most frequent hesitation in comparison with repeats and false starts. In chronological sequence, this result has been found for spontaneous narratives told by subjects 12–18 years of age, but not for subjects 6–10 years of age (Kowal et al., 1975, p. 199), for interviewers (Kowal, Bassett, & O'Connell, 1985, p. 10), for narratives told after a film showing (Kowal, 1989, p. 121), and for German and American politicians interviewed on TV (Kowal, 1991, pp. 140, 157, and 178; O'Connell, Kowal, & Dill, 2004, p. 194). Our interview research has also shown that, although interjections may indeed substitute for a sentence and stand alone as a turn, fillers may not. They function as hesitations with other verbal material, but do not constitute a turn by themselves. A free-standing *uh* indicates that an interruption has terminated a speaker's further development of his or her utterance, not that the speaker intends the filler as an utterance. Accordingly, Reagan's use of *uh* to signal that he was running out of material in the course of an interview turn can be considered a sort of self-interruption.

The Failure to Legitimize Fillers as Words

In a complete about-face from Clark and Clark's (1977) position that considered fillers as carriers of disfluency, Clark and Fox Tree (2002) have argued that *uh* and *um* are legitimate English words. The abdication of the concept of ideal delivery which is implicit in their argumentation is indeed gratifying, but their empirical evidence is not cogent. We have mentioned this research already in Chapter 2 as an example of a problem in methodology. We wish now to review the logic of Clark and Fox Tree with a view to putting it into historical perspective. Their purpose was to legitimize *uh* and *um* as words of the English language. To be words, they must fit into one of the word-type categories, the traditional so-called parts of speech, and they must serve an identifiable function for both speaker and listener. Hence, Clark and Fox Tree set out to engage the extensive London-Lund corpus of spoken English (Svartvik & Quirk, 1980) in order to find answers to their questions as to how *uh* and *um* are used. They hypothesized that *uh* would be followed by shorter pauses and *um* by longer pauses, and the existence of such pauses would serve to define the meaning of these lexical items as signals of coming delay. One may note here that their hypotheses emphasized pauses *following* fillers, whereas Maclay and Osgood (1959, p. 41) had emphasized the importance of pauses *preceding* fillers.

However, the pauses that Clark and Fox Tree found have proven empirically to be neither valid nor reliable (see Kowal & O'Connell, 2000; O'Connell & Kowal, 2004b; Spinos, O'Connell, & Kowal, 2002). The reason for the many false-positive and false-negative pauses perceptually identified by the expert

evaluators in the London-Lund corpus is the notorious inability of human subjects to perform this perceptual task accurately. There are simply too many other clues that offer a sense of closure in such a setting, and these appear to be the evidence actually used to identify pauses, even by language experts. The clues include falling intonation, syllabic prolongation, loudness, syntactic closure, completion of semantic meaning, and in some cases completion of a line of print (O'Connell, 1988, p. 225; see also our Chapter 12).

The experts end up using the same language habits that we all share to perform this task. In addition, it should be noted that, whatever these language experts thought they were doing, it was not what listeners do with pauses. O'Connell (1988, p. 213) has pointed out that one must be very careful to identify what "the use of pauses in various speech contexts on the part of the listener" really is: "Use in this setting does not mean reporting pauses, nor adverting to them; it simply means that the pauses make a difference in the understanding of an utterance." This usage of pauses on the part of the listener reflects the "transparency" of language use formulated by Hörmann (1981, p. 28). In other words, even if the experts had been perfectly accurate in identifying pauses, they would not have been doing with the pauses what the ordinary listener does with pauses. The listener cannot backtrack, re-listen, reflect, replay, or correct an impression; these are exactly what the "expert" does. The reader will notice that we have finally surrounded the "expert" with quotation marks in recognition of the fact that he or she cannot play the role of an expert in any relevant sense in this instance. Stuckenberg and O'Connell (1988, p. 27) have summarized the matter as follows:

> The use of perceptual reports for the identification of pause occurrence and the estimation of pause duration to the exclusion of instrumental measurement, is not justifiable. The practice has led to questionable data and misleading interpretation of data for many decades.

The empirical evidence that pauses do not systematically and predictably follow upon *uh* and *um* has been made even clearer from O'Connell and Kowal's (2005b, p. 567) replication of Clark and Fox Tree's (2002) research. We used a corpus of media interviews of Hillary Clinton to measure with the PRAAT software, www.praat.com, the actual occurrence and duration of pauses following *uh* and *um*. We argued, moreover, that the use of an experienced and highly qualified public speaker constituted a conservative test of Clark and Fox Tree's hypothesis. Overall, we found that only 24% of these fillers were followed by silent pauses. This finding confirms our earlier filler research listed in the previous section, in all of which fewer than 50% of the fillers were consistently found to be followed by pauses. For the Hillary Clinton corpus, then, the use of *uh* and *um* by listeners to predict the occurrence of a silent pause would be wrong 76% of the time. Hence, the assertion that *uh* and *um* are conventional English words used by speakers to signal their intention to initiate a pause is not empirically warranted.

Finally, we wish to interpret the failure of this hypothesis somewhat further. Of all the various phenomena that we have been analyzing in this book as characteristic of spontaneous spoken discourse – fillers, interjections, laughter, repeats, false starts – fillers stand at the periphery of written discourse more than any of the others. To put it another way, the legitimization of words exemplified in Table 13.1 is in terms of a standardization of orthography and function within written language. This is a set of conditions that fillers do not fulfill. They remain characteristic only of spontaneous spoken discourse, not of *The world on paper* (Olson, 1994). Hence, the legitimization of fillers as words on the part of lexicographers has been a gradual process. That lexicographers have tended to do so and to designate them as interjections is to be discussed further in the following chapter.

Chapter 14
Interjections

Primarily language does not express thoughts or ideas, but feelings and affections (Cassirer, 1944, p. 25).

Chapter Prospectus

Chapter 14, *Interjections*, engages a phenomenon that has been historically a marginal, thoroughly neglected linguistic category. Their independence of ambient sentential syntax makes interjections largely irrelevant to a linguistics deeply immersed in grammar, and their association with spontaneous spoken discourse makes them of little interest to research based on a written language bias. Recent research offers promise that prototypical characteristics of interjections can provide a basis for empirical analyses of their various forms and functions. In particular, the relationship of interjections to other particles, e.g., fillers and HA-HA laughter, must be clarified. Empirical evidence shows that, in their role as expressions of current emotion, interjections have strong links to both medial and conceptual orality rather than to medial and conceptual literacy. Conventional ("tame") and nonconventional ("wild") primary interjections (Rhodes, 1992, p. 222) differ from one another not only in their acceptance as words, but frequently as well in the level of their accordance with the phonological rules of the language in which they are expressed. Frequent themes of secondary interjections (i.e., interjections including at least one lexical item) include the deity and other supernatural entities (e.g., *Oh God*). The chapter ends with an analysis of the similarities and differences between interjections and fillers and a review of our own recent research on interjections in media interviews, reading aloud, and dramatic performances.

Some History

The nineteenth and twentieth centuries have not been kind to interjections. There were indeed writings about "natural sounds" (e.g., Winteler, 1892; our translation), "distress calls" (Kluge, 1902; our translation), "babbling words"

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_14, © Springer Science+Business Media, LLC 2008

(Oehl, 1933; our translation), and even interjections in classical languages (e.g., Schwyzer, 1924). But we have not found similar materials for this period written in English or about the English language. The first systematic linguistic investigations of primary interjections were, to our knowledge, those of the Swedish researcher Ideforss (1928), followed by those of Tesnière (1936) and Karčevski (1941), both in French. Reisigl (1999) has summarized much of the history of research on both primary and secondary interjections in the languages of continental Europe; his references, however, are limited to research presented in the German language, and his own research has been largely in terms of secondary interjections.

In particular, the period of time since the beginnings of generative grammar and mainstream psycholinguistics has had little room for entities that have no structural relationship to ambient sentential syntax and are to be found primarily in spontaneous spoken discourse rather than in written materials: An *oh* tells us nothing about the syntactic structure of the sentence that follows it; and, even in its written form, it is a simulation of the spoken.

And so, we have looked to the lexicographers for some information about interjections and their history. We may return now to Table 13.1 from Chapter 13 to examine the interjections listed there as entries (+) in *Merriam-Webster's collegiate dictionary* (11th ed., 2003) and as non-entries (–) used colloquially in the United States. Although not a comprehensive listing, the table reflects the long history of interjections in the written tradition of the English language. For example, *ha(h)* is listed as originating before the twelfth century as an expression of surprise, joy, or triumph. All of the entries are, in fact, listed as interjections, and all of them can be considered primary interjections. We have not included secondary interjections (e.g., *boy*, *God*, *man*). One could argue that *gee* and *zounds* might be considered secondary interjections, but the references to Jesus and to Christ's wounds, respectively, are only etymological. The non-entries are not designated as to function. The source of our "function" entry is the internet version (www.Merriam-WebsterCollegiate.com) of *Merriam-Webster's collegiate dictionary* (11th ed., 2003), and the function listed there is "interjection" for all of them.

What does it mean to say that the function of these words is to be an interjection? It appears to refer to the fact that they are indeed interjected – tossed or thrown into an utterance – without being integrated in any way with the ambient sentential syntax. But this is at best a function by collocation only. And in fact, this independence of ambient syntax appears not to be a linguistic universal, insofar as at least Chinese interjections can be related to adjacent syntax (see Yang, 2004). But for what purposes are interjections used? This is the real question regarding the function of interjections and what this chapter is all about.

Table 13.1 is also part of the modern landscape of interjections. We have already noted in Chapter 13 that seven of the interjections listed in the table were added during the twentieth century, in fact, as recently as 1971 in the case of *uh-oh*. What is to be said of the non-entries still waiting in the wings to be called out upon the

stage as interjections? In some respects, they literally set the stage for the understanding of the entered items insofar as the rationale for their non-entry offers us clues as to what is required for a word to be acknowledged by the lexicographers as an interjection. It seems clear that there are several reasons for their exclusion from the world of interjections: (1) They have not been found by the lexicographers in written materials (e.g., *piyu*; also spelled as *phew* and as *P-U* [Berdy, 2003]); (2) their orthography is not acknowledged as conventional by the lexicographers (e.g., *huh-uh*), even though they do occur in written materials; or (3) they are considered to be fillers and are therefore arbitrarily excluded from the systematized written language by the lexicographers (e.g., *uh*), again even though they do occur in written materials (see *Example 13.1* in our Chapter 13).

The first instance above (1) involves a non-entry (*piyu*) for which there is no conventional spelling at all. Nonetheless, native speakers of American English will generally recognize our somewhat arbitrary spelling of *piyu* as a reasonable representation of a common oral expression in reaction to an exceedingly offensive odor. The second instance (2) has been mentioned already in Chapter 13 as perhaps the more frequent form relative to *uh-uh,* as an expression of negation. The non-acknowledged form represents an initial *h* sound. Chafe (2007, p. 20) has noted a similar sound in HA-HA laughter, and indeed a sound "responsible for the idea that laughs can be written 'ha ha ha'." The form *uh-uh*, listed as an entry, expresses an initial glottal stop. As for the third instance (3), we have already discussed in Chapter 13 the fact that *uh* is far more frequent than *um*, but is nonetheless not acknowledged by all lexicographers as a legitimate interjection, whereas *um* is acknowledged as such – and precisely as an interjection, not as a filler. Hence, frequency of occurrence in *spoken* English is for the lexicographers by no means an adequate criterion for considering these items as words.

Interjections in Modern Language Sciences

Less than two decades ago, the mention of interjections as a focal research category in any of the language sciences would have been ludicrous. Interjections were literally one of "the ragbag categories" (Harris, 1980, p. 20) for European grammarians. They have been a thoroughly "unpopular subject in linguistics" (Ehlich, 1986, p. 1; our translation). And according to Burckhardt (1998, p. 492; our translation), interjections have suffered a "shadowy existence in the grammars of the twentieth century," and that despite a turn in the 1970s toward a more pragmatic approach in linguistics. Hansen's (1998, p. 41) diagnosis has been summed up by O'Connell, Kowal, and King (2007, p. 2) as follows:

> The neglect of interjections in modern linguistics can be related to its written language bias as well as to its focus on the referential function of language, a focus which is in itself not inclusive of the emotional aspects of language use and therefore not inclusive of interjections.

What little research did exist before the twenty-first century has proceeded from a linguistic standpoint (e.g., Fries, 1988, 1990, 1992). A special issue of the *Journal of Pragmatics* edited by Ameka (1992a) and a special issue of the *Zeitschrift für Semiotik* edited by Kowal and O'Connell (2004a) have now begun to focus attention on the pragmatic and the psychological functions of interjections, respectively.

But before the advent of the generative era of linguistics in mid-twentieth century, there was an earlier psychological interest in interjections. Wundt (1900/1911) himself distinguished the categories of primary and secondary interjections and acknowledged onomatopoeic interjections, although he did not include the division of primary interjections (introduced only much later by Rhodes, 1992, p. 222) into conventional ("tame"; e.g., *oh*) and nonconventional ("wild"; e.g., *piyu*). Nübling's (2004, p. 17; our translation) recent emphasis on "emotionality and expressiveness" as the basic criteria in the determination of "ideal types of interjections" had already been anticipated by Jespersen (1922, p. 415): "The usual interjections are abrupt expressions for sudden sensations and emotions." Karčevski (1941, p. 62 f.; see also Eastman, 1992; Müller & Posner, 2004) has emphasized the involvement of the whole human body in expressive spontaneous spoken discourse. Hence, interjections have also been related to gestures, and gestures or other nonverbal expressions can substitute for interjections (Ameka, 1994, p. 1712). Apart from several German-language studies (e.g., Schneider, 1959; Burger, 1980; Kleemann, 1980), the empirical investigation of these phenomena has not been engaged, to our knowledge, until the last several years.

According to Nübling (2004, p. 11), "interjections are often considered a dumping ground for particles which are otherwise difficult to classify." On the basis of a variety of functional, pragmatic, and formal characteristics, she has posited an "interjectional spectrum" of particles, with primary interjections at the "prototypical center"; this spectrum excludes the two extremes of discourse markers and baby talk. In this taxonomy, the onomatopoetic is relegated to a separate word class on the basis of functional criteria (Nübling, 2005, p. 606). In an earlier publication, Nübling (2001, p. 20) has investigated the diachronic development of interjections, and in particular, the development of secondary interjections into prototypical primary interjections. The title of her article is a German-language example of this interjectionalization: "Von *oh mein Jesus!* zu *oje!*" ("From *oh my Jesus!* to *gee!*"; our translation).

Recent Empirical Research

Kowal and O'Connell (2004b) began their empirical research on interjections with an investigation of a German-language interview of Katarina Witt by Günter Gaus in his TV series *Zur Person*. We chose that interview quite deliberately after looking in vain for interjections in political interviews. What

we had found was that politicians are generally cautious for the sake of their constituency: They wish to say the right thing, and such a stance hardly encourages the spontaneous expression of emotion. The result is a general paucity of interjections in political TV interviews. For example, throughout seven of Gaus's political TV interviews (with the German politicians Adenauer, Kohl, Schily, Scharping, Schröder, Schäuble, and Vollmer), lasting a total of five and a half hours, only 15 interjections occurred overall. By contrast, Katarina Witt, a world renowned ice skater turned movie actress – certainly not a politician – in her interview of only 45 min, uttered 59 interjections, while her interviewer uttered only two. The imbalance of $59 > 2$ interjections on the part of the interviewee pinpoints the relative conceptual orality of the interviewee's role in relation to the more conceptually literate role of the interviewer. In more concrete terms, while Katarina Witt expressed herself with considerable emotion in an unconstrained manner, her interviewer continued to pose carefully formulated questions in a very serious, objective manner. Her most frequently used interjections were the primary tame interjection *na ja* and the primary wild interjection *pff*, whereas she used only one secondary interjection, *Mensch*. Meanwhile, Gaus made use of only two back-channeled instances of *m-hm*.

Witt's interjections provided an abundant database with which to test a number of hypotheses with respect to privileges of occurrence. For this purpose, Kowal and O'Connell (2004b, p. 85) have examined locations where interjections occurred as well as the durations of pauses preceding and following them:

> (1) *ÄH* belongs to the word class of interjections (Ehlich, 1986); (2) since interjections are not syntactically embedded, their production always involves isolation by preceding and following pauses from accompanying oral utterances (Ameka, 1992b, 1994); (3) the sentence substitutes *ja* and *nein* are functionally interjections when emotionally laden, and in this setting deviate in their phonetic realization from standard forms (Tesnière, 1936).

(1) As it turned out, Ehlich's hypothesis that the German filler *äh* is to be considered an interjection was not supported in the data. The *äh* was not used in the same locations as interjections or with the same pattern and duration of pauses before and after it as interjections. Since its privileges of occurrence did not match those of interjections, the filler could not be considered an interjection. In this respect, our findings are in accord with Nübling's (2004, p. 16) exclusion of *äh* as an interjection and with those of our investigation of the English-language fillers *uh* and *um* (O'Connell & Kowal, 2005b, as reported in our Chapter 13; but see Rasoloson, 1994, for the opposite position). (2) Nor did Ameka's (1992b, 1994) hypothesis that temporal isolation – the occurrence of pauses both before and after interjections – will parallel their syntactic isolation receive support in our data: Less than 20% of the interjections were thus isolated, and more than twice as many (43%) were instead embedded, i.e., with no pause before or after them – the complete opposite of isolation. (3) The third hypothesis could not be tested for *ja*; there were no cases in which its phonetic realization was not standard. For *nein*, however, 25 of 26 occurrences

deviated from standard phonetic realization (mostly as *nee*), and so the hypothesis could be tested: The mean duration of pauses before and after it was much shorter than for interjections (0.96 < 1.51 s). Hence, there was no evidence for considering *nee* as an interjection in this case.

At present, we are continuing our investigation of *yes* and *ja* in the original English and in the translated German versions of audio recordings of Molly Bloom's soliloquy in the last chapter of James Joyce's *Ulysses* (1922/1960). The evidence indicates quite overwhelmingly that it is not a phonological deviation from the standard form of *yes* or *ja*, but the duration and other prosodic characteristics of these items that manifest their shift into interjections, *if and when* they are used for emotional expression – or for what has in this setting been referred to as "a very feminine 'yes' " (Lenoski, 2001, p. 8) – rather than for simple affirmation. Using the case of the affirmative particle *ja* as it shifts into an interjection, Nübling (2001, p. 28 f.) has provided a charming example of a gradual interjectionalization of *ja*, taken from a printed advertisement of a man shaving while trying to terminate a phone call with his mother: "Ja Mutter! – Jaa Mutter! – Jaaa Mutter! – Jaaaa Mutter! – Jaah Mutter! – Jaaahaa. . ."

What we did find from the Katarina Witt interjections, was not only that they reflected her penchant for emotional expression, but that they also served primarily the function of initializing speaking turns and turns of reported discourse, i.e., utterances explicitly designated as quotations. This finding confirms the claim of Reisigl (1999, p. 27) that the location of interjections to initiate, continue, or conclude an utterance is instructive for their functional analysis. At the beginning of turns, Witt's interjections were typically (i.e., 65% of the time) preceded by a pause. This finding suggested that interjections should be understood more as being interposed dialogically at locations between turns rather than within turns.

In order to pursue the matters of emotional expressiveness and initialization further, a dramatic corpus of interjections was sought out. And since George Bernard Shaw's (1916/1969) *Pygmalion* makes Eliza Doolittle's use of a single interjection crucial to his very emotional narrative, O'Connell and Kowal (2005c) have made use of it to compare the written interjections of the play with the spoken interjections of the motion picture *Pygmalion* (Pascal, Asquith, & Howard, 1938) based on the play. They found that the actors spoke unchanged only 30% (57/187) of the interjections in the written text of the play. Moreover, the actors substituted 29 other interjections for the written ones. But what was truly astounding was the finding that the actors added fully 106 interjections on their own initiative in other locations. These substitutions and additions have strongly indicated that, in Goffman's (1981, p. 226) terms, the actors were participating in the role of *author* and transcending the role of *animators* to become *principals*, parties "to whose position, stand, and belief the words attest." Most of the wild or nonconventional primary interjections that were spoken (21/24) were added spontaneously by the actors. Shaw's (1916/1969, p. 11) own use of a seven-syllable written wild interjection as Eliza's signature utterance – "ah-ah-ah-ow-ow-ow-oo" – is both hyperbole and

caricature. The actors in the roles of Eliza and Professor Higgins reduced the corpus of 20 of these signature interjections from an average of 5.1 syl in the written source to an average of 1.3 syl, i.e., generally to a single core vowel. In other words, they defied the written text with its caricature of interjections and instead employed realistic interjections as expressions of their own emotions – an act of authorship. This finding is in accord with Nübling's (2004, p. 24) generalization that spoken interjections are typically short, even monosyllabic. The data have also confirmed the findings of O'Connell and Kowal (2005c) that interjections are used primarily to express emotional involvement and that they fulfill an initializing function: 56% of the spoken interjections in the motion picture version were in initializing positions, compared with only 8% in final positions.

O'Connell et al. (2004, p. 185) have pursued the notion of relative conceptual orality on the part of an interviewee vs. a relative conceptual literacy on the part of an interviewer. As expected of politicians and their interviewers, there were not many interjections: Throughout eight interviews, a total of only 35 were found. Of these, 88% were produced by the interviewees, confirming the conceptual orality found in Katarina Witt's interview. The fact that 71% of the interjections occurred either at the beginning of a turn or before reported speech confirms our hypothesis regarding the initializing function of interjections. In addition, of the 35 interjections, 86% were produced by women. The most frequently used interjection in English was the primary interjection *oh* or the secondary interjection *oh yes*. In German, the most frequently used interjection was *na* or *na ja*, a result that has confirmed the finding for Katarina Witt.

O'Connell, Kowal, and Ageneau (2005) have continued the use of media interviews with a set of six interviews given by Hillary Clinton and a single interview given by the actor Robin Williams. The hypothesis of Ameka (1992b, 1994) that syntactic isolation of interjections would be paralleled by articulatory isolation, i.e., by the presence of pauses both before and after interjections, was once again found to be without evidence: The average percentages of articulatorily isolated interjections in the set of Hillary Clinton interviews and in the single interview of Robin Williams were only 12% and 19%, respectively – hardly characteristic of interjections in general in the corpus. Again, interviewers displayed characteristic conceptual literacy insofar as Hillary Clinton's interviewers used an average of fewer than three interjections per interview compared to Hillary Clinton's 34 interjections, and Lipton used only 11 interjections in his interview with Williams, who used a total of 128 interjections – especially of the nonconventional or wild primary type. Once again, this result has confirmed that actors are more emotionally expressive than politicians. Nonetheless, both Hillary Clinton and Robin Williams as well as all their interviewers used the primary tame interjection *oh* most frequently. The initializing function of interjections was also confirmed: For the same sets of data, the overall percentages of interjections used initially, i.e., "at the onset of various units of spoken discourse," were 71% and 73%, respectively. This finding in turn confirms the generalization of Bres (1995, p. 85 ff.) regarding a

much broader initializing function of interjections in the articulatory realization of a speaker's intentions. According to Bres, an interjection is more immediately accessible at the first moment of articulation than are other verbal units; it occurs rapidly and may lead directly into the realization of a longer phrase. On the other hand, an interjection may stand alone as a sentence substitutable unit, as linguists have long recognized. Nübling (2004, p. 30; our translation) has referred to this characteristic of interjections as "syntactic autonomy and holophrasis." Our own data exemplify both the initialization and the syntactic autonomy. For example, the secondary interjection *Oh God* was clearly used as a sentence substitutable complete turn (O'Connell et al., 2005, p. 169). It is important to note that sentence substitutability is definitely not characteristic of fillers (see our Chapter 13).

In their summary, O'Connell et al. (p. 153) have sought to generalize regarding the use of interjections:

> *The onset or initializing role of interjections reflects the temporal priority of the affective and the intuitive over the analytic, grammatical, and cognitive in speech production. Both this temporal priority and the spontaneous and emotional use of interjections are consonant with Wundt's (1900) characterization of the primary interjection as psychologically primitive. The interjection is indeed the purest verbal implementation of conceptual orality.*

We would add too that wild or nonconventional primary interjections are particularly salient in their affective expressiveness.

O'Connell et al. (2007) have recently compared oral literary readings with artistic performances of actors with regard to their use of interjections in English. They have found the addition of spoken interjections that did not occur in the written text far more extensive in the acting performances than in the readings: 79% > 26%. Again, in Goffman's (1981, p. 226) terms, literary readers are for the most part only *animators,* whereas performing actors are mostly *principals*. As has been the case in our previous research, O'Connell et al. (2007) have found that Ameka's (1992b, 1994) hypothesis regarding the isolation of interjections by preceding and following pauses could not be confirmed; only 39% of all interjections were thus isolated. But the initializing role of interjections could indeed be confirmed: 77% of all interjections were initializing in this corpus of 667 spoken interjections. The primary tame interjections *oh* and *ah* were the only interjections common to all the corpora.

Interjections are also used in a number of other settings. Some instances of what Karl Bühler (1990, p. 176) has referred to as "*empractical* naming and pointing using isolated language signs" are interjections. The murmuring and mumbling that occur as individuals watch TV drama or a sporting event consist partly of interjections (see Baldauf, 2002). And many of the expressions of live audiences by way of approbation and affiliation on the one hand and disapprobation and disaffiliation on the other are also interjections (see our Chapter 18). For example, *hurrah*, *bravo*, *boo*, *"Hear! Hear!*," (*Merriam-Webster's collegiate dictionary*,

11th ed., 2003, p. 574), and many other orthographically and phonetically wild, nonconventional utterances proper to such settings are interjections.

In summary, the empirical research clearly indicates that the use of interjections is both psychologically and socio-culturally relevant as well as constituting an orderly phenomenon worthy of further investigation.

Chapter 15
Referring

Practically all utterances are in a frame of reference for which the speaker's ego provides a center of gravity. Since language invariably involves interpersonal instruction, it is always blended with deixis (Hörmann, 1981, p. 307).

Chapter Prospectus

Chapter 15, *Referring*, calls the reader's attention to a number of indicators that speakers use to refer a listener to someone or somewhere or sometime (e.g., all three, as in *Jose* went *there in June*). We have used the generic term *referring* so as to include in this chapter discussions of deixis, anaphora (frequently a pronominal substitution for a preceding word or phrase), naming, and other forms of designation of objects, places, or persons in the environment or under discussion in spoken discourse. Deixis need not be verbal at all; pointing or even direction of gaze often suffices. Pronominal reference allows of much stylistic variation in choosing person (e.g., *you* do it this way or *one* does it this way) and number (e.g., *I* or an editorial *we*). We review empirical research on the implications of references in various settings (in particular, political). These research examples will especially clarify the importance of referring as an expression of the dialogical perspective of the speaker. The conscious ego of the speaker is at the center of all referring as *origo*.

Referring

Our old friend *Merriam-Webster's collegiate dictionary* (11th ed., 2003, p. 1045) offers the following definition of *refer*, as intended in the present chapter: "**~** *vi* **1. . . b :** to direct attention usu. by clear and specific mention <no one *referred* to yesterday's quarrel>." Or, as *Collins cobuild advanced learner's English dictionary* (2003, p. 1204) puts it: "If you **refer to** a particular subject or person, you talk about them or mention them." The term *indexical* is also used with respect to the action of referring. *Merriam-Webster's* (p. 633) definition of indexical is:

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_15, © Springer Science+Business Media, LLC 2008

"**2 a :** varying in reference with the individual speaker <the ~ words *I, here, now*> **b :** associated with or identifying an individual speaker <~ features of speech>." Accordingly, it is the *index* finger with which one points. In fact, a great deal of the referring in spontaneous spoken discourse is of a nonverbal nature. The pursing or protruding of the lips is yet another nonverbal means of pointing; it is widespread in Central and South America, Africa, and the Philippines. An additional way of pointing nonverbally is with the eyes, either by gaze coupling (looking at each other) or mutual gaze (together looking at the same person or object). One may note that our usage of *referring* always implies a speaker in a concrete setting; however, both speaker and listener may refer – either verbally or nonverbally.

Harley (2001, p. 423) has presented the concepts of *reference* and *referent* that are characteristic of mainstream psycholinguistics: The notions are subsumed under the general area of meaning and semantics – abstractly and without advertence to a speaker, although in his section on comprehension, he has indeed adverted to the listener (p. 322). Dietrich (2002, p. 135) has applied the term "**referentielle Besetzung**" (referential determination) to the speaker's role, but still within the context of theoretical semantics.

For purposes of empirical research, how a speaker refers makes a difference. The *hey-you-there* mode of address is very far removed from the *excuse-me-Ms.* style in a number of dimensions, all of which reflect the dialogical perspective of the speaker. Needless to say, the former example shows little respect; the latter, even though it addresses a second-personal entity, uses a first-personal pronoun. The fixing of the intended object or addressee is quite characteristic of spontaneous spoken discourse. Even when the discourse is relatively abstract, one speaks to the interlocutor, e.g., of *your* principle, or with reference to a third party's involvement as *her* intentions. Referring thus becomes an important means of anchoring the discourse to the here and now.

Deixis

In a more technical, linguistic context, Crystal (1997) has this to say of the concept of *deixis*:

> Every language has a set of lexemes which can be interpreted only with reference to the speaker's position in space or time. These are known as deictic forms (from the Greek word for 'pointing'), and the conditions governing their use have attracted especial attention in recent semantics. They fall into three main types.

- *Personal deixis* The use of pronouns, such as *I* and *you*, which identify who is taking part in the discourse.
- *Spatial deixis* Forms that distinguish the position of the speaker in relation to other people or objects, such as *this* / *that, here* / *there* (p. 99), *bring* / *take, come* / *go*. *Come*, for example, implies direction towards the speaker – *Come here*! (but not **Go here*!).

- *Temporal deixis* Forms that distinguish time with respect to the speaker, such as *now, yesterday, then*, and the various kinds of tense marker.

These are all familiar, commonsensible relationships that speakers and listeners use frequently in everyday life. One might add the use of proper names and roles (e.g., secretary of state and the labor negotiator) as methods of referring to people. But it should be noted additionally that Crystal has referred to "the speaker's position in space or time." The context, then, is spoken discourse, and the fact of the matter is that *I* is always in this setting the *origo* or central point of the universe from which all else radiates – what the classical Greek phrase 'όμφαλος της γης, the navel of the world, intends. William James (1891/1981, p. 278) has called attention to the underlying psychology of *me*:

> The altogether unique kind of interest which each human mind feels in those parts of creation which it can call *me* or *mine* may be a moral riddle, but it is a fundamental psychological fact. No mind can take the same interest in his neighbor's *me* as in his own. The neighbor's me falls together with all the rest of things in one foreign mass, against which his own *me* stands out in startling relief.

The relevance of James's generalization is that referring is always perspectivized in terms of this primacy and centrality of *me* and what is *mine*, as set off from others. It is always *my* perspective from which I set out, as we shall discuss in detail in Chapter 20 under the topic of *having a perspective*. And if, along the way, I adopt or take another's perspective, it is either because it is found to be in accord with mine or because I am trying it on, so to speak, taking the perspective of another tentatively, or for the sake of argument, or to be civil. This *taking, setting*, or *assuming* another's perspective is ordinarily not quite as definitive as it might seem to either interlocutors or analysts; instead it sometimes remains quite tentative and/or superficial. The cunning second-hand-car dealer may not be taking your perspective at all; he is simply professing his "Irish" because he noted your name and saw the shamrock in your lapel. Citation is an additional case of explicitly taking another's perspective, as we have noted in Chapter 9. And as we noted there, the spoken enactment of the written conventions for indicating this shift in perspective is not without subtle risks.

In a similar vein, Karl Bühler (1990, p. 117) has emphasized the speaker as the center or "origin" of the referential field, which is defined by the here, the now, and the I or ego. Hörmann (1981, p. 242) has summarized Bühler's argument as follows:

> The groundwork of language communication is constituted neither by the lexicon as a register of symbolic designations for things, nor by any set of rules for stringing together these designations, but by the ego engaged in an incessant confrontation with a world of human beings.

This anchoring "in the conscious ego as the *origo* of the phenomenal field," at the intersection of *me*/*here*/*now*, is reflected in deictic words, which "are

expedient ways to guide the partners" (Bühler, 1990, p. 121). Thus, the unitary self-consciousness of the speaker constitutes the center of the communicative space from which referring emanates (Hörmann, 1981, p. 115; see also Wunderlich, 1972, p. 81).

Laughter as a Nonverbal Self-reference

It could well be argued that *emotional expressions* reflect especially well the primacy and centrality of the *me*. Edmondson (1987, p. 29 f.) has emphasized in this regard that laughter is "both personal and relational," and that "it belongs to the general domain of the pronominal." Thus, laughter is shown to be self-revelatory: "The first person pronoun is an inevitable core of what it means to laugh." And yet, revealing as laughter might be of the ego, it can be potentially very ambiguous, as "in neutral, noncommittal, tentative or interrogative laughter." Furthermore, Edmondson has maintained that laughter is "distinguishable from other paralinguistic pronominal declarations, such as cries of grief and pain by its sociability: it is a direct appeal for mutuality." Accordingly, he has distanced laughter from "the autism of weeping, surprise, expletive and command" and has aligned it instead with

> the interjections of cheering and booing, of socially shared joy and anger. Laugh and the world laughs with you – for to laugh alone (or to be the sole non-laugher in a group) is a form of temporary ostracism with immediate person significance.

All this self-referential function of laughter adds up to the implication that it is "a declaration of individuality, possibly the most individualized of the human uses of sound." In any event, laughter certainly must be acknowledged to be strongly personally perspectivized.

Some Recent Research on Referring

In a recent volume edited by Enfield and Stivers (2007), *Person reference in interaction: Linguistic, cultural, and social perspectives*, Stivers, Enfield, and Levinson (2007) have reviewed the literature and "universal principles that govern this domain" (p. 2). They have also characterized the research in the book as emphasizing a central issue: the question "why some particular mode of reference rather than another has been chosen" (p. 6) in various empirical situations. Unmarked referring to persons is accomplished in English with an individual's name, whereas through marked referring, "speakers perform actions relative to the culture in which they operate. . . . It is the departure from the unmarked form that conveys that the speaker is doing something special with the action" (p. 18). For example, *Colin Powell* would be an unmarked reference to the former secretary of state of the United States; *Secretary Powell* would be a marked reference. In short, the empirical projects reported in this

book indicate that referring is not just "about giving and receiving information but about navigating social relations" (p. 19).

In Chapter 5, we have presented empirical evidence for a shift toward a more conversational style in U.S. presidential inaugurals of the late nineteenth century through the use of various rhetorical devices (Kowal et al., 1997). Thus, a shift in the use of first-person pronominals from the dominant use of the singular *I* to the plural *we* can now be seen in a more specific light: It reflects a shift in referential perspective to a more informal public self-presentation on the part of the presidents themselves.

More recently, Suleiman and her colleagues have engaged referring in the context of political discourse, particularly in the context of the Middle East conflict. Suleiman and O'Connell (2003, p. 419) have found considerable differences in the way in which Colin Powell, in a TV interview with Larry King, has referred to parties involved in the Middle East: "Powell's discourse divided the world into 'us' – the USA and its allies (Russia and Israel) – and 'them' (Iraq, Palestine, and Terrorists)." Atkinson (1985, p. 167) too has called attention to the use of *us* and *them* in the rhetoric of political speeches. He has found the use of *us* and *them* one of the most frequent rhetorical contrasts to effectively elicit audience applause (see also our Chapter 18).

Suleiman, Lucas, Blum-Kulka, Kampf, and Liebes (2001, September) have observed notable differences in the way in which the CNN and BBC television networks have reported events in the Middle East in terms of amount of coverage. More specifically, in referring to the intifada during October 2000, CNN afforded Israelis more TV time: "Across all categories of measurement – appearance, turns, number of individuals – Israelis predominated." Suleiman et al. have inferred perspectival imbalance from the imbalance of time coverage. They have found no such imbalance in the BBC coverage. More recently, Suleiman and O'Connell (2007) have pointed out Bill Clinton's perspective in media interviews, as shown in his way of referring to the Israeli and Palestinian points of view by designating them with first-personal and third-personal pronominals, respectively. Again, the *we* and *they* mentality prevailed – good guys and bad guys. Finally, notable differences in the way in which Hillary and Bill Clinton referred to their interviewers in their media interviews (Suleiman & O'Connell, 2008) have been found. Hillary Clinton addressed all five of her interviewers (male and female alike) with their first name, whereas Bill Clinton addressed none of his interviewers in this manner. It worked the other way around too: None of the interviewers addressed Bill Clinton with his first name. However, Hillary Clinton was addressed with her first name by three of her five interviewers.

These findings can be considered in light of a broader research context. In general, women have been found to use language more properly than men in keeping with the occasion (e.g., Labov, 1972; see also Bourdieu, 1991). It may be that Hillary Clinton simply wished to build camaraderie with her interviewers. Along these lines, Cohen (1987, p. 122 f.; see also Suleiman & O'Connell, 2007) has observed that interviewees use first names with their

interviewers rhetorically to give the illusion that they are closer in position or perspective to the interviewers than they really are. One American journalist reported to Cohen regarding interviewees' use of first names: "Sometimes people try to rub off against your credibility by showing some intimacy that is simply not there." But as for the interviewer's use of the first name of the interviewee, British reporters viewed the practice as destroying "the distance that an interviewer should have from an interviewee" and as implying "an over-familiarity between the interviewer and the respondent."

Referring is clearly not a neutral tool, but instead serves as a revelation of the speaker's perspective toward the person or institution or object of his or her discourse. Wagner, Kako, Amick, Carrigan, and Liu (2005, p. 639) have summarized the functioning of anaphora in spoken discourse as follows: "Discourse anaphora is rule-governed, but the rules make reference to more than just linguistic structures." The relationship between a news interviewer and a political interviewee is typically a very sensitive and subtle one. The way they refer to one another can therefore be very revealing. Perhaps that is why one German reporter replied – when Cohen (1987, p. 122 f.) suggested that interviewers do indeed sometimes use an interviewer's first name – with an emphatic and incredulous: "What?! Me?! Never!!!" The cultural differences between American and European reporters in this regard have been pinpointed in Cohen's final statement to the effect that the consternation at his suggestion "was unreservedly shared by all the reporters I spoke to in Germany," whereas he found the least objections to the practice among American reporters.

The empirical findings of Suleiman and her colleagues have clearly added substantive evidence for Enfield's (2007, p. 119) generalization: "Formulations of person reference in any language system . . . make publicly overt and thereby instantiate and stabilize cultural values about persons and their social relations."

The unavoidable usage of perspectivized personal reference in spoken discourse can also be considered in view of Clayman and Heritage's (2002) claim that public news interviewers remain professionally neutral with respect to their interviewees. In light of the ubiquity of perspectival reference to interlocutors, such neutrality seems implausible (see also O'Connell & Kowal, 2006b, p. 163).

Chapter 16
Turn-taking

Everyone in the van was talking, talking loudly and at the same time, nearly shrieking, which is how Afghans talk (Hosseini, 2003, p. 72).

Italians interrupt one another.... Americans speak in turns (Eco, 1986, August 29, p. 51; our translation).

Chapter Prospectus

Chapter 16, *Turn-taking*, moves us to a new level of analysis. Chapters 11 through 15 – *Pauses, Prosody, Fillers, Interjections*, and *Referring* – all have made the individual speaker somehow the focus of our considerations, even though all these variables have their place in a dialogical setting. In the long run, such an undifferentiated focus would undermine the interactive, dialogical thrust of the book as a whole; for spontaneous spoken discourse is essentially interactive. Now we turn our attention to the sequencing of speakers and listeners in dialogue. There is indeed a complex interactive sequencing in real time, involving smooth exchanges, pauses, overlaps, successful and unsuccessful interruptions, simultaneous speaking, and back-channeling. The specifics of moment-to-moment sequencing are determined by both the intentions of the interlocutors and the socio-cultural setting in which they are embedded. Methodologically, any discussion of turn-taking in the disciplines that deal with language use must take into account Sacks, Schegloff, and Jefferson's (1974) classic article "A simplest systematics for the organization of turn-taking for conversation." But it is important to emphasize that there is – independent of the socio-cultural setting – neither a single *normal form* nor a single *ideal form* of sequencing of speakers and listeners in dialogue.

Categories of Turn-taking

Before we engage either the historical or the current empirical literature on turn-taking, the various categories of possible dialogical sequencing of verbal material in real time must be clearly set forth. We present these categories in Figure 16.1.

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_16, © Springer Science+Business Media, LLC 2008

Figure 16.1 Possible Temporal Sequences of a First Speaker's (A) and Second Speaker's (B) Turns; Horizontal Bars (----) Indicate On-time, Triple Ellipses (. . .) Incompletion

Temporal Sequence		Sketch of Temporal Sequence
1. Smooth without Pause	A	----------
	B	----------
2. Smooth with Pause	A	----------
	B	----------
3. Overlapping with Completion by A	A	----------
	B	----------
4. Successful Interruption of A by B		
(a) Without Overlap	A	----------. . . . .
(without Pause)	B	----------
(with Pause)	B	----------
(b) With Overlap	A	----------. . . . .
	B	----------
5. Back-channeling While A Speaks	A	----------
	B	--------
6. Simultaneous Speaking	A	----------
	B	----------
7. Unsuccessful Interruption of A by B	A	----------
	B	----------. . . . .

The rationale for our definitions is temporal; a turn is herein limited to the words spoken, without regard for prosodic or nonverbal elements. Although our presentation emphasizes the dialogue of two interlocutors ($n = 2$), similar principles of turn-taking can be applied to dialogue with more than two interlocutors ($n > 2$).

The most straightforward sequencing in time is the completion of a first speaker's (A) turn and the immediate uptake of speaking by a second speaker (B). This is a *smooth exchange without pause* (1.); it is sometimes referred to in the literature on turn-taking as *latching*. A pause may also intervene before the second speaker begins to speak; this sequence is a *smooth exchange with pause* (2.). If the second speaker encroaches upon the time of the first speaker's words, but allows him or her to complete the turn, it is *overlapping with completion* (3.). If completion on the part of the first speaker is not allowed, the category is a *successful interruption* (4.). This may be accomplished either (a) without overlap – without or with pause – or (b) with overlap. *Back-channeling* (5.) on the part of the second speaker may occur intermittently during the first speaker's turn. *Simultaneous speaking* (6.) of the first and second speaker – i.e., with equal access to the floor – can involve the same

text, in which case a choral quality may emerge, or different text, in which case a conflictive, competitive quality may emerge. And finally, the *unsuccessful interruption* (7.) involves the initiation and breaking-off of speech on the part of the second speaker while the first speaker continues (or resumes) and completes his or her turn.

Most of these categories seem to be intuitively understandable. However, category 4, successful interruption of A by B, may require some explanation insofar as a successful interruption can occur without any overlapping at all, if the first speaker is pausing but has not yet completed his or her turn. Category 5, back-channeling, may also occur during a pause on the part of the first speaker, but such back-channeling is not usually considered by researchers to be a turn on the part of the second speaker, since there is no intent evident on the part of the second speaker to take the floor away from the first speaker.

The fact of the matter is that, both historically and currently, *none* of these categories have enjoyed a commonly accepted operationalization in the literature on turn-taking. The crucial concepts that have found their way into this wasteland include: turn, back-channel, interruption, completion, and pause. The search for the definitive operational definition of all of these concepts has been in vain. The overall consequence of this situation is that the comparability of findings from extant empirical studies is diminished. This limitation of comparability is also to be found for the functional interpretation of the various categories of turn-taking given above in Figure 16.1. A case in point is taken up in a section below in this same chapter under the title "Turn-taking, Gender, and Power": The interpretation of interruptions as the exercise of (masculine) power over women is discussed there.

Is There a Normal or an Ideal Form of Sequencing?

We began this chapter with brief epigraphs from two international litterateurs, Khaled Hosseini and Umberto Eco. The complete passage from Eco (1986, August 29, p. 51; our translation) reads as follows:

> Italians interrupt one another. Everybody gets all excited and tries to make his views prevail by preventing the other from speaking and by trying to prove that he is a fascist or communist. Americans speak in turns. (It is no accident that the pragmatic theory of 'conversation turns' originated in the United States. Italian researchers who write articles about this matter treat it as an excavation from Mars.) Each offers his opinion and says to the other (whom he considers to be an idiot): 'of course I understand your point of view. . .' At the stroke of nine everybody stops and goes home.

Granting Eco's penchant for hyperbole, we do find many cultural alternatives to smooth transitions between turns. Tannen (e.g., 1983) has pointed out family situations in which everyone seems quite comfortable with talking

simultaneously and interrupting a great deal. Ervin-Tripp (1979) has cited a case of Indian-English colloquial conversations in which unsmooth turn-taking is used to emphasize positive affect:

> Interruptions are not taken as offensive behavior. Instead, they are interpreted as a gesture of cooperative talk, a way of supplying more information and an indication of attentive listening. Moreover, everybody is interrupting each other. Not only that interruptions are *in*-offensive behavior, they are also *expected in* some cases by the speaker. (Agrawal, 1976, p. 68; cited in Ervin-Tripp, 1979, p. 395)

According to Kohonen (2004, p. 7), "the tolerance of overlaps and interruptions is high in the French communication style." She has also contrasted the French with the German style that might interpret as aggressive what the French would consider "lively and a sign of active participation." At the other extreme, Philips (1976, p. 91) has characterized the conversational style of one American Indian group as manifesting "absence of a requirement for immediate response."

For Cutler and Pearson (1986, p. 139), however, "the turn-taking procedure, i.e., the smooth interchange of speaking turns between conversational partners" has become mandatory:

> For a conversation to function successfully, each speaker's turn should not go on too long, and should be accomplished without interruption; and at the end of one speaker's turn another speaker should take over without too long an intervening pause.

In a similar vein, Wilson and Wilson (2005, p. 957) have quite recently asserted: "Turn-taking is ubiquitous in conversation and is the normal case against which alternatives, such as interruptions, are treated as violations that warrant repair." We would note that interruption *is* a form of turn-taking, not an alternative to turn-taking, and, as indicated above, it may in some cultural settings indeed be "the normal case." One may also find in such prescriptiveness a reminder of the written language bias. For example, in the written mode, sequential text may not be reduced to the same space on a page – the spatial analogue of temporal overlap in speaking – without losing its identity or at least legibility. But there is no absolute physical limitation on spoken discourse with respect to the dimension of time: One may be able to process brief input from an interlocutor while speaking, although it may be difficult to process complex or extended input under such conditions. Hence, if the model of well-formed dialogical discourse becomes smooth exchange, it does so as a function of a socio-cultural sense of appropriateness and politeness, not as a physical requirement "for a conversation to function successfully."

As we have stated already in Chapter 7, the criterion for a successful dialogue must be whatever the interlocutors wish it to be and decide for it to be. The requirement of "smooth interchange of speaking turns" is not at all a necessary criterion for a successful conversation or dialogue. For example, a political TV debate between candidates may have as its finality not smooth exchange, but the presentation of the opinions and positions of the candidates for the viewing audience of voters. Whether smoothness or its opposite is more conducive to effective communication in this setting remains an empirical question.

Some History

Before we proceed to discuss empirical studies of turn-taking, we must note for the reader that there is one reference to which all other modern research must have recourse because of its historical import: Sacks, Schegloff, and Jefferson (1974, p. 696) set out "A simplest systematics for the organization of turn-taking for conversation" to provide a formal description of the generalizable characteristics of turn-taking, and the mandatory rules that govern it. Some 16 years later, O'Connell et al. (1990) published a comprehensive critique thereof. This will serve as the basis of much of our presentation of Sacks et al. (1974).

One should note from the outset that Sacks et al. did not proceed empirically but rather formalistically:

> Instead of investigating in their own right the variables relevant to turn-taking, such as politeness and cultural norms, probabilistic speaker and hearer cues, expectations, motivations, purposes, and situational exigencies, the turn-taking research tradition has introduced a confusing array of purely formalistic terms such as signals, rules, devices, procedures, and systems under the general aegis of the "turn-taking procedure." (O'Connell et al., 1990, p. 347)

Our detailed criticisms (p. 346 ff.) of Sacks et al. (1974) can be considered under two major headings: their *assumptions* and their *operationalizations of crucial concepts*.

Among their problematic assumptions are the following: There exists some ideal and mandatory organization of turn-taking; a purely formal approach to the turn-taking system is feasible; the metaphor of an economy of time – not using time, but saving time – is appropriate to such an approach; the locations of transition-relevance places (TRPs), i.e., the end of a turn-constructional unit at which "speakers may change" (Levinson, 1983, p. 297), are projectable on the part of listeners and researchers from syntactic structure alone; "Conversation should be considered the basic form of speech-exchange system" (p. 730); and "Someone's turn must always and exclusively be in progress" (Goffman, 1964, p. 136; see Sacks et al., 1974, p. 697).

Similarly, many of the concepts introduced or made use of by Sacks et al. (1974) lack appropriate operational definition, conceptual clarity, and/or empirical foundation. These would include, as O'Connell et al. (1990, pp. 350–354) have noted:

(1) The *transition-relevance place* (TRP) and *projectability* (Sacks et al., 1974, pp. 702 ff.) are defined syntactically. But there appears to be no evidential basis for the assertion that these concepts, thus defined, have any relevance at all "from the perspective of the interlocutors" (Lenz, 1988, p. 150; our translation). In addition, Ellis and Beattie (1986, p. 194) have essentially changed these definitions by including as their "distinguishing features" a number of "verbal, prosodic and kinesic elements."

(2) Strangely enough, the concept *turn* itself is defined in terms of being somehow a right and an obligation. This vague definition of turn allows of strained identification of turns such as the following example – the initiation of a phone conversation – claimed by Schegloff (1979, p. 37) to constitute *three* turns:

Example 16.1
C: Hello?
Y: Hello Charles.
(0.2)
Y: This is Yolk.

Levinson (1983, p. 327) has argued that the first contribution of Y is "the first part of an adjacency greeting pair." His argument is that the second Y is a repair presented "after a significant pause has developed" (p. 328), during which C failed to respond. Levinson's argument relies on an estimated pause duration of 0.2 s. Apart from the impossibility of accurate estimation of pauses of this magnitude, such a pause is extremely short, even well below the radar screen of cut-off points for much of the classical literature on pause durations – hardly "a significant pause." In other words, the logic of turn conceptualization as expressed in this example is seriously flawed.

(3) No criteria are provided to distinguish a turn of short duration from back-channeling that is not considered a turn.
(4) This lack of criterion itself affects who is to be considered to be the speaker during a brief instance of back-channeling and to whom a silence is to be allocated, since, in accord with one of the assumptions mentioned above, it must always belong to one of the interlocutors.
(5) Both overlaps and interruptions are considered to be disruptive elements – "errors" and "violations" (Zimmermann & West, 1975, p. 115; but see Oreström, 1983, pp. 161 and 165, and McTear, 1985, p. 161).
(6) Perhaps the weakest concepts of all in Sacks et al. (1974) and their followers are the various temporal categories for off-time: *pause, gap, silence, lapse,* and even the null temporal category of *latching*. None of them are specified by physical measurement, but subjectively and perceptually. For example,

> A silence after a turn in which a next has been selected will be heard not as a lapse's possible beginning, nor as a gap, but as a pause before the selected next speaker's turn-beginning. (Sacks et al., 1974, p. 715)

No empirical evidence accompanies this generalization; it remains an empirical question. Nor is it at all clear by whom these categories "will be heard"; it is quite uncertain whether they might be categorized identically by researchers, experimental subjects, the speakers and/or listeners involved in the conversation, or others. This failure to use physical measurements of time reflects an underlying failure to acknowledge that researchers, insofar as they are native speakers of the language under investigation, are subject to similar perceptual

biases as are the interlocutors. This limitation of methodology has been and remains quite widespread (e.g., Bergmann, 1982; Ervin-Tripp, 1979; Garvey & Berninger, 1981; Jefferson, 1973; Psathas & Anderson, 1990; Tannen, 1986). If indeed "transitions are finely coördinated" (Sacks et al., 1974, p. 47) and there is "precision timing" (Jefferson, 1973, p. 47), then off-time as well as on-time categories must be subjected to instrumental physical measurement. We hasten to add that stopwatch measurement is also thoroughly inadequate for the assessment of sub-second temporal magnitudes (for further detail, see our critique of Schegloff [2007] below and our Chapter 11).

There are additional problems with the approach of Sacks et al. (1974). Both Oreström (1983) and Power and Dal Martello (1986) have objected to the anecdotal use of data characteristic of the conversation-analytic research tradition. Nor has a solution that is both ethical and methodologically sound been devised for deriving data in this research tradition. Crystal and Davy's (1969, p. 96) principle that surreptitious recording is the "only safe way of obtaining data" must be rejected. In general, the two extremes of surreptitious and in-your-face auditory recording must be avoided at all costs. And finally, the dynamics of spontaneous spoken discourse with $n > 2$ interlocutors has hardly been touched by Sacks et al. (1974); fully 71% of all their anecdotal examples are cases of $n = 2$ interlocutors. A quarter of a century ago, Oreström (1983) summarized his objections to Sacks et al.: They had neglected the relevance of contextual factors for turn-taking, the probabilistic nature of turn-taking cues, and the effect of emotional involvement on turn-taking. O'Connell et al.'s (1990, p. 363) summary critique of Sacks et al. was the following:

> Their approach is neither sociologically nor psychologically enlightening. The purely formalistic, anecdotal speculations, unsupported by temporal measurement, by comprehensive presentation of corpora rather than mere examples, or by appropriate descriptive and inferential statistics, were destined from the beginning to fail.

Nonetheless, conversation-analytic researchers have canonized Sacks et al. (1974); uncritical acceptance thereof is still today quite common. But, despite all the criticisms that have been voiced above, one must acknowledge that Sacks, Schegloff, and Jefferson have, more than any other individual or group, been responsible for making turn-taking a focal concern for all the disciplines related to dialogical language use. The heuristic value of their article has made it an historical landmark.

Over the decades, there has been a search for *the definitive unit of analysis* in the turn-taking setting. This search must be considered illusory, as the following chronological litany of some definitive units of spontaneous spoken discourse suggests: some "independent method of defining encoding units" (Maclay & Osgood, 1959, p. 23); the phonemic clause (Boomer, 1965); on–off patterns (Jaffe & Feldstein, 1970, p. 17); a "turn-constructional component," which is a syntactically defined unit type (Sacks et al., 1974, p. 702); "intonation-marked" phonemic clauses (Duncan & Fiske, 1977, pp. 166 ff.); "*idea units*" (Chafe, 1980a,

p. 48); and the tone unit (TU), "a point of prosodic completion" (Oreström, 1983). Beattie (1978, p. 33) had previously used as unit of analysis "speech units with demonstrable cognitive significance, and possible interactional importance"; these had to be at least 30 s in duration. A noteworthy – and astounding – consequence of Beattie's choice of unit is that it would have excluded fully 95% of the turns analyzed by Oreström (1983) from data of the London-Lund corpus of English.

Schegloff (2007)

Recently, Schegloff (2007) has published a volume entitled *Sequence organization in interaction: A primer of conversation analysis I*. Therein, the organization of "talk-in-interaction" (p. xiii; the term he prefers to "conversation") remains essentially unchanged from the traditional Sacks et al. (1974) version. The reader might note that none of the many voices of criticism that have been presented over the years have even been mentioned in this newly published volume, and none of their arguments seem to have been considered at all relevant. In fact, most of the references in Schegloff's bibliography are from the "conversation-analytic community" (p. xii). Hence, we wish here to offer some comments about the book from outside that community.

Before we proceed to our critique of Schegloff (2007), we must note a very welcome component of his presentation – a procedure that can radically change the way data are managed and evaluated in research texts in the future: He has offered his readers a rare opportunity in his Appendix 2 (p. 270), in which he has presented an internet address (http://www.cambridge.org/9780521532792) where one can find the audio (and video, where relevant) recordings that correspond to the transcripts used as examples throughout the book. We have made use of this material to re-analyze two of his transcripts by means of the PRAAT software (www.praat.org): 1.01, the first transcript in the introductory chapter (p. 23) and 14.01, the first transcript in the summary chapter (p. 253 f.). The rationale for our choice of these two transcripts includes their salient use by Schegloff, their reasonable duration for presentation in this book, and their derivation from audio recordings that are easily decipherable.

The outcome of these re-analyses is that both 1.01 and 14.01 are inaccurate in a number of respects that conversation-analytic researchers themselves consider essential to the validity of their research. The most notable of these inaccuracies can be listed as follows:

Transcript 1.01

(1) The introductory phrase actually consists of three voiced syllables, not unvoiced inhalation followed by one voiced syllable.
(2) The following "It is" is actually articulated as "It's."

(3) The question "Would somebody like some more ice tea" is articulated as "s'more" instead of "some more."
(4) The notation for the question in (3) "((* = voice fades throughout TCU))" should be just the opposite: The acoustic intensity of the second part, "s'more ice tea," is notably and consistently greater than that of the first part.
(5) In the second part of the adjacency pair, "some more" should again be "s'more."
(6) Both pauses that occur are underestimated: ($0.8 < 0.95$ and $0.4 < 0.78$ s).

Transcript 14.01

Since transcript 14.01 is quite long (68 lines), inaccuracies are noted herein by line.

> Line 1: The second occurrence of "thee" is actually pronounced "the". "ABC" should be "at ABC."
> Line 6: The "um:" at the beginning is actually longer in duration than the "u:m::::" in line 10 ($0.70 > 0.66$ s).
> Line 10: The line ends with a semi-colon, a notation not included among the transcript symbols given in Appendix 1 (pp. 265–269).
> Line 22: The substitution of "now" for "not" completely reverses the intended meaning of the utterance and thus distorts the meaning of the entire transcript.
> Line 35: The utterance in this location is marked simply as inaudible, but it is quite audibly articulated as "dutch" and is important to understand the meaning of "Sure" in the following line.

Insofar as the conversation-analytic tradition has always insisted on the importance of accurate estimates of temporal phenomena, a re-analysis of the pauses noted in transcript 14.01 is also warranted. We have divided these pauses into three categories: (1) The pauses designated with an estimated duration (p. 253 f.) are 8 in number, with a mean measured duration of 0.50 s. However, in every instance, the perceptually estimated duration is shorter than our physical measurements for these intervals, and the mean estimated duration is correspondingly shorter than the mean measured duration by one third of a second ($0.50 < 0.84$ s). (2) The pauses designated by "(.)" as "micropauses" – "hearable but not readily measurable; ordinarily less than 0.2 second" (p. 266) – are 4 in number. The measured mean for these intervals is 0.52 s. Note that this measured mean is greater than the mean duration of those pauses that were actually estimated ($0.52 > 0.50$ s) and is therefore not in accord with the operational definition given by Schegloff himself for a "micropause." (3) The third category is in many respects the most remarkable; the number of measurable pauses overlooked in the transcript is 26, with a mean duration of 0.32 s and a range of 0.10–0.83 s.

Table 16.1 sums up the descriptive statistics for the three pause categories as well as for the 8 pauses for which estimates are actually given in transcript 14.01. Pauses were measured to a minimal cut-off point of 0.10 s. One can safely say that transcript 14.01 simply does not accurately reflect the spoken dialogue it is

Table 16.1 Means (in s), Standard Deviations, and Number of *Measured* Pause Durations for Omitted Pauses, Micropauses, and Estimated Pauses in Transcript 14.01 of Schegloff (2007, p. 253); *Estimated* Pauses from the Transcript in Parentheses after their Measured Durations

	Categories of Pauses		
Descriptive Statistic	Omitted	Micro	Estimated
Mean	0.32	0.52	0.84 (0.50)
Standard Deviation	0.17	0.13	0.38 (0.28)
Number	26	4	8 (8)

supposed to represent. To personify for a moment, a transcript that does not know what it is saying cannot serve as the basis for a scientific explanation of talk-in-interaction. On the last page of his summary chapter, Schegloff (p. 264) himself has described his presentation of sequence organization in interaction as something he has "made to sound like an assembly line." This well may be the case, but the presentation does not explain the striking differences between our analysis of the data and Schegloff's. In any event, perceptual estimates would appear to be quite inadequate in comparison with careful physical measurements of temporal phenomena (see our Chapter 11). In his final statement, Schegloff has tried to throw some light on where he thinks this traditional approach, the "adjacency pair technique" to sequencing in dialogue, has arrived:

> The adjacency pair technique, in providing a determinate "when" for it [close ordering] to happen, i.e., "next", has then means for handling the close order problem, where that problem has its import through its control of the assurance that some relevant event will be made to occur. (p. 264; cited from Schegloff & Sacks, 1973, p. 297)

But as we understand this citation, nothing in the doctrine of adjacency pairing effectively engages "handling the close order problem." It is obvious from our re-analyses of representative transcripts (1.01 and 14.01) that the pauses notated therein cannot account for the "when" of the "next": They are – all of them – simply inaccurate. The "when" and "next" of close ordering appear increasingly to be more like mythology than like an assembly line. Instead, "some relevant event will be made to occur" in keeping with the complex intentions and purposes of the interlocutors, not by any pair-adjacency formula. The conversation-analytic community seems to be no closer now to solving "the close order problem" than was Levinson (1983, p. 327) in search of a "significant pause" without any evidence, or than were Schegloff and Sacks (1973) a third of a century ago.

Turn-taking, Gender, and Power

The association of turn-taking with an ideology of gender and power may – in the longer view of the history of a psychology of language use – be not much more than an ephemeral focus, a passing research fad. Nonetheless, it must be

engaged. The collocation of "language, gender, and power" (Samel, 1995, p. 42; our translation) and the specific designation of a *Contrastive feminist linguistic* (Hellinger, 1990; our translation) are formulations that have been used within German linguistic circles. At the same time, the Proceedings of the Second Berkeley Women and Language Conference were being published as *Locating power* (Hall, Bucholtz, & Moonwomon, 1992). McKendrick's (2003, p. 2) more recent summary of the thrust of this research tradition is quite straightforward, even blunt:

> Interruptions are commonly defined as "a violation of speakers' turns at talk" (Tannen [, 1994, p.] 56). This definition, combined with restrictive ideas of what comprises a violation of turn and research that is directed toward a preconceived result, leads to a misleading interpretation of gender differences in conversation. Typically, the authors of studies on interruption equate male gender with status and dominance (Chamblis &[Feeney, 1992;], Kennedy [& Camden, 1983;], Key, 1986; Kunsmann[, 2000; and], Zimmerman [& West, 1975]). Sometimes, the desire to prove this hypothesis can lead to authors concluding that men are more likely to interrupt despite the evidence of their own research (Kennedy [& Camden, 1983, p.]107). Removing the political overtones and considering a different model of conversation can lead to quite different conclusions.

Basing much of her own assessment of the gender research tradition on James and Clarke's (1993) "Women, men, and interruptions: A critical review," Kowal (1998, p. 364; our translation) has warned her readers in her review of Samel (1995):

> By reason of the multiply documented findings to the contrary, the research on gender-typical conversational performance is still a long way off from establishing empirically "the procedures of exercising power . . . or dominance in conversations" (Samel, 1995, p. 145).

Recent Research

The difficulty of obtaining an assessment of interruptions independent of cultural influences has been emphasized by Okamoto, Rashotte, and Smith-Lovin (2002, p. 38). They have suggested that gender and subcultural differences actually affect "*how coders construe interruptions*." In a similar vein, Kohonen (2004, p. 30 f.) has acknowledged that different language groups have different communication styles; these include the rules of turn-taking, overlaps, and pauses. Contrary to her initial hypotheses, Kohonen found that, for a corpus of French conversation, interruptions were one of the least frequent forms of overlapping talk in the corpus (0.9%). Most of the overlaps served instead the purpose of maintaining a natural flow of interaction, of facilitating interaction, and of fostering participation by means of simultaneous starts or turns and shared laughter. Our own nomenclature would not categorize interruptions as a form of overlap, but the point that interruptions occur quite infrequently is well taken.

A brief reference to a recent study of turn-taking is instructive as to the persistence of methodological problems in this area of research. ten Bosch,

Oostdijk, and de Ruiter (2004) have noted that telephone conversations have much shorter inter-turn silences than face-to-face conversations. But they have failed to note that this finding has been in the archival literature for many years (e.g., Norwine & Murphy, 1938; Brady, 1965, 1968). Another, unpublished paper also by ten Bosch et al., posted in the internet, http://www.hcrc.ed.ac.uk/comic/documents/publications/SPECOM_2004_tenBosch.pdf., has argued that the role of both speakers in their corpus of phone conversations must be identical because the conversations are informal. But symmetry of this kind cannot be logically inferred from informality. The same authors have also argued that turn-taking style is an adaptation of interlocutors to one another. They have used Garrod and Pickering's (2004) concept of an alignment process to describe this convergence. The difficulty is that their time data are limited to averages that do not take into account any kind of convergence *over time* and hence remain inadequate as evidence for such an adaptation over time.

Our Research

The categories of sequencing in turn-taking that were introduced by Kowal, Barth, Egemann, Galusic, Kögel, Lippold, Pfeil, & O'Connell (1998) have been applied to a number of English-language and German-language radio and TV interviews. In general, smooth turn transitions with or without pause accounted for most of the turn transitions: 74% in Kowal et al. (1998), 93% in Kowal et al. (1997), and 83% in O'Connell and Kowal (1998). In the same studies, the frequencies of interruptions on the part of interviewers and interviewees were, respectively: $40 > 4$, $3 > 0$, and $14 > 0$. Of these, 80% (49/61) were interruptions on the part of male interviewers. And in turn, 57% of these (28/49) were due to only two of the male German interviewers. However, since 64% (9/14) of the interviews were conducted by men, the confounding of interviewers with males makes it impossible to ascribe the phenomenon of interruption to males as such. Similarly, the paucity of interruptions on the part of interviewees (4 altogether) does not allow a significant gender discernment: Both male and female interviewees interrupted. Finally, we might note that the role of interviewer is, among other things, to manage the temporal constraints of programming; hence, it is hardly unusual for an interviewer to interrupt an interviewee who talks too long or to bring the program to a close in a timely way through an interruption.

A Perspective on Turn-taking

A focus on turn-taking transcends the role of the individual speaker and directs our attention to the interactive nature of spoken discourse. Still, it is worthwhile to note that this shift to the interactive, though undoubtedly appropriate, has not been without its own Scylla-and-Charybdis-like problems: A formalistic approach

has radically disregarded socio-cultural relevance; an American WLB has been prejudicial to other cultures; surreptitious and in-your-face audio recordings have been methodologically at odds with one another; perceptual assessments of time durations have been inadequate in comparison with physical measurements; a conversation-analytic emphasis on a logic of anecdote and example has been in conflict with an insistence on inferential statistical logic; qualitative and quantitative methods have been pitted (unnecessarily) against one another.

Turn-taking constitutes the very moment of overt, spoken interaction; its systematic importance for any empirical study of spontaneous spoken discourse is incontrovertible. Mainstream psycholinguistics has been busily engaged with speaker and listener phenomena in laboratory settings that are minimally interactive in any realistic sense; there has been precious little interest from that quarter in the psychological components of the genuinely interactive settings of spontaneous spoken discourse.

Perhaps, it is high time for a psychology of language use to deliberately walk away from the methodological jungles and start anew to address turn-taking without the baggage of the past. Graumann (2006, p. 61) has suggested that Wundt himself insisted that a methodology had to transcend the physiological and individual preoccupations of the psychology of his day, if the interactive and creative elements of language use were to be engaged. Speaking leaps beyond the individual's psychology in a number of senses: What one speaks becomes the property of the listener and allows him or her not only to interpret, but to change or even falsify what has been spoken. As Graumann has put it, a speaker places him- or herself necessarily under the control of the interlocutor. This is a far cry from the information-transfer approach of mainstream psycholinguistics and suggests that indeed a new approach to the psychology of spontaneous spoken discourse is long overdue.

Chapter 17
Laughter

Voltaire has said that heaven has given us two things to counterbalance the many troubles of life: hope and sleep. He could well have added laughter (Immanuel Kant, cited in Geier, 2006, p. 110; our translation).

Chapter Prospectus

Chapter 17, *Laughter*, pinpoints yet another phenomenon characteristic of spontaneous spoken discourse. In writing, the simulation of laughter is stereotyped as HA-HA, whereas the phonetics of actual laughter covers quite a range of sound (e.g., *he*, *ha*, *hm*, *ho*, or the simple, unvoiced expiration *h*). There has been historically a confusion of the concept because of the identification of humor and comedy with laughter, mostly on the part of philosophers and litterateurs. The empirical investigation of laughter itself – as a behavior – has languished, as has even more so the empirical investigation of its opposite, weeping. Linguistics has categorized laughter as a nonverbal phenomenon and therefore peripheral to its interests; similarly, mainstream psycholinguistics has found little interest in laughter. But much laughter is overlaid on speech itself and therefore carries a specific semantic message along with it. By way of contrast, the conversation-analytic researchers have consistently included laughter as an important component of talk-in-interaction. Laughter can be a subtle, nuanced rhetorical tool in the hands of public speakers, actors, and politicians, and even in the hands of the stereotypic country bumpkin who is master of his or her native language in its spoken form.

Science Times

As we sat down to write this chapter, there was a brand new article on the table that promised to throw light on current research regarding laughter. In the Tuesday *Science Times* section of the *New York Times* (2007, March 13, pp. D 1 and D 6), an article entitled "What's So Funny? Well, Maybe Nothing," by John

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_17, © Springer Science + Business Media, LLC 2008

Tierney explained the latest research on laughter. Typically, *Science Times* articles are based on only a few references (in this case, Provine, 2000 and Panksepp, 2005) and interviews with putative experts. The approach is, of course, popularized rather than scientific, and the outcome of such an approach becomes generalized interviewee comments such as "Laughter seems to be an automatic response to your situation rather than a conscious strategy" (p. D 6). Obviously, the emphasis of the article was on the instinctive: "It's an instinctual survival tool for social animals, not an intellectual response to wit" (p. D 1). But in both these quotations – one by an interviewee and one by John Tierney himself – contrasts are made between nonalternatives. And both sets of alternatives seem to be inappropriate: The alternative in laughter is neither between automatic and conscious, nor between instinct and intellect; neither dichotomy touches the essence of laughter. And the fact noted by Tierney that different people react differently to the same joke in different settings is not entirely new or surprising.

Curious as to the audience reaction, we went to John Tierney's blog and found that 89 entries had been filed in the first 21 h, 43 min to his question: "But Seriously: Why is There a Gender Gap in Laughter?" After perusing the comments submitted by Tierney's readership, we noted that the readers' everyday experiences of laughter did not generally reflect the tenor of his article or of Provine's (2000) research referenced therein. At the very least, these reader responses have suggested a further look at the positions presented in Tierney's article.

The other reference, Panksepp (2005, p. 62 f.), is quite a different story. His final argument, based on his biologically oriented research, seems to be overgeneralized and a bit cute: "Although some still regard laughter as a uniquely human trait, honed in the Pleistocene, the joke's on them." Panksepp and Burgdorf's (2003, p. 533) earlier research had suggested a relationship between their rat data and "primitive human laughter" and perhaps "human childhood laughter." Panksepp's (2005) research is clearly state-of-the-art biology, but his generalizations have little to do with the complicated, human, adult phenomenon that constitutes the behavior we call laughter. Among others, cynical, rhetorical, mendacious, and hypocritical laughter do not reflect the "joy and woe" described by Panksepp. Psychologically, when laughter is accompanied by emotion, it is not necessarily a pleasurable or positive emotion. Laughter is instead a multifaceted observable behavior, ranging from a muffled chuckle to a hugely raucous belly laugh. And this is what our chapter is all about.

Some History

Unfortunately, much of the history of the concept of *laughter* is not about laughter at all, but about wit, humor, comedy, and the ludicrous. Philosophers and litterateurs wanted to know from the very beginning what makes us laugh,

rather than what the behavior of laughter is. Rhetoricians too have been primarily interested in what elicits laughter (e.g., Quintilian, ca. 95/1958). As Seibt (2002, p. 760; our translation) has put it:

> In the three most recent centuries, hordes of philosophers, aesthetes, theologians, litterateurs, and people at large have reflected on the laughable and the comical. Only laughter itself never came into view; the focus remained fixated instead on its source.

Books about laughter have even carried in their titles the contradiction, as in the case of Bergson's (1914) *Laughter: An essay on the meaning of the comic* (original French: *Le rire. Essai sur la signification du comique*, 1900).

Only in the twentieth century has the acoustic, respiratory, and electronic equipment needed for the identification, measurement, transcription, and analysis of the behavior of laughter become available. And gradually, the books that were in fact about humor (e.g., Gregory, 1924; Greig, 1923/1969; Piddington, 1933) have given way to books that are actually about the behavior of laughter (e.g., Glenn, 2003; Partington, 2006; Provine, 2000). In this context, Chafe's (2007) *The importance of not being earnest: The feeling behind laughter and humor* and Martin's (2006) *The psychology of humor: An integrative approach* must both be considered a return to a focus on humor rather than on laughter, even though both books engage laughter as well.

One may note that *Merriam-Webster's collegiate dictionary* (11th ed., 2003, p. 560) has listed *ha-ha* as an interjection "used to express amusement or derision," whose history as part of written English reaches back to before the twelfth century and to the Old English *ha-ha* (see Table 13.1 in our Chapter 13).

Definitions

In view of the confusion surrounding the definition of laughter, we have begun our research with a common-sense definition, as have Bachorowski, Smoski, and Owren (2001, p. 1582): "Any perceptibly audible sound that an ordinary person would characterize as a laugh if heard under ordinary everyday circumstances." Provine (2000, p. 43; see also Jefferson, Sacks, & Schegloff, 1984, p. 18) has stipulated an additional characteristic: Laughter is "the quintessential human social signal" and therefore it "is about relationships." But to the contrary, Glenn (2003, p. 14) has noted that laughter "is both a solitary and a group form of expression." Hopper (1992, p. 179) has added the essential element of exhalation: "Laughter is a sort of involuntary, vocal exhalation that signals amusement, farce, or the unexpected." However, Chafe (2007, p. 22 ff.) has included inhalatory pulses as part of laughter as well; and there is no evidence at all that laughter must be involuntary or that it must be limited to the functions Hopper has listed. Trouvain (2003, p. 2793) has succinctly described laughter as "an every-day, human-specific, affective, nonverbal vocalization." But, on the one hand, this definition includes too much, namely the

phenomenon of weeping in addition to that of laughter, and, on the other hand, it would appear to exclude laughter overlaid upon the speech stream, a form that is suprasegmental, insofar as it is "a vocal effect extending over more than one +segment" (Crystal, 1997, p. 438) and is thus distinctly verbal. In fact, the semantic neutrality of the segmental HA-HA type of laughter must be contrasted with the semantic commitment on the part of laughter overlaid on speech. It should be clear that a psychological investigation of laughter cannot legitimately begin with the assumption that laughter necessarily involves pleasurable feelings, just as it should be clear that a psychological investigation of weeping cannot legitimately begin with the assumption that weeping necessarily involves sad feelings. Provine (2004, p. 215) has defined laughter as: "instinctive, contagious, stereotyped, unconsciously controlled, social play vocalization." We would argue that none of these characteristics is an essential characteristic of laughter as such, and in particular of complex adult human laughter.

Methodology

There are currently three basic research approaches to the behavior of laughter, and in practice, they overlap with one another: those of phonetics, of psychology, and of sociology (or more specifically, sociolinguistics in the context of conversation-analytic research). Until quite recently, all three were concerned with stereotypic HA-HA laughter and not at all with laughter overlaid on speech (but see Duncan & Fiske, 1977). The latter type occurs, however, frequently and in both spontaneous and nonspontaneous spoken discourse. That is to say that it can occur as an overlay on the articulation of passages read aloud as well as in ordinary conversation and other spontaneous spoken discourse. Both types of laughter can be conveniently unitized into calls and bouts. Calls (C) are pulses analogous to and co-extensive with syllables, whereas bouts (B) constitute a continuous sequence of calls. Hence, calls per bout (C/B), calls per syllable spoken (C/syl), and calls per second (C/s) become relevant normalized measures of the frequency of occurrence of laughter. Such measures as these are necessary if a quantitative assessment of laughter is to be made.

It should be noted that these units are not synonymous with Provine's (2000, p. 26) "*laugh episodes,*" defined by him as follows: "A laugh episode consists of the comment immediately preceding laughter, and all the laughter occurring within one second after the onset of the first laughter." Provine wanted to investigate whether the relationship of HA-HA laughter to humor is tenuous. And he (Provine, 2004, p. 215) did indeed find that "only 10–15% of prelaugh comments were remotely humorous"; for the rest, he found that they were "banal comments." However, his methodology (Provine, 1993, p. 293) was not adequate for such a broad generalization. He had sent his graduate students

out into the streets of Baltimore with clipboard in hand. Their observations were surreptitious and from a distance, without benefit of audio recordings or interviews of the observed laughers. Apart from the ethical problems of such research, an important methodological problem was the short duration of the probes, which were limited to "within 1 s preceding the onset of the initial laughter." O'Connell and Kowal (2005a, p. 286) have provided an example of an instrumentally measured 1 s probe preceding the onset of Bill Clinton's (underlined) laughter overlaid on speech: "*that uh yeah* <u>*that's right*</u>." Clearly the duration of the "prelaugh comments" (Provine, 2004, p. 215) in this case is too short for the researcher to know anything about the topic at hand. A judgment about either the banality or the nonseriousness of the prelaugh verbal context would appear to be beyond the scope of a 1 s probe.

Psychological Approaches

Already in the nineteenth century, there were some empirical approaches to laughter. Hecker (1873) had attempted a psychophysics of laughter. G. Stanley Hall (Hall & Allin, 1897) had engaged laughter in conjunction with tickling and the comical. But the twentieth century was not exactly an oasis of laughter research. Washburn's (1929) study of infants' smiling and laughing behavior in the first year of life is an early landmark. Duncan and Fiske (1977, p. 80) recorded many hours of *Face-to-face interaction* and concluded that "a laugh was a fairly rare event" therein, but that "the context of each laugh deserves investigation," especially since a partner's laughter proved to be a frequent accompaniment. And as recently as 1996, there has been research that has had to do not with laughter, as the author's title "Faculty Laughter" indicated (Coser, 1996; our translation), but with the humor of jokes.

The advent of the twenty-first century brought with it a serious interest in laughter itself. Provine's (2000; see also Provine, 2004) was the first book on laughter to appear in the new century. There and in his 2004 article, he has brought together his research findings. He has contended that laughter generally follows upon nonhumorous and banal comments and that it occurs in accord with syntactic junctures:

> Laughter punctuates speech. Laughter of both speaker and audience occurred at the end of phrases or sentences. This punctuation effect was extremely strong; a speaker's laughter interrupted speech in less than 1% of laugh episodes. The prominence of this effect is so striking that it may be confirmed by cursory observations of social conversations. (Provine, 1993, p. 296)

Furthermore,

> Laughter is spontaneous and relatively uncensored, thus showing our true feelings. Laughter, like crying, is difficult to produce on command and, therefore, an honest signal. We cannot deliberately activate the brain's mechanism for affective expression – laughter is an unplanned response to social, cognitive, and linguistic cues. (Provine, 2004, p. 216)

Further Questions about Psychological Research

Our own research in the area of laughter started from our questions about the generalizability of Provine's results – and indeed from our questions about their basic plausibility in some instances. Washburn's (1929) research had clearly indicated stereotypy in infants' laughter. But it was our assumption that, in the course of intervening years, the adult human has superimposed a huge amount of learning and experience upon the simple infantile phenomenon of laughter. And so, we wished to have a further look at adult laughter in light of these years of experience. Because of their articulateness, we chose to concentrate our research on the laughter of both politicians and actors.

Our first two studies (O'Connell & Kowal, 2004a, 2005a) were closely related to one another. After the publication of their respective memoirs (H.R. Clinton, 2003; B. Clinton, 2004), both Hillary and Bill Clinton had scheduled a series of media interviews to advertise their books. The corpus of Hillary Clinton's interviews included 4 TV and 2 radio interviews; the corpus of Bill Clinton's interviews included 11 TV and 2 radio interviews. Fortunately for purposes of comparison, 3 of the TV and 2 of the radio interviews were with the same interviewers (Couric, King, and Letterman [TV], and Gross and Williams [radio]).

In a number of basic respects, the results of the two studies were in complete accord: There was no evidence of Provine's (1993, p. 296) "extremely strong" punctuation effect, and laughter did not generally follow either banal comments (Provine, 2004, p. 215) or nonserious comments (Chafe, 2003a, 2003b; 2007). Our analyses indicated that the most common topic to elicit laughter on the part of both Hillary and Bill Clinton was her possible candidacy for the presidency of the United States (for further comments on this research see Partington, 2006, p. 20). In his interviews, further topics that Bill Clinton laughed frequently about included criticisms leveled against him and aspects of his life that were embarrassing to him, as *Example 17.1* below shows. Our interpretation of all these instances of laughter has been that it was being used "as a deliberate, sophisticated, and rhetorical device" (O'Connell & Kowal, 2005a, p. 275). This usage of laughter appears to us to be a far cry from laughter that is "relatively uncensored, thus showing our true feelings" (Provine, 2004, p. 216). As regards the interviewers, some of them in both corpora manifested their own perspectives by joining in the laughter – contrary to Clayman and Heritage's (2002) requirement of neutralism on the part of media interviewers.

In some respects, however, Hillary and Bill Clinton's laughter was strikingly different from one another. Her laughter generally occurred while her interviewers were speaking and was a back-channeling of the HA-HA type; in other words, she typically laughed about something that was being said by an interviewer, not by herself. His laughter generally occurred as an overlay upon his own speech; in other words, he typically laughed about what he himself was saying. Hillary Clinton's laughter therefore remained generally semantically

opaque, whereas Bill Clinton's was generally semantically quite transparent. A similar gender-related finding has been reported by Dreher (1982; cited in Kotthoff, 1996, p. 127) for conversations between students: The male students' laughter was generally self-referential, whereas the female students' laughter was generally concerned with other people. Dreher has also found that female students laughed twice as often as did male students. A similar finding, also quite in accord with gender research, has been reported for Hannah Arendt in her interview with Günter Gaus, in which she laughed more than six times as much as did Gaus (O'Connell & Kowal, 1998).

Jefferson (1984, p. 346) has investigated the laughter of people who talk laughingly about their troubles. She found that the listener in such situations "does not laugh, but produces a recognizably serious response." In our study of Bill Clinton's laughter described above (O'Connell & Kowal, 2005a, p. 286), we too found this to be the case. In the following example, in which the interviewer Charlie Rose (CR) introduced a serious and delicate topic, Bill Clinton (BC) responded laughingly – with HA-HA laughter, indicated as HU, and overlaid laughter, indicated by underlining:

Example 17.1
CR well there was also this you were gettin' beat up so bad at home that you were anxious to get to the office
BC that's right I said that uh yeah that's ri-hi-hi-hight HU HU HU HU I probably was more attentive to my work for several mo-honths just because I didn't want to have to attend to anything else

Charlie Rose did not join in the laughter, but immediately continued with a new question. In general, O'Connell and Kowal found that Bill Clinton laughed mainly about his personal problems, whereas his interviewers hardly ever responded to these woes with laughter.

But we wished to further pursue the frontiers of adult laughter (O'Connell & Kowal, 2006a). An extreme case presented itself in *The third man,* a classical *film noir* (Korda, Selznik, & Reed, 1949). Stereotypically, one might expect more weeping than laughter in such a cynical, malevolent setting. In addition, there was only very little laughter noted in any of the printed background materials to the film: the novel itself (Greene, 1950), the screenplay (Greene, 1984), a www.geocities.com transcript, a German translation of the novel (Greene, 1962), and a partial German transcript (Timmermann & Baker, 2002); and such laughter as was noted was of an innocuous type – embarrassed, ironic, humorous, or pleasant laughter.

Much to our surprise, however, negative laughter (cynical, hypocritical, and mendacious) was both clearly evident and limited almost entirely to the criminal characters in both the English- and German-language versions of the film:

> This laughter constitutes a notable change from both the medial and conceptual literacy of the novel and other written versions to the medial and conceptual orality of the film itself as a portrayal of spontaneous spoken dialogue. (O'Connell & Kowal, 2006a, p. 305)

Thus, we confirmed, in an unlikely source, the perspectival nature of laughter and its use as a deliberate and skillful rhetorical device. Such laughter is far removed from the "instinctive, contagious, stereotyped, unconsciously controlled" HA-HA type laughter described by Provine (2004, p. 215) and from the laughter overlaid on speech that he has described as a "curious hybrid" (p. 216) thereof, as well as from Chafe's (2003a, 2003b, 2007) nonserious laughter.

Finally, in a phonetic study of laughter, Pompino-Marschall, Kowal, and O'Connell (2007) have analyzed the well-nigh hysterically dramatic laughter of Jane Austen's character Mrs. Bennet as interpreted by actress Alison Steadman in the BBC TV mini-series *Pride and prejudice* (Birtwistle & Langton, 1995). They found therein a great deal of abnormal phonetic features in combination, including expanded pitch range, laryngeal whistle, octave jumps, and ingressive phonation, along with reduced supralaryngeal articulation of mainly central vowels with no controlled consonantal articulation. Accordingly, they have characterized laughter phonetically as vocalization with high pulmonary and laryngeal tension combined with low supralaryngeal tension.

Conversation-analytic Research

Conversation-analytic researchers have investigated laughter not only in private settings, but also in the context of public settings, e.g., as audience responses (see also our Chapter 18). In an early study, Schenkein (1972) had emphasized the use of laughter on the part of the speaker as an adjunct to instruct listeners regarding an appropriate action. Jefferson (1979, p. 80; see also Glenn, 1991/1992) has introduced an important distinction between two types of laughter. In invited laughter, the "*speaker himself indicates that laughter is appropriate, by himself laughing*"; the invitation may be accepted or declined. In voluntary laughter, the initiative to laugh comes from the listener. In a subsequent study, Jefferson, Sacks, and Schegloff (1984, p. 1) have emphasized the function of reciprocal laughter in the establishment of conversational intimacy when "'improper' talk" has occurred, i.e., talk that violates conventional standards of "courtesy, propriety, tastes, tact, ethics, commonality, etcetera, etcetera." Adelswärd (1989, p. 129) has undertaken to question Jefferson's assumptions that laughter is generally elicited by something funny and that laughter is typically contagious. In a corpus of institutional interaction (e.g., job interviews and post-trial interviews with defendants), he has found that laughter is not necessarily a response to something that is presented as funny and is not necessarily social insofar as "we often laugh alone." In accord with the gender differences mentioned above, Adelswärd also found that women laugh more often than men.

Chafe (2007)

The year 2007 has been a very stimulating one for the study of laughter. In February, Wallace Chafe published his book *The importance of not being earnest: The feeling behind laughter and humor*. In August, an Interdisciplinary Workshop on The Phonetics of Laughter was held in Saarbrücken, Germany, with a keynote address by Wallace Chafe. A noteworthy aspect of Chafe's approach is that he has indeed deftly combined interdisciplinarily the psychological ("*feeling*") and the linguistic (phonetics) research on laughter.

We are particularly indebted to Chafe (2007) for his section on the phonetics of laughter, entitled "Part One **How we laugh**" (pp. 17–58). It constitutes the most complete survey of the acoustic and phonetic characteristics of laughter that we have found to date. He has noted that what happens in the mouth is far less important for laughter than for speech, and that the vowels of laughter "are murkier than those of speech" (p. 21) because "the vocal folds do not close as decisively during the voicing cycle as they do while people are talking." The stated thesis of Chafe's book is also physical:

> The thesis of this book is that laughter hinders the person who is laughing from performing serious physical or mental activity. Laughter is physically disruptive. Not only does it interfere with breathing, it destroys the rigidity of the torso that is necessary for various physical acts. (p. 23)

Insofar as this thesis is not explained at length, one is tempted to wonder what demanding physical feats Chafe had in mind. Under ordinary circumstances, the laugher is hardly cognizant of physical impairment of any kind. Hence, it is surprising that Chafe has made it so explicitly his major thesis. And it is noteworthy that he has not further developed an argument for this thesis throughout the remainder of his book.

But his book is actually about something quite different: "This, then, is above all a book about a feeling or emotion" (p. 1). It is "not primarily a book about laughter" (p. 2). The decision to make a feeling the real topic of the book entails a shift into a genuinely interdisciplinary method and indeed into a psychological orientation. Chafe's basic tenet regarding laughter is the following:

> Both laughter and humor are derived from ... the feeling of nonseriousness, ... a mental state in which people are inhibited from taking some event or situation seriously, which is to say that they exclude it from their accumulating knowledge of how the world really is. (p. 137)

But what is this "feeling of nonseriousness"? It is certainly not a *per se notum* concept, nor a scientifically established one. The two major components thereof are listed as "disablement and pleasure" (p. 58). Elsewhere, "happiness" (p. 54) and "euphoria" (p. 53) are additionally listed as components. Chafe's alternative of "cheerfulness or seriousness" (p. 58) suggests the association of sadness, unhappiness, and dysphoria with seriousness.

Chafe's intuitions do not appear to be related to any systematic empirical analysis of his data, but to be the conviction with which he began his book. The contrary evidence reviewed above (e.g., Adelswärd, 1989; O'Connell & Kowal, 2004a, 2005a, 2006a; see also *Merriam-Webster's collegiate dictionary*,11th ed., 2003, p. 560) is not part of Chafe's coverage of the issue. Perhaps he would also be inclined to argue that cases that run counter to his theory of the universal nonseriousness of laughter are simulations of laughter (see Chafe, 2007, p. 55) rather than genuine laughter. And furthermore, "Fake laughs, like fake smiles, are easy to recognize" (p. 66). And so, for Chafe, laughter "expresses a feeling that insulates us from serious thought and action" (p. 145). But the research of Bloch (1996; cited in Tracy, 2002, p. 88 f.), in addition to the research mentioned above, seems to negate such insulation. Bloch has dealt with very real everyday narrative situations of her Danish subjects, and her findings have indicated that laughter overlaid on speech may carry either the positive meaning of joy or the negative meaning of shame.

Chafe's insulation of his theory against counter arguments does not appear to allow an adequate scientific understanding of laughter as such. The laughter of adult human beings has been subjected to a learning process of multiple decades. It appears to be far more subtle, versatile, pragmatic, rhetorical, and flexible in its scope than Chafe's concept of *nonseriousness* would allow.

Perhaps, George Steiner's speculative interpretation of Arthur Koestler's (1964; cited in Steiner, 1977, p. 146) concept of laughter in the latter's *The act of creation* might be a provocative view of laughter to leave with the reader. How precious little we know of this phenomenon of our everyday life:

> Koestler argues that laughter is creative. It results from the collision, deliberate or accidental, of two previously unrelated, apparently contradictory areas of reality, language or attitude. From that collision flashes a brief glimpse of the world newly ordered. Laughter is a "sudden glory" of intelligence; as in a child's game, man regains the faculty of spontaneous action, he says to reality "be otherwise."

Whatever is described here, it is surely not the "instinctive, contagious, stereotyped, unconsciously controlled" HA-HA type laughter proposed by Provine (2004, p. 215). Nor does it bear much likeness to the laughter overlaid on speech that he has described as a "curious hybrid" (p. 216) thereof, or to Chafe's (2003a, 2003b; 2007) nonserious laughter.

To anticipate some comparisons from Chapter 18, Clayman (1992, p. 35) has included among disaffiliative audience responses the "derisive laughter . . . used by audience members to disassociate themselves from speakers and their views." For an audience member to do so is tantamount to responding both adversarially and very seriously to a speaker, certainly not nonseriously and certainly not necessarily – or plausibly – with the feelings of pleasure and happiness.

Finally, a recent **op-ed** piece in the Sunday *New York Times* by Frank Rich (2007, September 30, p. WK 12) has occasioned outbursts of both late night TV humor and serious questioning of Hillary Clinton's manner of

self-presentation. Rich has emphasized both the automaton character and the implicit insincerity of Hillary Clinton's raucous, explosive laughter:

> Then there was that laugh. . . .
> Now Mrs. Clinton is erupting in a laugh with all the spontaneity of an alarm-clock buzzer. Mocking this tic last week, "The Daily Show" imagined a robotic voice inside the candidate's head saying "humorous remark detected – prepare for laughter display."

Nonetheless, Immanuel Kant was certainly correct, as we have indicated in our epigraph at the beginning of this chapter: Laughter is both a positive and psychologically healthy part of our everyday life, although it can also be used in very hurtful ways. Provine (2000/2001, p. 77) was right: "Above all, studying laughter provides new insights into what it means to be and act human."

Chapter 18
Applause and Other Audience Reactions

Even when the audience doesn't seem to do anything, it is in fact doing something: The audience is always co-author (Duranti, 1991, p . 137).

It's the silence that counts, not the applause. Anybody can have applause. But the silence before and during the performance – that's the best of all (Vladimir Horowitz; cited in Hanssen, 2007, July 23, p. 29; our translation).

Chapter Prospectus

Chapter 18, *Applause and Other Audience Reactions*, is concerned with the reactions of groups of listeners or audiences. In this context, applause is considered to be a nonverbal response to spoken discourse that is typical of larger and more public groupings. Depending upon cultural and situational differences, it may occur sporadically throughout a spoken presentation, only at designated times (e.g., at the end of an act), or only terminally, at the end of a presentation. It can be used positively as well as sarcastically or inappropriately. Other audience reactions may be verbal or nonverbal, approving (e.g., *bravo* or *encore*) or nonapproving (e.g., *boo*). Research on audience reactions has not been extensive and for the most part has only engaged political oratory. In this setting, the limitations of spontaneity and genuine interactivity between speaker and audience make political oratory a special case of turn-taking or dialogical group activity. Virtually all the other phenomena we have discussed thus far in Part III may also enter into audience reactions. Finally, applause and other audience reactions are a case in which the often neglected role of the listener becomes quite overtly prominent; this prominence allows the researcher in turn to assess the effectiveness of political oratory. The speaker's intentions can no longer be easily hidden in the face of an audience.

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_18,

Applause and Laughter: A Comparison

So far, most of the discussions in this book have had to do with one-on-one settings of spontaneous spoken discourse. We now turn our attention to settings that occur typically between a speaker (or performer) and large groups of people, usually in public. These groups are referred to as audiences or spectators. Audiences are present in quite diverse settings, including theaters, political meetings, concerts, and sports events. Since we are dealing here with the coordinated reactions of many people, the types of behavior that lend themselves to such simultaneous expression are both limited and in some settings highly ritualized. Audience reactions include both verbal (e.g., acclamation or booing) and nonverbal (e.g., applause, laughter, or whistling) behaviors, some of which are positive and affiliative, some of which are negative and disaffiliative, and some of which can be used in either way. Our emphasis in the following is on applause, since it is the most common audience reaction to a public performance of speaking (Atkinson, 1984a) and has been empirically investigated more than the other behaviors. Nevertheless, in order to enable a more specific treatment of audience reactions, we wish to begin our discussion of applause with a comparison of applause to laughter as another frequent audience response. The reader might note that the materials considered in this chapter are quite foreign to mainstream psycholinguistics.

In Chapter 17, we have distinguished two basic types of laughter, the HA-HA type and laughter overlaid on speech. The former is nonverbal and semantically opaque, the latter verbal and semantically explicit. Applause is similar to the HA-HA type of laughter insofar as both are nonverbal and semantically opaque, so that both are all the more apt to serve in a variety of roles; applause accompanied by verbal utterances such as *encore* and laughter overlaid on speech are similar to one another insofar as both cease to be semantically opaque. Both applause and laughter are the products of muscular movements and are acoustically perceptible; both are generally affiliative, but can also be used disaffiliatively (but see Heritage & Greatbatch, 1986, p. 111, who maintain that "applauding is invariably a display of *affiliation*"); both are perspectival with respect to the audience members who either applaud or laugh; both are typically reactive rather than initiating; and both may be live or canned.

But there are also a number of differences: Applause is manual, whereas laughter is vocal. As a consequence thereof, clapping, as an articulatory activity involving hands and arms, seems to have a longer latency for initiation than laughter (Atkinson, 1984a, p. 23), but can be sustained over longer periods of time because it is not constrained by the necessity to breathe (Atkinson, 1984b, p. 372); this characteristic difference in timing has turned out to be important for the empirical analysis of audience reactions. Applause is limited mostly to audiences, whereas HA-HA type laughter occurs also in one-on-one settings. Applause does not usually lend itself to the convenient measurement of units such as syllable-like calls and bouts as does the rhythmic phenomenon of

laughter. Whereas a clapping audience may in addition rise to its feet for a standing ovation, standing without clapping is not at all common for a laughing audience. Whereas a public performance typically ends in applause, laughter alone is rarely a closing reaction of an audience. Finally, it should be noted that applause and laughter may indeed co-occur.

Some History of Applause

According to Jenniches (1969, p. 571ff.), the first evidence of applause dates back to the Greek theater of the fifth century B.C. Applause both presupposed and at the same time accomplished the division of roles between actors and spectators. More differentiated forms of applause were developed only later by the Romans. Audiences applauded actors, orators, poets, and politicians by clapping in various ways in keeping with the situation: with the hands held flat, or by cupping the hands, while extending either their forearms alone or their entire arm and while sitting or standing. To indicate great enthusiasm, upper class audiences added visual expression by waving the edge of their toga, whereas lower class audiences, who were not allowed to wear a toga, waved a piece of cloth, presumably the precursor of the handkerchief. One may note in passing that Repp (1986), in an exploratory acoustic analysis of clapping with various hand configurations, has found that experimental subjects were able to perceptually distinguish claps produced with different hand configurations.

In Christian churches, applause had thrived at first but gradually fell into disrepute as a heathen practice. Interestingly enough, after the Second Vatican Council, the practice of clapping in church has re-emerged, whereas laughter, in particular the "Easter laugh" (Keller & Vornbäumen, 2007, April 8, p. S 1) has not.

Jenniches has also emphasized that applause has had a colorful history in the theater: In the sixteenth and seventeenth centuries, spontaneous audience reactions (e.g., applause, shouting, laughter, and whistling) were common in Central Europe. In the eighteenth century, more sophisticated audiences showed less emotion and more discipline. Applause had now turned into a stereotypical affiliative response. The first Bayreuth Festival of 1876 brought with it Richard Wagner's absolute prohibition of applause during operatic performances. This practice gradually led to the eventual conviction within German operatic circles that such applause was indeed in bad taste. In turn, opposition to this practice further encouraged the development of claques: Claqueurs were paid for producing well-planned applause, a practice continued even today in the recorded radio and TV production of canned applause, i.e., applause considered appropriate by the program director.

And even today, the ritualization of applause remains perhaps its most astounding characteristic. In the course of time, spontaneity has been replaced by discipline, audience noisiness by silence. And the silence of up to an hour that

sometimes precedes applause is the product of a long development in civilization chronicled by Jenniches and eloquently noted by Sennet (1977, p. 206):

> Restraint of emotion in the theater became a way for middle-class audiences to mark the line between themselves and the working class. A "respectable" audience by the 1850's was an audience that could control its feelings through silence; the old spontaneity was called "primitive."
>
> . . .
>
> In the 1750's, when an actor turned to the audience to make a point, a sentence or even a word could bring immediate boos or applause... By 1870, the audience was policing itself. Talking now seemed bad taste and rude. The house lights were dimmed too, to reinforce the silence and focus attention on the stage. . . .

Empirical Research on Applause to Political Oratory

As noted before, there is little empirical research available on the interaction between individual speakers and audiences. In fact, observable, immediate audience reactions such as applause have not even been considered in recent books in which the audience itself is presented as the main topic. Thus, Livingstone and Lunt's (1994) *Talk on television: Audience participation and public debate* does not consider applause or other audience reactions. Similarly, McQuail (1997), in a book entitled *Audience analysis*, does not mention applause or other audience reactions in his index, even while examining closely audience properties such as selectivity, utilitarianism, intentionality, resistance to influence, and involvement. The *non-present* or virtual audience is an entirely different phenomenon. Many TV programs play to no one other than a potential audience of viewers. How this audience influences and reacts has hardly been mentioned in the research to be reviewed here (but see, e.g., Baldauf, 2002).

With few exceptions (e.g., West, 1984), empirical research on audience reactions has been initiated and continued up until the present in the conversation-analytic tradition, and it has involved almost exclusively political oratory. Interestingly enough, almost all the studies available to the present authors were published during a period of only 10 years, starting with Atkinson (1983) and ending with Clayman (1993). A more recent conversation-analytic study by McIlvenny (1996) has concerned itself with heckling, i.e., with *individual* verbal reactions from members of a larger audience rather than with collective reactions. The results of Kurzon's (1996, p. 199) analysis of "semantic and paralinguistic strategies for eliciting applause" are difficult to interpret because of methodological problems (e.g., his quite idiosyncratic assessment of speech rates). More recently, Kuo (2001) has applied a conversation-analytic approach to the study of applause in Taiwanese political oratory, and Papaioannou (2003) has used a psycholinguistic methodology to analyze applause in eight U.S. presidential inaugurals and eight state-of-the-union speeches by the same presidents.

Atkinson (1983, 1984a, 1984b, 1985, 1986), the "founding father" of conversation-analytic applause research, has explicitly embedded his studies in a rhetorical tradition: Ancient and modern rhetoric had and has as its main purpose to identify effective means for designing a given speech so that it grips and holds the attention of listeners in a setting where there is, over extended periods of time, nothing else for the audience to do but listen and watch. The problem, however, is that rhetoricians have dealt almost exclusively with texts, i.e., with the written, rather than with the verbal and nonverbal rhetorical performance of a speaker. This limitation has made it difficult to assess the effect of rhetorical devices on an audience. In other words, the rhetoricians have treated speeches as monologues (Brodine, 1986, p. 172). According to Atkinson, applause and other immediate audience reactions, be they affiliative or disaffiliative, are behaviors which make possible an empirical engagement of political rhetoric. His basic premise has been that applause does not occur spontaneously or as a behavior initiated by the audience but as a behavior elicited by various aspects of the speaker's behavior or by rhetorical devices used in his or her speech: The audience is *invited* to applaud by the speaker.

In his study of affiliative audience reactions, Atkinson (1984b, p. 374) has observed that applause often tends to be preceded by vocal audience reactions such as whistles and that "bursts of applause tend to last for seven, eight, or nine seconds (i.e., eight plus or minus one second)." According to Atkinson (1984a, p. 28), applause in the range of 7–9 s should to be regarded as "normal." With respect to the temporal organization of speaker–audience interaction, he has noted that applause typically starts "*just before* or *immediately after* (i.e., with no gap) a possible completion point" in the speech of an orator (Atkinson, 1984b, p. 377). His database showed that in no instance did a speaker attempt to continue his or her turn after applause had started. He considered this disciplined, one-at-a-time sequential organization to be quite comparable to the norms for turn-taking postulated for conversation by Sacks et al. (1974) (see our Chapter 16). However, Atkinson (1985, p. 179) has observed that politicians sometimes refuse to stop speaking even after they themselves have invited applause, i.e., they continue to speak after the onset of applause. Though a rare event, this refusal seems characteristic of charismatic orators (e.g., J. F. Kennedy).

Atkinson (1984a) has identified three verbal rhetorical devices which have proven to be effective in eliciting applause: namings, lists of three, and contrastive pairs. He has assumed that these characteristics of the spoken text help the audience to anticipate or project completion points in the speeches at which collective action is appropriate. Atkinson (1986) has shown that parts of speeches which contained such rhetorical devices and were followed by applause had a better chance to be quoted in radio or TV programs and subsequently to become famous slogans. Quantitative analyses of 476 speeches at the 1981 British Conservative, Labour, and Liberal party conferences have confirmed Atkinson's claim that "political messages that are packaged in rhetorical formats embodying emphasis and projectability are more likely to

be applauded than messages that are not so packaged" (Heritage & Greatbatch, 1986, p. 140). Heritage and Greatbatch have also reported that contrastive pairs elicited by far the most applause (nearly 25% of all applause events). In addition, Atkinson (1986) has found that changes in gestures as well as an increase in speech rate were effective means of inviting applause (see also Atkinson, 1985, p. 167 ff., for a summary of techniques for inviting applause). Failure to use such techniques may have accounted for Michael Foot's poor performance against Margaret Thatcher in the British General Election campaign of 1983 (Atkinson, 1986, p. 52). Grady and Potter (1985, p. 182), in a similar comparison of Foot and Thatcher, have found that Foot did indeed elicit significantly less applause than did Thatcher, precisely because he used prosodic and timing cues that made it difficult for the audience to project places appropriate for applause. Grady and Potter have ascribed this problem partly to the fact that Foot, an accomplished speaker who previously "did not rely on written texts while speaking" finally yielded to pressure from journalists and adhered woodenly to his written text during this campaign.

Bull's (1986) analyses of the speeches from the same British General Election of 1983 have been carried out from a psychological point of view. He has distinguished two categories of applause: *Sustained* applause involves clapping of a large part of an audience, whereas *isolated* applause involves clapping by only one or two people. It should be noted that the latter is not in accord with Atkinson's (1984a, p. 21) criterion that clapping becomes applause only "when several do so repeatedly and at the same time." Bull's (1986, p. 111f.) in-depth analysis of a speech delivered by Arthur Scargill – "by far the most successful in evoking applause from the audience" – has confirmed Atkinson's finding that applause is elicited by one or more of his list of rhetorical devices. Bull has also found that Scargill talked through isolated applause and always resumed speaking before sustained applause ended. Most of the time, the audience began to applaud before Scargill had finished a sentence, and he in turn continued speaking inaudibly through the applause, i.e., he showed his refusal of invited applause – a ploy we have mentioned earlier that characterizes charismatic speakers (Atkinson, 1985). In view of Sacks et al.'s (1974) requirement of smooth turn-taking as a norm, it seems remarkable that both Atkinson and Bull have noted charismatic and successful speakers, respectively, who consistently violate this norm in their interaction with their audiences. In addition, Bull (1986, p. 115; see also Brodine, 1986, p. 190) has noted that Scargill's use of the various rhetorical devices was always accompanied by hand gestures calculated to either encourage or inhibit applause: "Arthur Scargill actually seems to conduct his audience." Heritage and Greatbatch (1986, p. 120) have observed that across party lines "speakers were applauded much more often for attacking statements. . .than for constructive policy assertions." They have related this negativism on the part of the audience to an observation by Simmel (1902; cited in Heritage & Greatbatch, 1986, p. 121) regarding mass actions: "The motives in individuals are often so different that their unification is possible in the degree in which their content is merely negative and destructive."

Atkinson (1984a, p. 85) has mentioned preliminary evidence suggesting that applause-elicitation sequences work similarly in France, Germany, and the Netherlands and has claimed that these findings may have "cross-cultural applicability far beyond the English-speaking world." However, the research of Kuo (2001, p. 212) on applause and laughter in the 1998 Taipei mayoral debates did not confirm the effectiveness of Atkinson's three rhetorical devices (namings, three-part lists, and contrastive pairs) for the elicitation of applause insofar as they occurred only rarely in Kuo's data.

With respect to duration of applause, Kuo has found that it lasted an average of 4.2 s rather than Atkinson's 8 s norm. She has concluded that "results of further research from non-western languages" are needed to test the generalizability of Atkinson's claims. Kurzon (1996, p. 223), who – confirmatory of Atkinson – has observed an overall mean applause duration of 8.72 s for public performances of politicians in the U.S., has acknowledged, however, that the wide variability of applause duration for individual speakers "is probably a good enough reason for considering such figures as somewhat suspect." Papaioannou (2003) has found that the duration of applause for people named by the speaker was significantly longer than applause about issues, in both U.S. presidential inaugurals (12.8 > 7.8 s) and state-of-the-union addresses (23.8 > 10.6 s). Only the applause about issues in the inaugurals was anywhere close to Atkinson's (1984a) norm of 8 s. Her data have also shown that the audience applauded for longer periods of time and more enthusiastically (e.g., with cheering sounds, whistling, stomping, affiliative laughter, and vocal encouragement such as *yeah, yeah*) in the state-of-the-union addresses than in the inaugurals. In other words, the duration of applause as well as the additional reactions of the audience seem to be idiosyncratic to the setting, speaker, and type of speech. As an example of a remarkable deviation from Atkinson's 8 s norm, Papaioannou has cited the address to the joint houses of congress that was delivered by President George W. Bush immediately after the September 11, 2001 attacks: Applause averaged 22 s in duration. At the very least, it would appear that Atkinson's 8 s norm of applause duration can be considered normative only in a very weak fashion; the deviations from this duration seem to be circumstantial or situational.

Without reference to the conversation-analytic tradition, West (1984) has engaged exploratory research on audience reactions from a political science perspective. His data have included applause, laughter, cheers, and boos to speeches of U.S. politicians during the 1980 presidential election. Although it is not clear how he derived his counts of various audience reactions, they have shown that, across 10 politicians, by far the most frequent audience reaction to their nomination speeches was applause (62.3 %), followed by laughter (35.7 %), cheers (1.6 %), and boos (0.4 %); these percentages have been derived from West's Table 1 (p. 39). However, there was wide variability for individual politicians with respect to relative amount of applause, ranging from a high of 95% to a low of 44%, mainly due to the correspondingly high variability of the laughter component (from 55.6% to 5%).

Empirical Research on Laughter as an Audience Reaction

As a collective audience reaction, laughter has been studied empirically less often than applause, although, as West's research has indicated, it may be a rather frequent reaction to political rhetoric. Chafe (2007, p. 118) has called attention to the fact that "audience laughter has a unique sound. . .quite different from the sound of a single person laughing." Clayman (1992), in his study on disaffiliative audience reactions during three 1988 U.S. presidential-election debates, has discussed both affiliative and disaffiliative audience laughter. Affiliative laughter occurred when a candidate criticized his opponent but managed to do so humorously and therefore made the audience laugh *with* the candidate at his opponent. Such laughter was also often followed by applause. By contrast, disaffiliative laughter occurred when a speaker talked uncritically about himself and did so in a nonhumorous tone, but was not convincing in defending himself against critique or accusations. Disaffiliative laughter was seldom followed by applause.

Montgomery (2000, p. 122) has studied audience laughter in a British TV chat show in which talk is treated "as an occasion for laughter." The planned invitation to laugh consisted of an accelerated tempo on the part of the speaker. Neither the host nor the guest attempted to speak during audience laughter. In confirmation of Clayman (1992), Montgomery noted that longer laughter was often followed by applause.

Empirical Research on Booing as an Audience Reaction

According to Clayman (1993, p. 110), booing as a form of collective behavior may be considered "the quintessential display of disapproval in the public speaking context." However, this would appear to be something of an exaggeration insofar as departure can be clearly the most expressive display of disapproval. It is important to note that *boo*, unlike applause, is both disaffiliative and verbal; it is, in fact, a conventional or tame primary interjection (see our Chapter 14). Clayman has analyzed a total of 33 speech excerpts in which booing occurred. The excerpts were taken from a large variety of public speaking settings, including U.S. presidential debates, British party conference speeches, and TV talk shows. He has found that booing occurred most often when a speaker made an unfavorable remark about a political adversary. This is in stark contrast to Brodine's (1986, p. 174) finding in a database of U.S. national party conventions in 1984: "Instances of booing in our data were invariably *in agreement* with the speaker (e.g., in response to an insult directed at the other party)." The apparent incompatibility of the results may be resolved by a closer look at the corpora of the two studies. A part of the corpus studied by Clayman was chosen because its settings were adversarial to the speakers, whereas the setting studied by Brodine was partisan to the speakers.

Such a comparison of datasets indicates one direction for future research on audience reactions: It should more carefully distinguish various speech settings in view of their potential to facilitate or inhibit specific audience reactions (see also Clayman, 1992, p. 55). A step in this direction is to be found in Papaioannou's (2003) research, mentioned above, on applause in U.S. presidential inaugurals vs. state-of-the-union addresses. The two sets constitute very different genres of speech: The inaugurals are shorter (mean duration ca. 20 min) and are delivered in the open to a general audience, whereas the state-of-the-union addresses are three times as long (mean duration ca. 1 h) and are delivered indoors to a joint session of the U.S. congress. In the former, applause is sporadic and spontaneous; in the latter, applause is to some extent ritualized according to political party, with Democrats or Republicans rising to their feet as separate groups in accord with the rhetorical thrust and content of the speech. These differences in genre account quite plausibly for the results discussed above.

To return to Clayman's (1993, p. 116) data, the onset of booing had the following characteristics: (1) a "substantial time lag" of at least 0.5 s between a speaker's disapproved utterance and onset of booing in more than two thirds of his data; (2) other affiliative or disaffiliative audience reactions (e.g., clapping or shouting) between the disapproved utterance and onset of booing; or (3) other affiliative or disaffiliative audience reactions while the speaker was still talking before onset of booing. More specifically, Clayman noted that many disaffiliative audience reactions that preceded booing tended to be individual rather than collective. In addition, he found that booing seldom lasted for more than 3 s and that speakers showed a general tendency to ignore booing, i.e., for the most part they continued to talk through booing.

An anecdote from the *New York Times* (2007, June 2, p. WK 7) throws additional light on the complexity of disaffiliative booing. As a recent Miss Universe finalist, Rachel Smith was repeatedly greeted "with one chorus after another of boos." But one must be very careful about identifying who or what was being booed. There was nothing personal about the booing at all; "it's part of a dialogue with the neighbor up north": "So, Mexicans say, the booing at this pageant was never about Miss U.S.A. herself. It was those letters on her sash." The booing in this instance was entirely symbolic, and the recipient of the booing was a surrogate for the U.S.A.

Audience-to-audience Responses

On closer inspection, a number of observations in conversation-analytic research reveal that applause or other audience responses may be elicited not only by the speaker herself or himself, but also by reciprocal activity among audience members. Whether onset of the audience response is sudden or gradual would appear to be crucial in this instance. The assumption is that a sudden burst of applause indicates that the speaker's behavior or the content

of the speech itself has elicited the applause, whereas a gradual onset indicates the influence of the audience itself.

Heritage and Greatbatch (1986, p. 111, footnote 6) have observed that an audience seldom takes up applause initiated by only one or two individuals and that these individuals themselves typically cease to applaud after a couple of seconds. By contrast, applause that begins as a burst of collective audience reaction attains a maximum intensity within a second and is considered by Heritage and Greatbatch to be the norm for applause initiation. But they do not explicitly consider the possibility that the gradual emergence of applause may indicate the potential of an audience-to-audience interaction. It seems that such a project would demand a more sophisticated methodology than has been used for applause research thus far. For one thing, the temporal organization of speaker–audience interaction according to Heritage and Greatbatch involves time units that, in keeping with psycholinguistic research (e.g., Spinos et al., 2002), cannot be reliably perceived. For example, Heritage and Greatbatch (1986) have reported that in their data the "physical initiation of clapping from rest takes approximately 0.2 seconds" (p. 116) or "0.3 seconds" (p. 112); but they do not indicate how they determined these reported durations with the exactitude they have noted. Montgomery (2000, p. 125, footnote 3), who has also worked within the conversation-analytic framework, has acknowledged that "it is difficult to be precise in the timing of applause or laughter."

Without pursuing the matter, Atkinson (1984a, p. 21) has pointed out indirectly that applause initiation may be a result of audience-to-audience interaction:

> When we look more closely at how applause gets under way, it emerges that a main function of other affiliative responses is to prompt audiences to start clapping. Even when no official cheer-leaders have been appointed, individuals who whistle, laugh, cheer or shout 'Hear hear', effectively perform the same task of leading the rest of the audience into a collective response.

Such audience-to-audience influence is clearly planned in the case of claqueurs, who, as putatively legitimate audience members, are hired in order to initiate reactions from the "real" audience. Their behavior would lend itself to a more thorough investigation of applause onset as a product of speaker–audience as well as audience-to-audience interaction.

Clayman (1993) has explicitly taken up the issue of speaker vs. audience influence on the elicitation of coordinated audience reactions by contrasting the typical initiation of booing with the typical initiation of applause in his corpus. He has argued that there may be two different processes involved which facilitate coordinated audience reactions: "*independent decision-making*" (p. 111) and "*mutual monitoring*" (p. 112). Independent decision-making implies that individual audience members are guided in their reactions by characteristics "within an ongoing speech, places that stand out conspicuously from the talk thus far" (p. 112) and by their assumption that other audience members will find similar places appropriate for a response. By contrast, mutual monitoring

implies that members of the audience, in their decision to react, "may also be guided, at least in part, by reference to the behavior of other audience members" (p. 112). According to Clayman, audience responses based on independent decision-making should begin suddenly and with a large part of the audience responding simultaneously; audience responses based on mutual monitoring should begin gradually with only a few audience members involved who subsequently lead others to respond. Clayman has emphasized that the two onset processes are not mutually exclusive. He has reported that, according to extant empirical research, applause onset is typically guided by independent decision-making, whereas onset of booing seems to be guided by mutual monitoring. Or to put it more succinctly: "Clappers usually act promptly and independently, while booers tend to wait until other audience behaviors are underway" (p. 124). He has explained these findings with reference to the conversation-analytic concept of "preference organization" developed for the analysis of conversations. Preference organization implies that affiliations in general "tend to be produced promptly, in an unqualified manner, and are treated as requiring no special explanation or account" (p. 125), whereas disaffiliations tend to be "delayed, qualified, and accountable" (p. 125).

As indicated above, for the empirical analyses of onset of audience reactions, the temporal organization is of crucial importance in conversation-analytic research. Durations in the order of milliseconds are assumed to determine the standard for affiliative or disaffiliative audience responses. From the point of view of psychology, conversation-analytic research on audience responses should indeed be pursued, but with measurement refinements characteristic of psychological methodology.

Whereas onset of audience responses has turned out to be of both theoretical and empirical interest, the same has not as yet been shown for offset of audience responses. In this respect, we are left with a warning from Alexander Solzhenitsyn's *The Gulag Archipelago* (1973, p. 69 f.; cited by Atkinson, 1984a, p. 29f.), a warning that applies only to situations where inordinate power is being exercised over public behavior: "Don't ever be the first to stop applauding."

The Listener's Golden Opportunities

Nowhere in the entire arena of dialogical activities does the listener come into his or her own more prominently than in the use of applause and other audience responses. The vast resources of interjections, somaticized expressions of affiliation and disaffiliation, laughter, marching, silence – everything in the armamentarium of dialogue – are at their disposal. Baldauf (2002, p. 25) has given the example of the theater audience that may make use of vocal forms and structures that have played no role to date in analyses of dialogue; these devices include sighing, crying out, and whistling. One has only to imagine the hysterical welcome visited upon Frank Sinatra and Elvis Presley in their day; or

Pope John Paul II; or the funerals of Princess Diana and Mother Theresa; or the teenagers of Germany vying with the teenagers of Italy in acclaiming Pope Benedict XVI with BE-SECH-ZEHN instead of BE-NE-DE-TO. Of course, BE-NE-DE-TO won. A similar example from the concert world is given by Hanssen (2007, July 23, p. 29; our translation): The "active formation of public opinion" in an audience may take place via competitive *bravo* and *boo* calls that lead occasionally even to brawling. As researchers, we have precious little understanding of these dynamics. In fact, our inclination is still to look for an explanation in the speaker rather than in the creativity of the listener and/or in the dynamic of the setting. The term "crowd control" is evidence of the need for discipline of group reactions in these settings, precisely because the sky's the limit and socio-cultural habits are not always adequate to restrain enthusiasm, even in Mecca, and more particularly among soccer hooligans. Sentential syntax, the favorite structure of the linguist and psycholinguist alike, disappears amidst all the other candidates for prominence in the determination of such behaviors. The setting and society impose the controls, not a researcher in a laboratory.

Some Comparisons between Individual Listeners and Audiences

Throughout our discussions in this chapter, one notion in particular has implicitly threaded its way: The listener, through the metamorphosis into audience finally comes into his or her own. The behavior of an audience as listener becomes *overt*. Some of the specific differences that are built into this dramatic shift include the collective potential of an audience for a variety of reactions, the limitations of such audience reactions to partly ritualized, expected ones, the public nature of the setting, and the change in speaker-audience relationship from a one-on-one to a one-on-many relationship. The analogue of turn-taking in the one-on-one setting becomes in the public setting a complicated set of speaker-audience, audience-speaker, and audience-to-audience shifts; the audience never becomes the speaker as such, but has multiple other means at its disposal for interaction. Obviously, the spontaneity of spoken discourse is limited in the public setting, sometimes by scripting of the speaker's text, sometimes by the various relationships listed above; but it is never the informal, free-wheeling back-and-forth of an informal conversation.

Chapter 19
Intersubjectivity

Dialogue is both logos and concrete, embodied intersubjectivity (Luckmann, 1990, p. 53).

Chapter Prospectus

Chapter 19, *Intersubjectivity*, engages the most basic psychological and social prerequisite for spontaneous spoken discourse. Intersubjectivity does *not* include consciousness of the content of discourse, but focuses on the mutual and reciprocal consciousness of one another on the part of interlocutors. That is to say that the interlocutors are aware not only of the interactive presence of other interlocutors, but also of their own awareness of them. Failing this prerequisite, spontaneous spoken discourse ceases – or fails to commence. Intersubjectivity itself is not verbal, but provides the conscious underpinning for both speaking and listening.

The *Principle of Intersubjectivity* as a Problem in Literary Stylistics

At the very outset, one must carefully distinguish the psychological notion of intersubjectivity as intended in this book from "the principle of intersubjectivity" presented in his discussion of stylistics by Taylor (1997, p. 98). Because of his high profile among language scientists, the danger exists that the principle of intersubjectivity critically discussed by Taylor be assumed without question to define our discussion of intersubjectivity. We must therefore clarify the matter before engaging a very different concept of intersubjectivity.

Taylor (1997) has been intent upon refuting a notion of intersubjectivity that he assumes to underlie a number of prominent stylistic models in the domain of literary scholarship, all based upon Charles Bally's (1951) original model. The essence of Taylor's critique pertains to two aspects of stylistics: (1) the distinction between the stylistic content of verbal materials and their expressive aspects, i.e., their linguistic meaning, on the one hand, and (2) the central question in stylistics à la Bally as to how the expressive aspects can be related

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_19, © Springer Science + Business Media, LLC 2008

systematically and objectively to the content aspects. Whereas expressive aspects of a given literary text are open to inspection, content aspects are not; they are accessible only to the analytic scholar him- or herself. According to Taylor (p. 97), "there is no criterion by which the sameness or difference of the stylistic contents of two expressions may be determined." Thence results a dilemma that constitutes the basic problem of such models of stylistics: In order to identify relevant expressive characteristics, one needs criteria to determine when a specific stylistic content is being communicated. The way out of the dilemma – a solution criticized and rejected by Taylor – *is* the principle of intersubjectivity which claims that "we all *do* experience the same pattern of samenesses and differences of content" (p. 98) and that therefore "the communication of stylistic content works". By contrast, Taylor has argued that the principle of intersubjectivity may not even be true and is, in any event, superfluous.

A Psychological Concept of *Intersubjectivity*

A psychologically realistic concept of *intersubjectivity* has been proposed by Rommetveit (1974, p. 29) more than a quarter of a century ago in terms of an "*intersubjectively established, temporarily shared social world*." The concept is both epistemological and phenomenological. It is not the product of empirical analyses of data, but the very starting place for empirical research on communication. It defines the psychological reality which has to be realized in the face-to-face, one-on-one dialogue of spontaneous spoken discourse. Every dialogue must begin with a mutual acknowledgement of presence, not just physical presence, but psychological, conscious presence *to one another*. If this consciousness lags, dialogue is diminished; if it ceases, dialogue ceases instantly. And in addition to the mutuality, whereby each is aware of the other, there is also a reciprocality, whereby each knows that simultaneous to consciousness of the other there is also a consciousness that the other is conscious of oneself.

Then only one may engage spontaneous spoken discourse, when this complex setting is established. And yet, this process, for all its complexity, is almost instantaneous and is comfortable to the universal family of man. Intersubjectivity is a social, partial, temporary, interactive consciousness that sometimes reaches an initiating crescendo of "hi." If one falls asleep, becomes distracted, is diverted to activity that interferes with this awareness, the dialogue languishes, sputters, ceases. If it is maintained, it allows understanding and, above all, appreciation of one's fellow man. Sadly enough, intersubjectivity also makes possible an abundance of griping, hurting, lying, and disdaining of one another – both verbally and physically.

Mainstream psycholinguistics has almost totally ignored this fundamental prerequisite to all dialogue. One can set up the most perfect experimental conditions to elicit a dialogue; but if intersubjectivity is not evinced, there is

no dialogue. The use of one's native language in the spoken mode usually provides for the enactment of this intersubjectivity; but little children in multilingual settings sometimes do without their native language and still inaugurate interactive games of following, leading, chasing, being caught, hiding, and above all laughing and smiling, touching and beckoning – without benefit of words. The only prerequisite seems to be the maintenance of intersubjectivity; and in turn, the maintenance of an affective relationship – whether one of benevolence or even of malevolence – further fosters intersubjectivity. The same is true in the adult world, even though it is frequently hidden by the prevalence of the verbal in our interrelationships.

It is of the essence of one's consciousness that no other person can commandeer all of it or commandeer it permanently: Dialogical intersubjectivity is, in other words, partial and temporary. The commonality of dialogue is never complete, nor does it last forever. Indeed, were it complete, there would be no need for words; were it nil, there would be no possibility of dialogue.

That is to say that the imperfection inherent in human dialogue reflects the very limitations of our human nature and consequently of our social interaction with one another. The refusal to establish this prerequisite of dialogue by otherwise directing one's attention, and the thoughtlessness, meanness, or injustice involved in allowing a dialogue to terminate – or insisting that it terminate – also involve the violation of verbal integrity as discussed in Chapter 22. It is in this sense, that dialogue can be seen to recapitulate the moral responsibilities of our social lives. For it is the mandatory venue for all social interaction. Other less immediate or less complete contacts (e.g., phone, internet, snail-mail) can be delegated to fulfill these social responsibilities, but they remain derivative, partial solutions.

The fact that intersubjectivity always involves only a partial complementarity does not detract from its importance at all. In actuality, this partial complementarity provides the fundamental rationale for engaging dialogue with one another: We need to be complemented by one another in every aspect of our social lives. The mechanism to engage this enrichment of our private existence is precisely dialogue with one another. It is in this most basic sense that no man is an island.

Hence, intersubjectivity is psychological bedrock, the *conditio sine qua non* of all dialogue and the rationale which makes dialogue attractive and useful for humans. We have commented earlier on the relationship of this intersubjectivity to what Hörmann (1976) has referred to as the "transparency" of language use:

> The intersubjective characteristics of discourse are for the most part transparent to the interlocutors, in the sense intended by Hörmann (1976, p. 58, our trans.; see also Hörmann, 1981): "One sees what is meant, as it were, right through the words and sentences." Language use does indeed bridge from consciousness to consciousness transparently. Both speaker and listener are ordinarily not concerned with the specific tools used (see O'Connell, 1988, p. 67). It is precisely this transparency that is typically disrupted in mainstream psycholinguistic experimentation. (O'Connell & Kowal, 2003, p. 203 f.)

In spontaneous spoken discourse that is one-on-many or many-on-many, intersubjectivity must be allocated or distributed differently than in one-on-one dialogues. A group of listeners may vary from one individual to another in their intersubjective relationship to an ongoing dialogue. Similarly, a speaker's intersubjective relationship to an audience is quite different from his or her relationship to an individual interlocutor. And the eavesdropper, overhearer, or bystander may not be part of the consciousness of a speaker-listener pair at all. The first author recalls, for example, a dinner with a graduate student at the Georgetown University Faculty Club. Toward the end of the evening, a woman at an adjacent table rose, approached our table, knelt beside it, and proceeded to thank us very graciously for providing a most engaging conversation, in which she had obviously – unilaterally and completely unnoticed – shared with great enjoyment. Her covert, unilateral participation in our conversation as an eavesdropper is very far removed in terms of mutual and reciprocal consciousness from public media interviews, in which the interviewer and interviewee must mandatorily relate intersubjectively to one another *and* to a remote audience over an extended period of time, if anything intelligible is to be produced for an audience.

Finally, intersubjectivity applies most especially to the situation in which two – or perhaps more – people share a long silence together. It occurs more often than one might imagine while friends or lovers go walking or sit before a fire or simply gaze at each other for a period of time. In this setting, a dialogue involving verbal interaction remains both potential and unneeded; the interaction is thoroughly nonverbal, intimate, and deeply affective. It is not at all unrealistic to assert that intersubjectivity reaches to the very depths of our human inclination to communicate with and relate to one another.

Chapter 20
Perspectivity

Perspective is one of the components of reality. It is not at all to be assumed that perspective distorts reality; quite the contrary, it orders reality (Curtius, 1949/1993, p. 224; our translation).

Chapter Prospectus

Chapter 20, *Perspectivity*, asserts our conviction that every utterance necessarily carries with it a perspective of the speaker. One speaks from one's own point of view and/or one deliberately or indeliberately assumes that of interlocutors or others. Speakers have reasons, intentions, goals; listeners similarly have their own purposes in listening. Language use always has some finality; it does not just occur, but occurs for a reason. The goal of discourse is *not* simply the pursuit of consensus, as some theorists currently hold; such a conviction has its origins in an oversimplified informational analysis of communication. Nor is perspectivity on the part of a speaker detrimental to intersubjectivity, but constitutes instead a firm foundation for the enrichment thereof.

The Basic Concept of *Perspectivity*

Perspectivity is very much akin to the intersubjectivity, which was our topic in the previous chapter. Both are psychologically so very close to us as speakers and listeners that we find it difficult to back away far enough to get them into focus. Metaphorically, we speak and listen from a place, and it goes without saying that the place is defined by one's bodily presence, one's situatedness in the here and now. Socially considered, that place is with someone (or a number of others). Psychologically considered, it is from some point of view, some vantage point, some *Weltanschauung*, some set of priorities that vary from a set of current preoccupations to a lifelong philosophy of life – *my perspective*. As Eelen (2001, p. 223; cited in Mills, 2003, p. 35) has put it:

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_20, © Springer Science+Business Media, LLC 2008

> Each individual also has a unique individual history and experiences the "common" world from this unique position. The common world is thus never identical for everyone. It is essentially fragmented, distributed over a constellation of unique positions and unique perspectives.

Linell (1998, p. 48) has pinpointed perspectivity in similar terms:

> Each thought or utterance views aspects of the world from some particular vantage-point, thus telling us (as recipients or analysts) something not only about the things talked about but also about the actor's background.

These perspectival orientations in turn are rooted in the fundamental psychology of the individual, a fact not saliently emphasized either in modern psychology in general or in mainstream psycholinguistics in particular. But William James (1891/1981, p. 380 f.) has pinpointed it over a century ago in a classic passage:

> *My experience is what I attend to*. Only those items which I *notice* shape my mind – without selective interest, experience is an utter chaos. Interest alone gives accent and emphasis, light and shade, background and foreground – intelligible perspective, in a word.

This is just as true for listeners as it is for speakers, except that the perspectivity of the listener is generally less overt insofar as it is expressed less verbally.

The note of self-interest in James also pinpoints the fact that perspectivity involves affect. Hence, perspectivity is closely related to the use of several of the phenomena discussed in Part III: first-personal pronominals as self-referents, prosody as one of the "phenomena associated with affect" (Norrick, 2001, p. 86), interjections as "the expression of spontaneous reactive emotions or evaluations" (Nübling, 2005, p. 604; our translation), and laughter as "a declaration of individuality, possibly the most individualized of the human uses of sound" (Edmondson, 1987, p. 29). In the sequence in which they appear here, these phenomena signal a decreasing dependence on linguistic means and at the same time an increasing, though implicit expression of perspectivity. It is in this sense that the audience reactions of booing and applause might well be added to laughter as expressions of perspectivity.

Suleiman (2000, p. 30) has discussed a number of concepts related to perspectivity: *self* (e.g., Goffman, 1959, p. 252 f.); *face* (e.g., Goffman, 1967, p. 5); *footing* (Goffman, 1981, p. 128; Levinson, 1983, p. 283); *frame* (e.g., Schiffrin, 1994, p. 104); and *positioning* (e.g., van Langenhove & Harré, 1993, p. 84). All of these concepts constitute legitimate approaches to the orientation of speakers and listeners. However, it is our conviction that the concept of *perspectivity* is the most comprehensive and inclusive of them all, as well as the most basic and constitutive one in the existential order of *personal psychology*. Hence, our own considerations are all couched in perspectival terminology.

Mine and Thine

But, there is a preliminary clarification needed: The basic distinction between *having* a perspective and *taking, setting*, or *assuming* an interlocutor's perspective. Both perspective having and perspective taking are essential to the overall concept of perspectivity.

How often have we heard comments or questions such as: "I don't know where he's coming from" or "What's she have in mind?" The assumption is always – correctly – that people begin to speak with something in mind. They wish to communicate something to someone. They begin by having a perspective that is in turn meant to be communicated to others. Developmentally, existentially, and epistemologically, having a perspective is prior to taking a perspective. It is simply unavoidable: If we begin to speak, the action must spring from some motivation that is rooted in our personal ego as origin. In other words, the intention to start, stop, or continue speaking (or listening) must be informed by a perspective that is deeply rooted in the individual speaker (or listener). By contrast, setting or taking a perspective always presupposes a relationship to another person or persons, which is to be fostered, developed, and/or altered. Or, as Rommetveit (1990, p. 97) has expressed it: "*Perspective setting* in human conversation is essential for the transformation of human subjectivity into temporary states of *intersubjectivity*, i.e., of convergence of attention onto relevant aspects of the talked about state of affairs."

A glance at the research literature, however, shows that having a perspective has been subordinated to taking a perspective. Thus, Holtgraves (2002, p. 121) has recently described the process of perspective taking as follows:

> In order to construct an utterance that will be understood by a recipient, the speaker must try to adopt the hearer's perspective, to see the world (roughly) the way the hearer sees it, and to formulate the remark with that perspective in mind.

He has then proceeded to list a number of theorists who "have argued that it [perspective taking] is *the* fundamental task in language use (Brown, 1995; Clark, 1985; Krauss & Fussell, 1991a, 1991b; Mead, 1934; Rommetveit, 1974)." And in the same context, he has made Clark's (1996) insistence that it is coordination that is the essence of successful communication (Holtgraves, 2002, p. 122) virtually synonymous with perspective taking.

But Holtgraves' (p. 121) contention that "the speaker must try to adopt the hearer's perspective" is something of an exaggeration (see Alber, O'Connell, & Kowal, 2002, p. 258). *Merriam Webster's collegiate dictionary* (11th ed., 2003, p. 17) includes the following relevant definitions of "adopt": "to take up and practice or use . . . to accept formally and put into effect." These definitions make it clear that the concept of *adoption* is not quite in accord with Holtgraves' usage. In order to communicate with a listener, a speaker does not need to adopt the listener's perspective, but rather needs to *take it into account*. This then is the common meaning of taking, setting, or assuming the listener's perspective. In fact, the genuinely successful adoption of a listener's perspective can often spell

the death knell for any sort of creative or informative dialogue. If the finality of a dialogue becomes purely the successful adoption of the listener's perspective, what function does the speaker retain as contributor to the ongoing interaction? And as for coordination as the essence of successful communication, it is rather the prerequisite for successful communication (see Rommetveit, 1974, p. 52). Interlocutors may be marvelously coordinated as a speaker/listener unit and fail miserably to communicate. Again, the clarification of the concept can be of help. *Merriam Webster's collegiate dictionary* (11th ed., 2003, p. 275) tells us that to coordinate is "to bring into a common action, movement, or condition." Such coordination does not engage the communication of content per se. As Rommetveit (1974, p. 52) has insisted, it appears rather to be a prerequisite to successful communication, and to have more in common with intersubjectivity than with perspectivity.

Holtgraves (2002, p. 127) has further contended that "intersubjectivity is a state that is achieved to varying degrees through language." This too appears to be at variance both with the facts and with Rommetveit's (1974, p. 39) comments about the expansion, modification, and increase of intersubjectivity through language. Language itself is always already assumed by Rommetveit to be "nested on to only partially shared social realities." In other words, intersubjectivity precedes spontaneous spoken discourse as prerequisite; it may indeed be expanded, modified, and increased by ongoing discourse – but not created, not initiated thereby.

Dialogicality

Rommetveit (1990, p. 98) has contended that perspective is always under dyadic control. The fact is that the expression of either having or taking perspective generally emerges in a dialogical setting. By and large, perspective is reciprocal from the beginning of a dialogue; but it is further negotiated as dialogue proceeds. Accordingly, perspective taking has an essentially interactional character (Graumann, 1989, p. 99). For example, Suleiman et al. (2002, p. 285) have found that the *HARDtalk* perspective of British TV interviewer Tim Sebastian "is defined by the momentary relationships that emerge from the intersubjective setting itself." With an interviewee who shows her- or himself to be "the stereotypical politician," Sebastian becomes ruthless and penetrating. But the lion turns into a lamb when the interviewee presents himself or herself as a publicly recognized good guy, e.g., a war crimes investigator or a UNESCO humanitarian official. In other words, the expression of perspective taken by Sebastian develops interactionally in the course of the interview:

> Without a controversial topic, it is difficult to muster passion; without a challengeable interviewee, there can be no bullying; without a television audience interested in an issue and biased with regard to it, there can be no successful *HARDtalk*.

A further question then arises: To what extent can the expression of one's personal perspective be offset in the sense that it is subordinated by taking the perspective of someone else or of an institution (e.g., a TV network)? Clayman and Heritage (2002, p. 120) have claimed that the news interviewer sacrifices the expression of his or her own perspective through the cultivation of a professional *neutralism*. To the contrary, we would insist that such a vacuum of perspectivity is impossible. By way of illustration, we offer the following example. This chapter was being composed during the week of the horrendously tragic killings at Virginia Technological University, Blacksburg, Virginia on Monday, April 16, 2007. In such settings, the radio or TV announcer is limited by professional decorum to avoid as much as possible any expression of personal perspective. Still, the massacre is reported with a grim seriousness, whereas a children's festival is reported with a pleasant demeanor. However, both extremes are expressed through nonverbal means only; the words themselves are authored by the network itself. In other words, the journalistic institution – the network – is acknowledged as the ultimate arbiter as to what is verbally appropriate in such instances. More specifically, as we have already commented a number of times in this book, the news announcers play the role of Goffman's (1981, p. 226 f.) "*animator*, that is, the sounding box from which utterances come"; however, they assume personal responsibility for and therefore authorship of their nonverbal expressions of compassion or pleasure. In a certain sense, these instances of expression of one's own perspective can be looked upon as a sort of leakage that defies neutralism. By contrast, news interviewers not only express themselves nonverbally, but also make use of their own words, which are eo ipso more closely related to personal perspective.

The crucial point introduced by Clayman and Heritage (2002, p. 119 f.) with respect to neutralism on the part of news interviewers is clearly professionalism. This requirement leads in turn to the need for "*the interviewer's management of a 'neutralistic' stance towards the interviewee's statements, positions, and opinions*." However, the empirical evidence would seem to indicate that such a rigid neutralism is simply not adhered to by news interviewers. As O'Connell and Kowal (2006b, p. 155) have found, "Tim Sebastian's BBC *HARDtalk* program is an excellent example of the deliberate abdication of neutralism" (see also Suleiman et al., 2002), as are also Christiane Amanpour's CNN interviews with Yasser Arafat (see Alber et al., 2002). Cohen (1987, p. 19) too has noted that the interviewer "often has well defined opinions that he or she expresses in the course of the interview." Both answering instead of asking questions and laughing are also expressions of the personal perspective of the interviewer that violate a rigid neutralism (see O'Connell & Kowal, 2004a, 2005a). Hence, Clayman and Heritage's (2002) position with regard to neutralism would appear to be inadequate to challenge the conviction expressed by Graumann and Kallmeyer (2002, p. 1) that every utterance necessarily emanates from some perspective, i.e., from "a position from which a person or a group view something (things, persons or events) and communicate their views." Or, as Foppa (2002, p. 17) would have it: "Whenever two people are

engaged in a dialogue they are displaying their respective perspectives on the issue in question" (see also Ensik & Sauer, 2003). Foppa (2003, p. 50; our translation) has also recognized perspectivity as basic to Graumann's theory of dialogue and of human interaction itself:

> Whenever one expresses one's own perspective on a topic, it is necessarily presented also as a potential perspective for others. And even if my interlocutor doesn't accept my perspective, he or she must have recognized it as a potential perspective, if only to be able to reject it.

Perspectivity is indeed ubiquitous in spontaneous spoken discourse, and that presence in every utterance definitively excludes any kind of neutralism that is not so watered down as to be meaningless. But the degree to which perspective is openly expressed verbally and/or nonverbally may vary considerably. As regards research on perspectivity, it seems to be a matter of choosing a sufficiently sensitive methodology to investigate empirically such subtle variations in perspective.

Perspectivity and Intersubjectivity

Finally, the objection can be made that having a perspective interferes with a thoroughgoing intersubjectivity insofar as the two states of mind are somehow incompatible: The one is personal, the other is interpersonal. But such a conclusion is not warranted. One must keep in mind that intersubjectivity is always temporary and partial; the interlocutors are genuine interlocutors only insofar as they are seeking to expand their intersubjectivity, and this necessarily involves perspective taking. However, the most important safeguard to that finality is actually *having one's own perspective*. The honest interlocutor is forthcoming about the motivations and assumptions that inform his or her discourse. And this perspectival stance is precisely conducive to an expanded intersubjectivity. The politician who minces words and is on everyone's side at the same time does not expand intersubjectivity, but rather encourages a definitive cynicism regarding political rhetoric. In other words, it is conducive to greater intersubjectivity to know more completely where one's interlocutor is really coming from, what his or her personal perspective is regarding the topics under discussion. The on-going dialogue itself provides the further opportunity to *take, set, or assume new perspectives* presented in the course of the dialogical interaction – once again, a process that takes place in real time.

We began this chapter with a quotation from Ernst Robert Curtius in which he has claimed that having a perspective, far from distorting reality, orders and organizes it. Linell (2005, p. 209) has described a very different claim on the part of some social constructionists: That having a perspective creates its own truth. This perspectival creation of truth has been spelled out by Linell as follows:

> According to this position, there are countless, context-dependent perspectives on and versions of the world, and none can be said to be better or more true than another.

Linell has further contended that Shotter (1993), Gergen (1994), and Potter (1996) "sometimes come fairly close to this position." Wilfred M. McClay (2007, February 2, p. W 13), in his recent piece on the "Twilight of sociology" in *The Wall Street Journal*, has bluntly criticized the implications of proclaiming that no version of the world "can be said to be better or more true than another." It would appear that for him such sociological relativity is a logical and historical precursor of political correctness:

> And if, as many sociologists came to believe, all reality was "socially constructed," then nothing was grounded in nature, nothing was justified by tradition or custom, and nothing was to be treated as enduring. All things were provisional, and all could be reshaped, usually along predictable political lines. Thus academic journals and scholarly monographs were given over to supporting the reigning views of race, gender and class – and fiercely suppressing any inquiry that might challenge these views.

We wish to stand with Linell and McClay in this matter: Such a tendency on the part of radical constructionists can obliterate "the distinction between fact and fiction" (Linell, 2005, p. 209) and, indeed, seems to take a specific perspective for the definitive truth. It is our conviction that such a closed perspective regarding truth impedes genuine intersubjectivity.

Chapter 21
Open-endedness

Any human utterance of a certain weight contains more than the author may have been immediately aware of at the time (Ratzinger, Pope Benedict XVI, 2007, p. xix).

The discourse should be treated as an achievement; that involves treating the discourse as something 'produced' over time, incrementally accomplished, rather than born naturally whole out of the speaker's forehead, the delivery of a cognitive plan (Schegloff, 1982, p. 75; cited in Linell, 1990, p. 152).

Chapter Prospectus

Chapter 21, *Open-endedness*, emphasizes that spontaneous spoken discourse exists in an open, creative, learning, and didactic setting. The openness involved is not only the absence of ritualistic or written-text determination, but a readiness to listen to an interlocutor who co-determines the forward movement of dialogue in time. Accordingly, neither is open-endedness some kind of narcissistic readiness to await whatever may occur to *me*. It always moves forward in real time to a new focus point, and the direction and pace of this movement are essentially unpredictable. Open-endedness is closely related to intersubjectivity; and just as with intersubjectivity, the absence of open-endedness terminates genuine spontaneous spoken discourse. Reading aloud allows no open-endedness as to words, although expressive readers often introduce minor verbal changes. Readers are also free to exercise open-endedness as regards prosody and other nonverbal expressions. But a thorough-going open-endedness pertains primarily only to spontaneous spoken discourse, and consequently to both speaker and listener.

The Concept of *Open-endedness*

A danger in choosing the word *open-endedness* to express a property essential to spontaneous spoken discourse is that it be taken to be synonymous with emptiness, a passivity characterized as the tabula rasa, "the mind in its hypothetical primary

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_21, © Springer Science + Business Media, LLC 2008

blank or empty state before receiving outside impressions" (*Merriam-Webster's collegiate dictionary*, 11th ed., 2003, p. 1271). The definitions given for *open-ended*, however, are quite different: "not rigorously fixed: as **a** : adaptable to the developing needs of a situation **b** : permitting or designed to permit spontaneous and unguided responses" (p. 869). Open-endedness is, then, a readiness, an expectation of the future. It was St. Augustine of Hippo (ca. 397–400/1960) who, in Book 11 of his *Confessions*, wrote eloquently of the subjectivity of time, and especially of the expectation of the future. Such expectation can be either positive (hope) or negative (fear). And the open-endedness of dialogue with another person incorporates these affections of the soul of man. Hence, the open-endedness of spontaneous spoken discourse is both active and receptive, but not passive and empty.

Spontaneity

We come thus indirectly to the concept of *spontaneity* itself and to *Merriam-Webster's collegiate dictionary's* (11th ed., 2003, p. 1206) definitions of *spontaneous*:

> **[**LL *spontaneus*, fr. L *sponte* of one's free will, voluntarily**]** (1653) **1 :** proceeding from natural feeling or native tendency without external constraint **2 :** arising from a momentary impulse **3 :** controlled and directed internally : SELF-ACTING <∼ movement characteristic of living things>... **5 :** developing or occurring without apparent external influence, force, cause, or treatment **6 :** not apparently contrived or manipulated **:** NATURAL...

The tenor of this dictionary entry is clearly in the direction of a relationship between spontaneity and open-endedness. At every moment in all genuinely spontaneous spoken discourse, the verbal and nonverbal contributions of both the speaker and the listener in the next moment are not specifiable. They are not determinate and therefore cannot be designated in advance. Indeed, there are probabilities that stochastic statistical procedures thrive on and that sometimes yield a high predictability or, in conversation-analytic terms, projectability. But predictability is not the same as determinateness. In open-endedness, the next element in the sequence is *always indeterminate*. Researchers can find many such predictive elements – *post factum* – precisely because they are dealing with the completed product of spontaneous spoken discourse, *not* with the ongoing, moment-to-moment development of the dialogical interaction from the perspective of the interlocutors involved in the dialogue. At the level of local management, the same problem is reflected in the researcher's effort to determine whether an interruption has occurred; the difference between completion and incompletion is not always momentarily discernible. In general, the researcher is dealing with a corpus that has already been produced; the participants by contrast are dealing with the moment of real time. Open-endedness deals with each successive moment as a creative opportunity, the change of what is future into what is instantly past.

The reason for the indeterminacy is not really all that difficult to locate: Interlocutors are preeminently voluntary with respect to what they will decide to say next. They are free, in the richest sense of the word: They say what they intend to say. The freedom of open-endedness is not vacuous; it proceeds in an orderly direction, but not in a predetermined fashion. This is not to deny that there are indeed social pressures on interlocutors to be less than free and there are ritualistic components of spontaneous spoken discourse that constrain freedom. The melodramatic scene of the vicious criminal with a gun to the hero's head may not allow a great deal of freedom other than pleading when the bad guy inquires, "What did you call me?" But the good guy was, of course, able to speak freely at the time he referred to the scoundrel in the first place – accurately – as "a loathsome scoundrel." However, no one in his or her right mind would refer to the gun-to-head scenario as spontaneous spoken discourse. Spontaneity is clearly a prerequisite for the exercise of open-endedness.

The situation is quite different in the case of reproductive spoken discourse – reading aloud or reciting from memory without the written text at hand. Here spontaneity is severely limited insofar as the verbal text is a given. The only spontaneity that can be engaged must be by means of prosodic and other nonverbal expressiveness. However, as we have already found in our discussion of interjections, and especially the use thereof by actors (see our Chapter 14), a limited spontaneity in reproductive spoken discourse is practiced by additions and minor changes of the printed text: In this case, the inclination to spontaneous expression supersedes the given words. Such spontaneous changes are quite different from error proneness, which is certainly to be acknowledged as a component of indeterminacy and therefore of open-endedness, but not of spontaneity.

A Methodological Problem

As we have emphasized before, empirical researchers of spontaneous spoken discourse are typically confronted with a transcript and audio and/or video recording of a completed corpus of spoken discourse. These data do not give researchers immediate or direct access to the moment-to-moment psychological provenance of the sequential elements of such discourse. In other words, they have no immediate or direct access to the momentary spontaneous determination of each successive contribution to the corpus of speech: They have no access to the micro genesis (*Aktualgenese* in the German) of the dialogical interaction. The temptation in that case – and obviously the procedure engaged in many instances – is to analyze the corpus precisely as a completed behavioral entity. But almost all the psychological nuances of spontaneous contribution are thereby lost to analysis. And so, the moment-to-moment organization of the ongoing spoken discourse is lost. A similar methodological problem has plagued psycholinguists for decades in their attempt to decipher the cognitive processes

involved in sentence or language production. Thus, Aitchison (1983, p. 230) in her chapter on how "we plan and produce speech" has stated quite emphatically: "It is tantalizingly difficult to observe how anyone actually plans and produces speech... Clues as to what is happening are infuriatingly elusive." This impasse may well have contributed to the tendency of mainstream psycholinguistics to avoid "the collection and analysis of uncontrolled or genuine data" (Dietrich, 2002, p. 14; our translation).

And now, it would appear incumbent upon the present authors to provide a solution to this methodological impasse. But we have none: Our experience of psychological methodologies and our training therein have taught us to be modest in the face of the intrinsic limitations of hermeneutics. It is our conviction that these limitations constitute an impasse in empirical methodology (see also Taylor's critique of the *Principle of Intersubjectivity* in our Chapter 19). Occasionally, by courtesy of the interlocutors themselves, we are given an insight into a specific spontaneous moment, but by and large, such richness remains elusive to empirical methodology. The human spirit in its moment-to-moment decision processes remains largely inaccessible to empirical science (but see the American Association for the Advancement of Science, 2007); we still do not read minds.

This inaccessibility of the moment-to-moment is also related to a conviction that we have mentioned previously (see our Chapter 10): Conversation-analytic researchers have insisted that their transcripts must come "as close as possible to the experience of those actually participating in the interaction" (Psathas & Anderson, 1990, p. 87). At issue in this passage is specifically the conviction on the part of Psathas and Anderson of the relative unimportance of "'absolute' or 'clock-time' differences." The moment-to-moment inaccessibility of the interlocutor's "experience" adds one more methodological problem to their use of transcripts, insofar as their assumption of accessibility exemplifies William James's (1891/1981, p. 195) psychologist's fallacy: "The *great* snare of the psychologist is the *confusion of his own standpoint with that of the mental fact* about which he is making his report."

A further aspect of the psychologist's fallacy characterizes mainstream psycholinguistics rather than the conversation-analytic tradition of research. It consists in treating the completed transcript as the important unit of analysis, even though the moment-to-moment dynamic is the preeminently important locus of psychological engagement and therefore the richest mine for psychological insights as to the intelligibility of the ongoing spontaneous spoken discourse. The absence of an extant methodology with which to engage the moment-to-moment does not suffice as a justification for considering the completed block as more important. What happens is analogous to standing in awe of the hugeness of the Great Pyramid of Cheops without speculating about the human, scientific, engineering, political, and historical endeavors that contributed to it over time. The fact that our speculations are not rewarded with clear-cut closure should not prevent us from indulging them. Once again, what transpires in time is on that very account ultimately mysterious. The mysteriousness of the classic philosophical

definition of time – *mensura motus secundum prius et posterius* (the measure of motion according to before and after) – is ample witness to our inability as researchers to cope adequately with the interlocutor's experience of time. The mysteriousness itself is part and parcel of open-endedness.

The Psychology of Open-endedness

Ultimately, open-endedness has to do with our expectations, our eagerness, our readiness as human beings to engage one another in dialogue. The human spirit seeks always to go beyond itself, to go out to others, to transcend the here and now, the present of time and space. Our social inclinations, then, are the origin of genuine spontaneous spoken discourse: the desire to share, to learn, to teach, to rejoice with one another. Small wonder that the poets have enshrined dialogue as the high road to human fulfillment. At the same time, however, the very open-endedness of dialogue is not always a source of joy, much less of fulfillment. Fear of the unknown can effectively block the positive aspects of open-endedness. This is sometimes the case in dealing with a new, unfamiliar interlocutor or with new topics or issues. And as we emphasized in Chapter 19, psychological intersubjectivity does not automatically entail the predictability of such unknowns in the domain of semantic content in ongoing dialogue. If there is mystery to be found in dialogue, open-endedness is its usual source.

Chapter 22
Verbal Integrity

Morality is such a common and intrinsic quality of everyday social interaction that it is usually invisible to us, like glasses that provide a sharp sight of the area beyond although they themselves remain unseen (Bergmann, 1998, p. 280).

Chapter Prospectus

Chapter 22, *Verbal Integrity,* concerns the basic respect speakers and listeners must have for one another and for others in order to carry on genuine spontaneous spoken discourse. It excludes mendacity, disdain, haughtiness, arrogance, disregard, and disinterest, and the concept embodies as well the fundamental ethical dictum *bonum faciendum, malum vitandum,* the universal mandate to do good and avoid evil. The relevance of verbal integrity for this book is precisely that a comprehensive and adequate analysis and understanding of spontaneous spoken discourse from a psychological point of view is impossible without factoring in verbal integrity: For interlocutors in spontaneous spoken discourse, meaning is always embedded within moral responsibility. Interlocutors are not only psychological agents; they are moral agents. Our inclusion of the concept of *verbal integrity* as a building block for a theory of spontaneous spoken discourse makes it clear that our own discourse in this book has become explicitly evaluative. We are concerned here with *genuine spontaneous spoken discourse* and with genuine interlocutors. We acknowledge the necessity of integrating such concepts into a theory that can guide the empirical analysis of very diverse types of spontaneous spoken discourse. The genuine norm, then, is expressed precisely in the very title of this book: *Communicating with one another* – and in the present context of verbal integrity, we would add – *effectively*. This is indeed a daunting task, and it constitutes a goal far removed from the ideal delivery of mainstream psycholinguistics.

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_22, © Springer Science + Business Media, LLC

The Concept of *Verbal Integrity*

The concept of *verbal integrity* refers to the respect and consideration due on the part of interlocutors both for one another and for those about whom they converse. It is something of a corollary to the perspective inevitable in all spontaneous spoken discourse: Dialogical interaction, insofar as it springs from speakers' and listeners' orientations and takes into account the orientation of others, has also an inevitable socio-cultural finality and hence is embedded in a moral dimension of life. It is our obligation as interactive human beings to speak and listen responsibly, and there is no way to hide from that role without damaging genuine interaction itself. Correlatively, it is the obligation of the researcher to analyze responsibly. We characterize spontaneous spoken discourse not only as reflecting finality and not only as perspectivized, but would insist that both these characteristics also possess a moral dimension of social responsibility on the part of interlocutors, i.e., verbal integrity. The formulation of this integrity as *verbal* has been chosen by O'Connell and Kowal (2006c, p. 207) precisely to limit the morality in question to the context of spontaneous spoken discourse. It can clearly be expressed also in any of the nonverbal components of spoken discourse, but it remains dependent upon the verbal as its context.

The application of the concept to spontaneous spoken discourse is not at all original with us; in fact, it has been taken into account in a number of language related disciplines. More than 50 years ago, Goffman (1955; cited in Linell, 1995, p. 183) had already emphasized: "Utterances are produced not only 'with respect to' the other but also 'in respect for' the other (and oneself)." The citation from another sociolinguist (Bergmann, 1998) that introduces this chapter reflects a similar preoccupation with morality. The communication scientist W. Barnett Pearce (2002, p. 17) has been quite explicit: "Conversations are dances in which we and others move among positions in the moral order." And the social psychologist Ragnar Rommetveit (see Wertsch, 2003), to whom we have already frequently referred in this book, has for many years been moving more and more explicitly in this direction. He has become convinced that the moral dimensions of discourse must be considered essential components of a researcher's analyses. But for him, it goes much further than this: The researcher's own morality must enter into the analyses he or she engages. As Hagtvet and Wold (2003, p. 199) have expressed it:

> Over the years Rommetveit has become increasingly concerned with moral issues. If the external world may be described in a multitude of ways, and if meaning and truth are contextually dependent, scientists cannot any longer justify their activities by arguing that they are giving objective descriptions of external states of affairs. Nor is one perspective more "right," "neutral," or "God given" than others. The scientist's personal history, concerns, and agendas are then brought into focus, and a crucial question arises: From what perspective, or (moral) position, is this research conducted? To Rommetveit, the answer is to be found in the realms of ethics and morals as much as in rationalistic thinking.

It is thus that "Values, whether explicitly acknowledged or tacitly endorsed, determine the positions and perspectives of the researchers." In other words, the immersion of spontaneous spoken discourse in a moral, normative, value-laden socio-cultural world is part and parcel of the research setting itself. The researcher – immersed inevitably in the same setting – dare not disregard this immersion at the risk of distorting his or her data. Quite literally, what an utterance means in a given context and socio-cultural setting is partially determined by that very immersion in the world of human values. For example, as we have indicated in Chapter 16, researchers have all too often regarded interruptions in dialogue as aggressive behaviors of male interlocutors addressing women. It has been a short step of inference, then, to consider all interruptions simply as expressions of unadulterated power. Such an oversimplification manifests a serious neglect on the part of researchers with respect to the complexity of spoken discourse in socio-cultural settings. And in turn, it manifests those researchers' serious neglect of verbal integrity on their own part. This may indeed reflect one path by which myths acquire the status of facts (see Cameron, 2007).

The Authors' Preoccupation with Verbal Integrity

The reader might well argue that the preoccupation of the present authors is only a reflection of their own personal ethical convictions. Indeed it does reflect our ethical convictions, but it also reflects our many years of trying in vain to salvage mainstream psycholinguistics from its own objectivism and latter-day behaviorism in the guise of cognitivism. But the severe isolation consequent on such a philosophy of science and such a methodological solipsism cannot ultimately be justified; and hence, a critique of mainstream psycholinguistics has become a fundamental rationale for our preoccupation with the socio-cultural world of values: The typical psycholinguists' laboratory settings of unrealistic, often written "language use" are actually an eloquent exhibition of their neglect of the embeddedness of genuine language use in the real world of human values. This neglect constitutes in turn a serious failing in ecological validity.

The entire preceding paragraph might well be construed as a violation of verbal integrity on our part. The challenge of honestly and forthrightly criticizing one's colleagues while at the same time avoiding arrogance and meanspiritedness is indeed daunting. It is a balancing act that we as authors have sought to achieve, but may not have consistently attained.

Verbal Integrity as an Essential Component of Spontaneous Spoken Discourse

The various historical instantiations of a psychology of language use have been so intent on mimicking the natural sciences – with physics as the prototype and ideal – that they have seriously neglected the socio-cultural nature of language

use. Verbal integrity, then, is not to be considered something tacked onto the phenomenon of spontaneous spoken discourse, but rather something that inheres in every utterance. Wertsch (2003, p. 184) refers this acknowledgement back to Rommetveit:

> In his most recent writings and conversations, Rommetveit has focused increasingly on issues having to do with the ethical stances inherent in communication – including those that surface in the study of communication itself.

Ethical stances are indeed inherent in human communication processes. O'Connell and Kowal (2006c, p. 207) chose to apply the phrase *verbal integrity* to this phenomenon as follows:

> *Quite simply, every utterance of a human being in socio-cultural interaction with others through the medium of language must be a responsible interaction – on the part of both speaker and listener, and on the part of both writer and reader. The concept of verbal integrity is intended to be broader than that of truth itself. For example, it incorporates the obligation to take another person seriously – with respect and attentive listening – in every encounter. Disdain, frivolity, over-earnestness, prolixity, evasiveness, arrogance, and aloofness all violate integrity, even though their relationship to truthfulness is more remote. . . . Rommetveit would insist that utterances can simply not be understood comprehensively and adequately unless verbal integrity is factored in as a relevant element.*

In other words, the speaker's, listener's, writer's, or reader's accordance with verbal integrity radically affects the semantic, syntactic, and pragmatic import of an utterance for all concerned. Mainstream psycholinguistics has not thus far generally allowed entry to such a concept, even though it is urgently needed to supplement a sterile methodology by making the actual meaning of utterances and written texts accessible to a psychology of language use.

The temptation – not least of all for the researcher – is to consider everyday instances of verbal integrity (and its violations) as trivial, perfunctory, and insignificant parts of spontaneous spoken discourse. But the back-channeled *mm*s and *mm-hm*s, the *thank you*s, the *mornin'*s, the '*scuse me*s, the *sure*s, the *oh*s, the *yeah*s, and even the four-letter expressions of our impatience and consternation that sneak out under our breath are all warp and woof of everyday discourse. We cannot do without them precisely because they are of such great importance to express our momentary *stance* to the social world around us.

Malinowski (1923, p. 314) has referred to such "sociabilities" as "one of the bedrock aspects of man's nature in society," and even more dramatically, Ogden and Richards (1923, p.8) have characterized them as a universal duty for civilized man. But it appears to us that, in research on language use, they are often considered to be trivial, implicit, elliptical, momentary, unnoticed throwaways. The same is to be said of a whole array of ritualized usages of language, including the whole domain of politeness and diplomacy (see Brown & Levinson, 1987; Grice, 1975). In a similar vein, Glinz (1959, p. 104) has claimed that a conversation between two or more persons "interposes between them values which are already held in common or have yet to be accepted." And Cameron (1995, p. 163) has gone so far as to claim that "there is no language [use] without

normativity." Shotter (1991, p. 62; cited in Linell & Rommetveit, 1998, p. 465) has summarized much of the literature on the subject as follows:

> One of the most significant features of human actions in everyday life is not only their situated, contexted or occasioned nature, but also their normative accountability. That is, in some sense they are not only made but judged in their making, by both those who perform them and those around them, as to both their social and moral appropriateness to the circumstances of their performance.

Duranti (2003, p. 131f.) has come at the moral scenario in political discourse with yet another instantiation. He has cited the situation in which an audience interprets approvingly what a speaker has intended otherwise. Here the speaker must react and respond either by shifting and incorporating the audience's interpretation into his own continuation of his or her speech (an example of perspective taking) or stick to the intended interpretation and endeavor to bring the audience around to his own perspective. In either case, he or she must act responsibly: "This struggle over the right balance between pleasing others and asserting oneself reminds us of the centrality of morality in the construction of human agency through talk."

Some Examples

O'Connell & Kowal (2006c, p. 214) have provided two examples, illustrative in different ways of the importance of verbal integrity. The first of these is the simple utterance "I'll get 'im." Among its other meanings, this very short verbalization can serve two diametrically opposed functions: as an heroic offer on the part of a firefighter to enter a burning building to save a comrade, and as a threat to seek vengeance on an enemy. The moral dignity on the one hand and the moral shame on the other would be sharply distinguished in a concrete utterance not by the words – which remain identical – but by prosodic differences. The heroic version might be loud and firmly resolute, quick and definitive; the shameful version might be more drawn out in duration, deliberate, harsh, and grim. The point to be made is that it is precisely the moral orientations themselves that are psychologically important to the speaker with respect to how these words are to be articulated and to the listener with respect to how they are to be understood.

The second example is a quite different one. It is our diagnosis of a recent theory of cognitive linguistics as inadequate precisely because of the neglect of socio-cultural values. Croft and Cruse (2004, p. 329), in their book on cognitive linguistics, have introduced on the very last page a needed theoretical element that goes far beyond the cognitive. There they have acknowledged the necessity "for cognitive linguistics to engage with the social-interactional nature of language." At that point, it would appear to be too late to admit that such "a contribution to a theory of language" must be precisely "a theory of language that goes beyond cognition." It is not a question of going "beyond cognition,"

but of integrating from the beginning of one's theorizing an orientation toward the socio-cultural embeddedness of spoken discourse within human values. By their acknowledgement, Croft and Cruse themselves have suggested wherein the future of cognitive linguistics lies. And their suggestion further "pinpoints the fact that socio-cultural, dialogical, perspectival functions of every utterance must be incorporated into any viable comprehensive and adequate theory of language use" (O'Connell & Kowal, 2006c, p. 214). Years ago, Ogden and Richards (1923, p. viii) cautioned researchers against compacting the various functions of language use into "the symbolic," insofar as the conflation has led to "many controversies in the sciences." Today's orientation, synonymous with "the symbolic," is cognitivism: Theories of cognitive linguistics have forced cognitive processes into accounting for functions far beyond the cognitive domain. Verbal integrity is an important component of the domain beyond the reach of a reductionistic cognitivism.

One of the important implications of reductionistic cognitivism, Clark and Clark's (1977, p. 261 f.; Clark, 1996) conceptualization of the "*ideal delivery*" must be replaced, to our way of thinking, with genuine dialogue as the ideal (see Buber, 1958; see also Gadamer, 1975; as cited in Linell & Rommetveit, 1998, p. 468; Habermas, 1998):

> Whereas ideal delivery (the speech of the ideal speaker) is constructed from the structure of the language itself (syntax and the lexicon), the ideal dialogue is determined by contributions from socio-cultural norms entirely other than linguistic structure. (O'Connell & Kowal, 2006c, p. 218)

It should be clear that linguistic structures are relevant in spoken discourse, but their priority over socio-cultural embeddedness is not warranted.

Ethics

As a final note, we wish to enter a disclaimer. In emphasizing the importance of verbal integrity in the analysis of spontaneous spoken discourse, we do not wish to engage the philosophical discipline of ethics. In this respect, we find Linell and Rommetveit's (1991, p. 469) terminology regarding "studies of morality within the framework of discourse" infelicitous. We would rather speak of studies of discourse within the framework of morality. Ours is a "study of communication itself," cognizant of "the ethic stances inherent in communication" (Wertsch, 2003, p. 184). This appears to us to be the proper role of a psychologist of language use; attempting to play the role of ethicist oneself is not.

Chapter 23
Spontaneous Spoken Discourse

My theoretical explorations led me all the way from the propositional calculus to poetry – with a profound disregard for traditional academic boundaries and barriers . . .

The difficulties we encounter when we engage in dialogues across traditional academic boundaries should . . . carry promises of creative transcendence (Rommetveit, 1974, p. 2 and p. 128).

Chapter Prospectus

Chapter 23, *Spontaneous Spoken Discourse*, is our effort to pull some of the most important psychological characteristics of language use together so as to vindicate our own theoretical perspective regarding spontaneous spoken discourse. Its orderliness is not determined uniquely by syntactic (sentential) well-formedness. It transcends such a lawfulness derived only from the written with a complex orderliness to be found in the very elements and components we have here depicted. Above all, intersubjectivity, perspectivity, open-endedness, and verbal integrity offer to the psychology of language use the purview necessary to engage human spoken interaction empirically. Our conclusion then must be that a comprehensive and adequate engagement is otherwise not possible.

Syntax

In the spirit expressed by Rommetveit in our epigraph for this chapter, we can formulate the questions: Is syntax a valid starting point for an analysis of the structure and functions of spontaneous spoken discourse? Will syntax lead us to an adequate theory of spontaneous spoken discourse? Perhaps, a step back to the basic concept of *syntaxis* in the Greek language can help us to answer these questions. The original meaning of syntaxis is "to arrange together" (*Merriam-Webster's collegiate dictionary*, 11th ed., 2003, p. 1269). One of its principal uses was to describe an array of troops drawn up for battle;

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_23, © Springer Science+Business Media, LLC 2008

the syntaxis was a spatial array, a formation or collocation of soldiers. It is in this sense that syntaxis has come in the course of history to refer to the spatial allocation of words in a written sentence.

Syntax, however, is defined by *Merriam-Webster's* without direct reference to the sentence: "**1 a :** the way in which linguistic elements (as words) are put together to form constituents (as phrases or clauses)." The American Psychological Association (APA, 2006, p. 918) is more explicit:

> The set of rules that describes how words and phrases in a language are arranged into grammatical sentences, or the branch of linguistics that studies such rules. With MORPHOLOGY, syntax is one of the two traditional subdivisions of grammar.

What is not explicit in either of these definitions is the underlying metaphor of a spatial array. It is analogous to the mainstream psycholinguist's use of the terms *line* and *linearization* (see *Merriam-Webster's*, p. 723) without any explicit reference to the underlying spatial metaphor. Levelt (1981, p. 305) has even entitled one of his articles *The speaker's linearization problem*, but without adverting to this underlying metaphor that refers us back to the written. The spatial linearization actually characterizes his graphic materials; it is not really a *linearization* problem on the part of the speaker, but rather a problem of temporal *sequencing*.

The reference to the written becomes crucial when we seek to explain the syntactic structure of spontaneous spoken discourse. In their book *Spontaneous spoken language: Syntax and discourse*, Miller and Weinert (1998) have argued that, in learning their native language, children do not learn the "magnasyntax" (p. 378 ff.) characteristic of the written language:

> Children learn informal spoken language with its own simple structures. They do not learn written language, . . . (The irony is that the theory of grammar which has had the greatest impact on theories of first language acquisition is based largely on written language, and not just on the written language of children's stories but on the complex written language of well-educated adults.) Moreover, the language of the various corpuses of informal conversation is not degenerate. (p. 383)

Miller and Weinert have deliberately set out to discover this non-degenerate structure of spontaneous spoken language: "The syntactic structure of clauses and phrases in spontaneous speech and. . . the grammatical devices that play a central role in the organization of spontaneous spoken discourse" (p. 1). But these structures and devices must be carefully distinguished from those of the written language:

> The linguistic properties spring from the various general properties of spontaneous spoken language – but the pragmatic nature of the general properties does not mean that the syntax of spontaneous spoken language is to be treated as resulting from performance error. The properties of spontaneous spoken language and the properties of formal written language both reflect the conditions under which they are produced. (p. 23)

Essentially, their position is in accord with that of Halliday (1989) who has maintained that "sentences are inappropriate for the analysis of spontaneous

spoken language and works with clause complexes" (Miller & Weinert, 1998, p. 28). This effectively reduces the sentence to "a low level discourse unit of written language" and leaves the way open to the characterization of the structure of spontaneous spoken language as "blocks of syntax with little or no syntactic linkage and requiring from the listener a larger than usual exercise of inference based on contextual and world knowledge." Miller and Weinert have emphasized in their rejection of the sentence as a proper analytic unit for spontaneous spoken language their opposition to the approach of mainstream psycholinguistics: "Models of speech production, such as the one presented in Levelt (1989), are based on the analysis of the syntax of written language; the basic syntactic unit is the sentence" (p. 420). Their own presentation has been intended to "demonstrate that spontaneous spoken language has its own regular characteristics of grammar and discourse organization."

Beyond Sentential Syntax

As psychologists of language use, we begin our own characterization of spontaneous spoken discourse where Miller and Weinert's ends. Their focus remains with the clauses, phrases, and "phrase complexes," whereas ours concentrates on the "contextual and world knowledge." In fact, our own preoccupations go much further – to what we are tempted to refer to as the somaticization of syntax. For the very structure and functions of spontaneous spoken discourse are dependent on far more than even the listener's "inference based on contextual and world knowledge." Spontaneous spoken discourse, as it proceeds in real time, is inherently embedded – for both speaker and listener – in the material surround, including the corporality of the interlocutors themselves, and in the entire ambience of socio-cultural and personal life. This embeddedness need not be the object of a focal consciousness on the part of either speaker or listener; it is, however, always operative and must therefore be taken into account by researchers.

We wish, then, to begin with an example of the complexity of embeddedness. The first author's first telephone conversation with Carla ended abruptly with her articulation of "Enough!" as she walked away from the phone. On the face of it, such a description reads like a wholesale rejection of an interlocutor, but underlying this bald description is a delightful scenario. The import of this single embedded word is completely other than might be expected. One must know that the interlocutor with the last word was three years old, that she had never met the first author, and that it was hilariously funny to her entire family. Note too that the single word "Enough!" could indeed have been imbued (prosodically) with anger, disdain, disinterest, or failure to understand. But none of these was the case; the single word was a simple notification that the conversation was over for her and that young Carla was proceeding to other activities.

The example may perhaps seem somewhat overdrawn to the reader. But, in fact it is entirely typical of the richness of spontaneous spoken discourse, a richness *that is deleted from the individual linguistic elements (words, phrases, clauses, sentences) as the conversation is transcribed onto paper and becomes an object of research*. The intelligibility of spontaneous spoken discourse – along with its structure and functions – is simply not comprehensively transcribable. It lies elsewhere than in the words alone – particularly after they have been violently disembodied and disenfranchised onto paper. It is precisely for this reason that the telephone conversation with young Carla had to be *described* rather than simply *transcribed*.

In somaticized, embedded spontaneous spoken discourse, everything counts as a determinant of the structure and functions of the utterance: gestures, bodily movement and stance, prosodic variations, socio-cultural context, both mutual and common gaze, both idiolectal and institutionalized ritual, the physical setting, as well as a host of additional psychological and sociological conditions. And yes, then there are the words, phrases, clauses, and clause complexes mentioned by Miller and Weinert. But sentences as such prove to be foreign to spontaneous spoken discourse – odd imports from the written world, or, as Miller and Weinert (1998, p. 28) have referred to them, "a low level discourse unit of written language." If this all sounds quite radical, we acknowledge that it is intended to be. As Linell (1998, p. xiii) has pointed out, the tradition of the language sciences is diametrically opposed to such an approach to spontaneous spoken discourse:

> a long-standing tradition in the language sciences in general, and in linguistics in particular, to give priority to theories and methods that suit written language and text better than they do spoken interaction.

And yet, our radical thoughts are not without precedent. What we have referred to as the somaticization of syntax is precisely what, according to Rommetveit (1974, p. 62), Roman Jakobson has referred to as "meta-linguistic operations" and Birdwhistell has referred to as "the 'integrational' and the 'new informational' aspects of the entire, multifaceted process of interaction." The most recent expression of this expanded domain of function and structure in spoken discourse that we have been able to find is from Docherty and Khattab (2008, p. 603):

> In recent years there has been a sharply growing awareness that developing our understanding of how speaker performance is shaped by extralinguistic factors associated with particular communicative situations is fundamental in building models of speech production, . . .

Even when empirical research must divide and conquer in the sense that it must select specific details of this daunting array, it must still keep the overarching integrating principles of somaticization in perspective. Only thus will research be able to cumulate in a productive manner; otherwise, triviality and re-invention of the wheel ensue.

Back to Psychology

Our motivation throughout this book has been to reinstate genuinely psychological considerations to the analysis of spontaneous spoken discourse. The sole treatment of spontaneous spoken discourse as a product of cognitive machinery or of information processing or of conditioning or of neurological mechanisms or even of all of the above we would consider inadequate. They are all involved in speech production; that is not at issue. But none of them address the basic psychological characteristics of spontaneous spoken discourse that we have discussed above. This is why we have been so insistent on the neglected components of spontaneous spoken discourse that we have outlined in the chapters of Part IV Toward a Theory of Spontaneous Spoken Discourse: *Intersubjectivity, Perspectivity, Open-endedness*, and *Verbal Integrity*.

In this chapter, we wish to argue that these components of spontaneous spoken discourse are all continuously involved in the interaction of speaker and listener and determine both meaning and understanding. To the extent that one or another of these components is neglected by the researcher, the theoretical understanding of dialogicality languishes or ceases. All four of these components are in a sense facets of the same diamond and cannot be isolated one from another. For example, in terms of the speaker's and listener's roles, if one hears the plaint from the speaker, "You're not listening," most likely all four components are being violated by the listener: (1) Clearly, the dozing interlocutor is lacking in intersubjectivity; (2) he or she may be manifesting a rather narcissistic perspective; (3) the conversation is not going anywhere, is certainly not open-ended; and (4) the dozing is disrespectful of the speaker, in violation of verbal integrity. But such a diagnosis would, of course, be completely wrong on the part of the speaker if the listener had simply missed something because he or she is hard of hearing. Such a situation is indeed a not infrequent problem in verbal interaction, but is not eo ipso a violation of the conditions for spontaneous spoken discourse. It simply places an additional burden on both speaker and listener, respectively, to articulate more efficiently (in particular, with an adequate loudness level) and to attend and interact as alertly as possible. Long range, it may require additional means such as hearing aids, elimination of ambient masking noise, or even a writing pad, but the necessary characteristics of spontaneous spoken discourse can be maintained throughout this entire process of adjustment.

Spontaneous spoken discourse is preeminently presence to one another, but not just physical presence, or even interactive presence as in playing the game of H-O-R-S-E on the basketball court or tic-tac-toe with paper and pencil. One may note that both these games involve verbal means, even though they do not qualify as spontaneous spoken discourse. In order to qualify, the interaction must be focally verbal, that is, it must use verbal means to communicate with the person or persons present at the moment. *Merriam-Webster's collegiate dictionary* (11th ed., 2003, p. 357) has succinctly defined discourse as "verbal

interchange of ideas." But as a matter of fact, much more than "ideas" is shared in discourse. And once again, for the psychologist, the emphasis must be on the sharing interlocutors themselves.

Interestingly enough, few people seek to grasp the intelligibility of the world around them in the isolation of their own reflections; the eremitic contemplative and the isolated recluse are both oddities to most of us. Instead, we seek out the company of other human beings in order to share what we know, inquire about what we don't know, tell stories, plan the future, and spend what has come to be known as quality time with one another. Often, von Kleist's (1806/no date, p. 975; our translation) "gradual working-out of one's thoughts in the process of speaking" is involved – a process of thinking out loud, if you will, but in the presence of another person.

The first author recalls an excursion on the Rhine River with the families of Fulbright fellows. There were many toddlers on board who spoke a variety of languages other than German. Some of the adults noted that it took the children only a short time to begin the process of sharing, even though their interactions were limited to rather basic communications such as *hey*, *oh*, *hm*, not at all unlike Raschka's (1993, 1997) dialogue of two boys as was discussed in Chapter 6, and to lots of looking and touching and running and giggling. Even though it was clearly marginal, all of the adult spectators would likely have voted for including within the category of spontaneous spoken discourse this verbal interaction of children *without a common language*. All the mandatory characteristics of spontaneous spoken discourse were realized in their sharing and interaction – with the notable exception of sentential syntax.

A Theory

It is perhaps presumptuous to try to explain the structure and functions of such instances of verbal sharing by an entirely new theory that involves massive expansion of the meaning of syntax. But the evidence is equally massive: The meaning and coherence of spontaneous spoken discourse are dependent on what we have referred to as somaticization or embeddedness. This is the ultimate descriptor of spontaneous spoken discourse; it is preeminently human contact with the help of a verbal toolkit – whether the tools are those of the scholarly polyglot or those of the non-German-languaged children on our Rhine cruise.

Such an extension of syntax to the comprehensive setting of sequences of spontaneous spoken discourse is undoubtedly heterodox. Paradoxically, we are arguing for the incorporation of nonlinguistic elements into what has been, up until now, purely linguistic theorizing. The array of candidate elements for inclusion in such a theory is, in fact, infinite; there is no adequate criterion for excluding any factor that affects meaning and understanding. This virtual infinity does not exclude the possibility of focusing on salient factors in an

orderly fashion and cannot, therefore, be used as a legitimate counterargument to such a theory. For example, the occurrence of an occasional extrasystolic heartbeat might well lead to a momentary (sub-second) inadvertent change in articulation rate on the part of a speaker; but this is hardly the type of reliable occurrence to be incorporated into a theory of spontaneous spoken discourse. The lawfulness of adequately frequent re-occurrence is required for a scientific concept to be practically relevant, and this logic takes precedence over the relevance of the "occasional."

Throwaways

This brings us back to the 10 chapters of Part III Empirical Research on Spontaneous Spoken Discourse. Perhaps the reader has wondered why such odd, peripheral, and trivial response measures have been chosen for analysis. The answer is that we have found precisely these to be reliably frequent re-occurrences that merit incorporation into a new syntax of spontaneous spoken discourse. They do characterize spontaneous spoken discourse in a dominant way that has to date not been incorporated into a syntax of spontaneous spoken discourse. One reason for the thorough neglect of these phenomena has been the insistence that the syntax of spontaneous spoken discourse must emulate the syntax of written discourse.

Hesitation phenomena (repeats, false starts, fillers, pauses), interjections, and laughter are not throwaways in spontaneous spoken discourse; they are extraordinarily important sources of spontaneity, reflection, and the pursuit of communicative coherence in real time for both speaker and listener. They are in this function the very building blocks of a legitimate fluency and communicative effectiveness. Paradoxically, all of them are optional. One may use a number of different tools in order to make the best of the available time for speaking. For example, the same effect might be accomplished by the use of a repeat (e.g., the pronoun *he he*) or the insertion of a filler (e.g., *uh*), or even the combination of the two (i.e., *he uh he*). The fact that there are so many options makes the orderly analysis and interpretation of these response measures difficult, but not impossible. We have endeavored to show throughout the empirical chapters that the use of all these means can be both predictable and orderly, and above all productive of communicative effectiveness.

To characterize them as *throwaways*, as Erard (2004, January 2, p. A 13) has in the case of fillers, is not quite as negative as to plop all hesitations into the same pot as disruptive, deficient, and chaotic. But neither are they to be thought of as innocuous and irrelevant elements of spontaneous spoken discourse. They are, along with the entire available armamentarium of material, contextual, socio-cultural, and ritual devices, part of the somaticized syntax of spontaneous spoken discourse.

One need only imagine once again our young suitor who has just proclaimed his love and asked for the hand of a young woman in marriage. In the seconds between his proclamation and her response, her slightest move, gesture, sigh, smile, laughter, or tears are of utmost importance to him for his semantic peace of mind. It is literally a case of she-loves-me, she-loves-me-not. Should we counsel him to construe and decipher the syntactic well-formedness of her reaction? What if her first utterance is an *uh*? Wait a moment: What comes next? What if she begins to cry? Is this a bad sign or must it be construed along with all the other evidence? She may indeed be so overwhelmed with happiness that she must weep! Such are the subtleties of the syntax of spontaneous spoken discourse. It is not easy for the young man in his own anxiety to be patient and wait out her reply, without jumping to a rash conclusion. But her response may well require a verbal working-out of her deepest convictions and yearnings, and that may take some time. That the few seconds will seem an eternity to the young man is an eloquent witness to the subtle importance of time that we thoroughly neglect – most of the time.

We find it quite noteworthy that our search for examples typically returns us to the simplest, basic situations of life. One need not look afar, but it seems that we simply do not look – or listen. One could well retort that the response measures are not at all linguistic. The objection is quite legitimate: Some of these response measures are indeed not linguistic. However, the point to be made is that these very response measures are capable of profoundly affecting the meaning that a speaker intends and the corresponding understanding of a listener.

A more subtle and complex example that demonstrates such an influence on meaning and understanding can be found in the extended question posed by Walter Cronkite in a TV interview with then President Ronald Reagan. It has been discussed in Chapter 5 in the context of choosing effective devices for rhetorical performance in spontaneous spoken discourse; here we wish to examine it in more detail in light of our theory. The transcript excerpt includes silent pauses, measured in seconds according to duration, as well as repeats, fillers, and false starts. It should be noted that pauses are not necessarily hesitation phenomena; they also serve other functions such as breathing, rhetoric, and juncture. However, all these devices are instruments of *discontinuity*. Literally, what Cronkite asked was:

> What what really philosophically is different (.8) from (.3) our (.43) going down to help a a (.3) democratic government uh (.37) sustain itself against guerilla activity (.27) promoted from the outside (.43) uh Soviet and Cuban uh uh aid as we believe it to be (.27) or as your your (.33) administration says it is (.7) and (.43) Afghanistan (.27) uh the El Salvador is in our sort of geo (.4) political sphere of influence Afghanistan on the border of the Soviet Union is certainly in their geo (.43) political sphere of influence (.47) uh they went in with troops uh to uh uh to support a Marxist government friendly to them (.63) what what's where where's the where why isn't that a parallel situation. (Kowal et al., 1985, p. 15)

If we examine only the sequence of words on paper, they do not appear to be the well-formed discourse of one of the leading TV journalists of America at that

time. It is a stumbling articulation of a rambling question that is worked out in the course of time. And this course of time involves the use of a number of devices (7 repeats, 16 pauses, 4 false starts, and 9 fillers) distributed across 99 words, i.e., 36/99 or a rate of more than one such device for every three words. And our question must be: Does the very extension in time brought about by these devices of discontinuity contribute in some *positive* way to an appropriate and effective communication with President Reagan on the part of Cronkite? This is what our theory of the syntax of spontaneous spoken discourse would suggest. And surprisingly perhaps, that is exactly the case: The tactic worked. Whether some radically different tactic would also have worked must remain pure speculation. The question posed by Cronkite was a very sensitive one – almost an accusation on his part. Reduced to its essentials, he was asking: "Aren't we doing the same thing as the Russians?" The president could well have been expected to bristle and refuse to engage it, had Cronkite asked the question directly and bluntly in this reduced form. However, Cronkite's hesitant gradualism communicated a certain modesty and an almost apologetic tone. The consequence of this approach on the part of Cronkite was the following: "Reagan laughs aloud in the course of an extensive reply" (Kowal et al., 1985, p. 15). In terms of our theory, what Cronkite has accomplished in this rambling run-up to his question is: (1) to engage *intersubjectively* the potential defensiveness of President Reagan; (2) to recognize the corresponding *perspective* that the president must have and at the same time to work toward the expression of his own alternative *perspective*; (3) to cut a path to a possible *open-endedness* in further dialogue rather than experience a defensive dead-end; and finally, (4) to approach his interlocutor with *verbal integrity*, i.e., with modesty instead of the *HARDtalk* characteristic of a Tim Sebastian. This is not to say that there is no place for *HARDtalk* in political interviews, but only that it would have been both impertinent and ineffective in this setting. In addition to Reagan's laughing aloud and offering an extensive answer, he later concluded the interview with the epigraph we have used for Chapter 3: "Uh Walter that uh I know you're running out of time and here I am hemming and hawing" (Kowal, 1991, p. 147) – an eloquent, and hesitant, acknowledgment of Cronkite's successful interview.

In order to provide further empirical support for our interpretation of Cronkite's performance above, we offer the following comparison with Dan Rather, another well-known TV interviewer in an interview of his own with President Reagan (Kowal et al., 1985). An analysis of the normalized number of hesitations used in their respective interviews has shown that Rather used far fewer hesitations than did Cronkite ($1.97 < 6.66$ hesitations/100 syl; p. 10). Kowal et al. have concluded that Rather manifests here an interviewing style very different from Cronkite's, one "with very tight control over the interview," as shown precisely in his use of hesitations. Instead of using many rather long hesitations, he made use mostly of hesitations that were "syllable fragments" (e.g., "y-you"; p. 15). In view of Rather's tight control, it is not surprising that he locked horns with former U.S. President George Bush in a much discussed, unpleasant

stand-off. It has been analyzed by a number of scholars in the 1988/1989 volume (22) of *Research on Language and Social Interaction*, as a special issue.

One should note that one or another of Cronkite's means or devices could well have been deleted or substituted for by another means of time extension. But the overall pattern of gradual working-out what Cronkite had to say (or ask) was clearly effective for him to prevail in obtaining the ample reply he was requesting.

Almost 50 years ago, Maclay and Osgood (1959, p. 322 f.) provided an initial thrust to research on hesitation phenomena in spontaneous speech. In their long list of directions for further research, they included an emphasis on probability and location of occurrence. They also mentioned the importance of the listener, but they neglected the interactive nature of dialogue. It would appear that mainstream psycholinguistics is the current inheritor of this tradition, in which we too were involved for many years. Our own presentation above of Walter Cronkite's very hesitant question addressed to President Reagan manifests our very different current approach to the analysis of hesitant spoken discourse. Our theoretical emphasis now is on the intentional and purposeful integration of hesitation into a speaker's gradual working out of what he or she wishes to communicate, i.e., as a tool of effective communication and indeed of eloquent rhetoric – as Cronkite well exemplifies.

In precisely this vein, we have found a passage of eloquent praise for spontaneous spoken discourse from a master of written English style whose essays in the *New Yorker* magazine are well known. Louis Menand (2004, November 8, p. 104) has summed up much better than we could what is most important about spontaneous spoken discourse. And in doing so, he has described the beauty of this everyday human interactive behavior:

> Writers often claim that they never write something that they would not say. It is hard to know how this could be literally true. Speech is somatic, a bodily function, and it is accompanied by physical inflections – tone of voice, winks, smiles, raised eyebrows, hand gestures – that are not reproducible in writing. Spoken language is repetitive, fragmentary, contradictory, limited in vocabulary, loaded down with space holders ("like," "um," "you know") – all the things writing teachers tell students not to do. And yet people can generally make themselves understood right away. As a medium, writing is a million times weaker than speech. It's a hieroglyph competing with a symphony.

High praise coming from someone who is himself a professional writer! It is noteworthy that even in this lofty passage of praise for spontaneous spoken discourse, Menand has written about "space holders" that are really "time holders." Even here, we have not escaped the visual, written language bias of the language sciences. And – fairly obviously – neither has the litterateur. For our part, we can only hope the hieroglyphics and chirographics that are this book have not competed with but have instead displayed some of the complex beauty of the symphony.

We wish to emphasize, finally, as we have already several times in the course of this book, "*the gradual working-out of one's thoughts in the process of speaking*" (von Kleist, 1806/no date, p. 975; our translation) in a thoroughly dialogical and

oral setting. Thus, the gradualism of spontaneous spoken discourse, its construction of fluency through the very instrumentality of discontinuous devices, the somaticization of rhetorical procedures, and the use of abundant options – all these characterize the vast majority of our everyday uses of language in ways radically different from what mainstream psycholinguistics and other language sciences have to date been able to tell us.

Splendid Isolation?

Are we left then with a view of spontaneous spoken discourse that is, like British foreign policy in the late nineteenth century, cornered into splendid isolation? Or is it still possible to integrate our position with much current thinking in the social sciences and indeed with the very origins of modern psychology? In fact, we have been pleasantly surprised at the many scholars who have voiced theoretical positions very much in accord with the tenor of this book. Once again, the reader will note that most of them are European rather than American scholars.

According to Graumann (2006, p. 61; our translation) in his discussion of Wundt's relevance for a currently viable social psychology of language use, "every intentional act reaches beyond the consciousness that produces it," and this makes it difficult to determine the "borders of individual existence." These intentional acts are largely the spontaneous spoken discourse that has been our topic in this book – or, as Graumann has written, they are the "utterances" that can then never be withdrawn by the speaker. For utterances become the intellectual property of interlocutors – to be understood, changed, falsified by them as they wish. Thus, a new, conscious, social phenomenon arises from every utterance, something interactive in its very essence, a shared communication. Graumann has traced such a conceptualization back in time to both Wilhelm Wundt (1911) and George H. Mead (1980) as an interactive psychology of language use that transcends the individual psychology of a tradition that has as its outside boundary the consciousness of the individual experimental subject. Only beyond that boundary is it possible to conceptualize a truly dialogical discourse, "a movement of response that develops derived from a common movement" (Wundt, 1911; cited in Graumann, 2006, p. 62; our translation).

Linell (2005, p. 214), in his more recent incarnation of *The written language bias in linguistics* (see Linell, 1982), has given voice to similar conceptualizations:

> Dialogism is a general epistemology for conceptualizing human action, communication, cognition and language. It is primarily relevant for the cultural sciences, rather than the natural sciences, ... neither language nor cognition can be liberated from the body, emotions and preconceptual structures.

Indeed, language use cannot be "liberated from the body." Linell has referred this dialogical "inter-world" back to Merleau-Ponty (1955), to Rommetveit (1998a, 1998b, 1999), and finally to Voloshinov (1973, p. 26):

> By its very existential nature, the subjective psyche is to be localized somewhere between the organism and the outside world, on the *borderline* separating these two spheres of reality.

In addition, Linell has aligned himself with the conversation-analytic tradition with regard to its emphasis on "situated interaction," or, in his own words, "*embodied*" interaction (p. 19), but he has faulted that tradition for its over-emphasis on local management to the neglect of "situation-transcending social practices" (p. 217). He has also acknowledged others' terminology for these situation-transcending social practices, such as "*sedimentation*" (Berger & Luckmann, 1967), "*entrenchment*" (Langacker, 1987), and "*analogy*" (p. 217). We ourselves have preferred the term *embeddedness* because it reflects our emphasis on the somaticization of spontaneous spoken discourse.

In short, Linell (2005, p. 224) has opted for an approach to language, and especially "to spoken interactional language" based on "very different data than has been traditionally used in linguistics":

> Instead of written data, made-up sentences and linguistic intuitions, we need to work with situated talk, that is talk with which people do things in all the diversified situations of real life.

That is exactly what our book has been all about. We have made use of this incredibly rich database for a modest beginning of our process of learning something about spontaneous spoken discourse – both in theory and in practice. That process of learning will never be comprehensively complete within the empirical disciplines that address language use.

Chapter 24
Communicating in Print about Communicating Orally

> *We – readers of books such as this – are so literate that it is very difficult for us to conceive of an oral universe of communication or thought except as a variant of a literate universe. This book will attempt to overcome our biases in some degree and to open new ways to understanding (Ong, 1982, p. 2).*

Well then, is it at all possible to communicate scientifically, i.e., within the science of psychology, about communication that makes use of oral language? We think it is indeed possible, and that is why this book has been written. Although the variability both of language and of the use of spontaneous spoken discourse is socio-culturally colossal, orderliness and even scientific lawfulness can be discovered therein, not perhaps the lawfulness of the natural sciences, but a truly scientific understanding nonetheless. The paradoxical challenge for research on this topic, as we have emphasized throughout this book, is that researchers who are well trained, literate people have to struggle against the biases in both methodology and theory that result precisely from their own – our own – competence with written language. Only then can they do justice to spontaneous spoken discourse. A simple example will serve to pinpoint this problem: When one asks orally – and the reader should carefully note that we can only express the question here in the written mode – how horseradish is spelled, the correct, although somewhat awkward answer is that *it isn't spelled at all* – i.e., the spoken version itself is not divided up into letters of the alphabet. Nothing in the oral world is spelled, for the simple reason that orthography is precisely a written phenomenon; it is a graphing of sorts, a spatialization of something non-spatial.

Within the community of mainstream psycholinguists, we have thus far not found much interest in the use of spoken language by adults for purposes of both public and everyday communication. Quite the contrary, we have found that their controlled experiments generally lack a basic ecological validity insofar as such experiments do not involve genuine dialogue; their concentration on the written is biased rather than balanced; their concentration on sentential structure

D.C. O'Connell, S. Kowal, *Communicating with One Another*,
DOI: 10.1007/978-0-387-77632-3_24, © Springer Science+Business Media, LLC 2008

disregards the various structures contributed by non-sentential phenomena; their ideal of syntactic well-formedness reflects both monologism and conceptual literacy rather than openness to both dialogism and conceptual orality; and their neglect of the socio-cultural embeddedness of language use is an important limitation. Also, the current importance of neural-cognitive research – welcome as it is for the discovery of neural relationships to behavior – remains, at least for now, considerably distant from a truly psychological study of spontaneous spoken discourse. It is, of course, the case that mainstream psycholinguistics marches to the tune of another drummer; the *status quaestionis* is about the processing of a language system. The legitimacy of such a scientific endeavor is not at all in question, only its ability thus far to engage psychologically comprehensive questions of communication with spoken language.

We have also been rather critical of the conversation-analytic tradition, in particular of the following ethnomethodological characteristics: qualitative analyses, for the most part to the exclusion of quantitative analyses; a corresponding dearth of inferential statistics; problematic transcription notations; and limitation to the use of anecdotal evidence. Nonetheless, the conversation-analytic tradition has brought a breath of fresh air to the psychological investigation of communication by means of spoken language. These scholars may not as yet have solved the problems of surreptitious vs. in-your-face collection of data, but they do investigate genuine spoken discourse. Their emphasis on "local management" is only part of the story, but it is clearly a significant part of the story. Their own interest has not been per se psychological, and yet their evidential base is extremely important for a psychological approach to spoken communication – to talk-in-interaction, to use their own terminology. Their insistence on the centrality of conversation in research has been succinctly stated by Schudson (1978, p. 327 f.): "Conversation holds a primary place in our lives and our sense of reality which neither novels nor newspapers nor film nor television can match."

Clearly, some sort of rapprochement is needed, some *modus agendi*, some theoretical orientation that can bridge the gap between mainstream psycholinguistics and the conversation-analytic tradition, while avoiding the extremes and pitfalls of previous generations of research that have been grist for the mill of the preceding pages. We have mentioned Herbert H. Clark frequently in this book; it is to his credit that – to our knowledge – he, more than anyone else in this domain of research, has endeavored to distill the best of both the mainstream psycholinguistic and the conversation-analytic traditions. We ourselves have most certainly not devised such a psychological science of spontaneous spoken discourse for the future. But we are nonetheless quite convinced that we have found a firm footing for it – both in the history as well as in current research and theory on spontaneous spoken discourse.

A relatively small band of scholars – originally mostly psychologists and sociologists, and almost entirely in Europe rather than America – has been leading the way for several decades now. Their heritage includes Karl Bühler, Irving Goffman, Hans Hörmann, Moritz Lazarus, Walter Ong, Wilhelm

Wundt, and many less well known scholars; their current membership includes Klaus Foppa, Carl F. Graumann († 2007), Roy Harris, Per Linell, Ivana Marková, Ragnar Rommetveit, and Talbot Taylor. We are deeply grateful that they have led research in the direction of socio-cultural, field observational, qualitative as well as quantitative, inferential analyses of spontaneous spoken discourse. Above all, we are grateful for Hörmann's emphasis – more than 30 years ago – on the intrinsic dynamic of human communication: Both the intended meaning of the speaker and the creative understanding of the listener are dependent not only on linguistic structure but also on the very embeddedness of spoken discourse in the various antecedent and concurrent aspects of the nonlinguistic setting. It is, then, to Hörmann's theoretical stance that we owe our own interest in a number of phenomena quite marginal to traditional linguistics; these phenomena would certainly include interjections, fillers, and laughter. In other words, the estate that we have inherited as we set out to write this book has been a very rich one. What we have endeavored to pull together into one place has largely been written or said already – as eloquently witnessed by the veritable glut of citations in this book – but in scattered, mostly European, and in any event, unfortunately unread and unheard sources. And the reason for the lack of interest in the writings of these scholars to date is, for the most part, American provincialism.

We are reminded of Ronald Reagan's summary of rhetorical devices: Tell them what you're about to tell them, then tell them, and then tell them what you've told them. We are now in the position of telling the reader what we think has been most important in what we have told you. Here is the list for your inspection and criticism:

(1) Spontaneous spoken discourse is not inevitably deficient, incoherent, pathological, chaotic, and disorderly. Quite the contrary, when investigated in its own right, it is found to transcend in its structure the sentential syntax of well-formed written text. It has at its disposal a vast array of verbal, prosodic, temporal, concrete, situational, socio-cultural tools largely lacking in the armory of written language use. We have referred to this new view of syntax – again, not original with us – as embodied (somaticized) and embedded in order to emphasize its otherness from sentential structure.
(2) Among the tools of spoken communication are the phenomena that have served as material for our empirical chapters: the various hesitation phenomena (repeats, pauses, false starts), prosody, fillers, interjections, devices for referring, turn-taking, laughter, and both affiliative and disaffiliative audience responses. To these must be added a huge number of other nonverbal devices, including gestures, body movement, both common and mutual gaze, and silence that we have not engaged herein. They too are part of the actual embodied and embedded syntax of spontaneous spoken discourse that are the fundamental basis of meaning and understanding.
(3) Empirical methods are to be adapted creatively and imaginatively to the purposes of research on spontaneous spoken discourse. Hence, a

standardized notation system for transcription is *not* a desideratum; realistic, accurate (valid and reliable), useable, purposeful notation systems are. Since spoken discourse proceeds in real time, *all* the empirical phenomena discussed in Part III have important temporal dimensions which must be assessed instrumentally, especially when fine-tuned, sub-second durations are in question. Quantitative analyses are needed for purposes of statistical inference, while qualitative analyses remain necessary for the concrete exemplification of phenomena. Both remain subject to the interpretive integration and skillful use of hermeneutics on the part of a discerning scholar.

(4) Basic to our theory are the four neglected psychological components that underlie the empirical engagement of spontaneous spoken discourse: intersubjectivity, perspectivity, open-endedness, and verbal integrity. We have not introduced these concepts primarily at the level of empirical response measures; they are instead overarching necessary concomitants of all spontaneous spoken discourse. In this sense, they are pre-empirical: It is our assumption that all these components actively characterize human efforts to communicate with one another by speaking and listening. Nor should they be conceptualized as operating independently of one another. For example, the very perspectives of interlocutors are subject to change over time in accord with the lawfulness of open-endedness in the evaluative setting contributed by verbal integrity.

(5) There is no reason for the various disciplines that are concerned with research on spontaneous spoken discourse to be in conflict with one another. The empirical engagement should ultimately be a unifying endeavor. We ourselves have happily collaborated with colleagues in the areas of biology, drama, linguistics, literature, phonetics, semiotics, and sociology. Our ignorance has sometimes made cooperation a formidable challenge in such settings. But we are grateful for the expansion of our thinking and the enrichment of our research at the hands of these colleagues.

With Walt Disney's Porky Pig, we would take our leave as follows: "Th- th- th- that's all, folks." Interestingly enough, even Porky Pig has become the victim of the well-formedness and written language bias. The Google entry for Porky Pig, reads simply: "That's all, folks": Simplified, well-formed, sententially syntactic – but unreal, and certainly not genuine Porky Pig.

Carl Bernstein (Diening & Schuller, 2007, July 29, p. S 1; our translation), in an interview about his recently published biography of Hillary Clinton, has mirrored quite well our own feelings about this book: "Authors write in order to find out what they truly believe. Now I read for the first time the finished book as a whole, and I notice that I have moved on."

And from where have we moved on? Both of us must acknowledge that we are converted experimental psychologists who were deeply embedded in mainstream psycholinguistics for many years. Self-knowledge is a difficult virtue, and we still find ourselves from time to time thinking in the literate mode; as

Walter Ong has suggested in our epigraph to this chapter, such thinking is quite inevitable. Nonetheless, we wish to reflect in our written communication about orally communicating a striving toward a theory of ideal spontaneous spoken discourse.

Much of our book has been anticipated in germ by two sentences in *Advances in psycholinguistics* (Flores d'Arcais & Levelt, 1970, p. v), the report on the 1969 Bressanone Conference on Psycholinguistics almost four decades ago:

> The final selection of papers is in many respects not representative for psycholinguistics today. Rather, this volume will serve the purpose of directing the reader's attention to various new and lively trends in psycholinguistics.

This we have tried to do herein.

References

Abercrombie, D. (1965). *Studies in phonetics and linguistics*. London: Oxford University Press.

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh, Scotland: Edinburgh University Press.

Adelswärd, V. (1989). Laughter and dialogue: The social significance of laughter in institutional discourse. *Nordic Journal of Linguistics, 12*, 107–136.

Aitchison, J. (1983). *The articulate mammal: An introduction to psycholinguistics* (2 nd ed.). New York: Universe Books.

Akinnaso, F. N. (1982). On the differences between spoken and written language. *Language & Speech, 25*, 97–125.

Alber, J., O'Connell, D. C., & Kowal, S. (2002). Personal perspective in TV interviews. *Pragmatics, 12*, 257–271.

Allport, F. H. (1924). *Social psychology*. Boston: Houghton Mifflin.

Allwood, J., Nivre, J., & Ahlsén, E. (1990). Speech management – on the non-written life of speech. *Nordic Journal of Linguistics, 13,* 3–48.

Ameka, F. (Ed.). (1992a). Interjections [Special issue]. *Journal of Pragmatics, 18* (2–3), ix–301.

Ameka, F. (1992b). Interjections: The universal yet neglected part of speech. *Journal of Pragmatics, 18*, 101–118.

Ameka, F. (1994). Interjections. In R. E. Asher (Ed.), *The encyclopedia of language and linguistics* (Vol. 2, pp. 1712–1715). Oxford: Pergamon Press.

American Association for the Advancement of Science. (2007). Special section, decision making. *Science, 318*, 593–610.

American Psychological Association. (1985). Awards for distinguished scientific contributions: 1984. *American Psychologist, 40*, 285–300.

APA dictionary of psychology (2006). Washington, DC: American Psychological Association.

Arendt, H. (1992). Hannah Arendt to Karl Jaspers (Chicago, October 25, 1964). In L. Kohler & H. Saner (Eds.), *Hannah Arendt, Karl Jaspers: Correspondence, 1926–1969* (pp. 568–572). New York: Harcourt Brace Jovanovich.

Arendt, H. (1996). Fernsehgespräch mit Günter Gaus. In U. Ludz (Ed.), *Ich will verstehen: Selbstauskünfte zu Leben und Werk* (pp. 44–70). München: Piper.

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research, 32*, 25–36.

Atkinson, J. M. (1983). Two devices for generating audience approval: A comparative study of public discourse and texts. In K. Ehlich et al. (Eds.), *Connectedness in sentence, text and discourse* (pp. 199–236). Tilburg, Netherlands: Tilburg Papers in Linguistics.

Atkinson, M. (1984a). *Our master's voices: The language and body language of politics*. London: Methuen.

Atkinson, J. M. (1984b). Public speaking and audience responses: Some techniques for inviting applause. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 370–409). Cambridge: Cambridge University Press.

Atkinson, J. M. (1985). Refusing invited applause: Preliminary observations from a case study of charismatic oratory. In T. A. van Dijk (Ed.), *Handbook of discourse analysis: Vol. 3. Discourse and dialogue* (pp. 161–181). London: Academic Press.

Atkinson, M. (1986). The 1983 election and the demise of live oratory. In I. Crewe & M. Harrop (Eds.), *Political communications: The general election campaign of 1983* (pp. 38–55). Cambridge: Cambridge University Press.

Atkinson, J. M., & Heritage, J. (Eds.). (1984). *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press.

St. Augustine. (ca. 397-400/1960). *Confessions of Saint Augustine* (J. K. Ryan, Trans.). New York: Doubleday.

Austin, J. L. (1962). *How to do things with words*. Cambridge, MA: Harvard University Press.

B., R. (1997, August 3). EINSILBIG. *Der Tagesspiegel*, p. 4.

Baars, B. (Ed.). (1986). *The cognitive revolution in psychology*. New York: Guilford.

Bachorowski, J.-A., Smoski, M. J., & Owren, M. J. (2001). Not all laughs are alike: Voiced but not unvoiced laughter readily elicits positive affect. *Psychological Science, 12*, 252–257.

Bakhtin, M. M. (1981). *The dialogical imagination: Four essays by M. M. Bakhtin* (C. Emerson & M. Holquist, Trans. and Eds.). Austin: University of Texas Press.

Baldauf, H. (2002). *Knappes sprechen*. Tübingen: Niemeyer.

Ball, M. J., Perkins, M. R., Müller, N., & Howard, S. (Eds.). (2008). *Handbook of clinical linguistics*. Malden, MA: Blackwell.

Bally, C. (1951). *Traité de stylistique française*. Paris: Klincksieck.

Baron, N. S. (1981). *Speech, writing, and sign: A functional view of linguistic representation*. Bloomington: Indiana University Press.

Beattie, G. W. (1978). Sequential temporal patterns of speech and gaze in dialogue. *Semiotica, 23*, 29–52.

Bellebaum, A. (1992). *Schweigen und Verschweigen: Bedeutungen und Erscheinungsvielfalt einer Kommunikationsform*. Opladen, Germany: Westdeutscher Verlag.

Berdy, M. A. (2003). It's the little words that count. Retrieved August 3, 2003, from http://www.themoscowtimes.com/stories/2003/07/25/007-print.html

Berger, P., & Luckmann, T. (1967). *The social construction of reality*. Harmondsworth, UK: Penguin.

Bergmann, J. R. (1982). Schweigephasen im Gespräch: Aspekte ihrer interaktiven Organisation. In H.-G. Soeffner (Ed.), *Beiträge zu einer empirischen Sprachsoziologie* (pp. 143–184). Tübingen: Gunter Narr.

Bergmann, J. (1998). Introduction: Morality in discourse. *Research on Language and Social Interaction, 31*, 279–294.

Bergson, H. (1914). *Laughter: An essay on the meaning of the comic* (C. Brereton & F. Rothwell, Trans.). New York: Macmillan. (Original work published 1900).

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Biber, D., Conrad, S., & Leech, G. (2002). *Longman student grammar of spoken and written English*. London: Longman.

Birdwhistell, R. L. (1971). *Kinesics and context*. London: Allan Lane.

Birtwistle, S. (Producer), & Langton, S. (Director). (1995). *Pride and prejudice* (DVD of TV mini-series). UK: BBC.

Blank, A. (1991). *Literarisierung von Mündlichkeit: Louis-Ferdinand Céline und Raymond Queneau*. Tübingen: Gunter Narr.

Bloch, C. (1996). Emotions and discourse. *Text, 16*, 323–341.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart, & Winston.

Blumenthal, A. L. (1970). *Language and psychology: Historical aspects of psycholinguistics*. New York: Wiley.

Boas, F. (1940). *Race, language, and culture*. New York: Macmillan.

Bock, K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin and Review, 3*, 395–421.

Boersma, P., & Weenink, D. (2005). *Praat: Doing phonetics by computer (Version 4.3.14) [Computer program]*. Retrieved May 26, 2005, from http://www.praat.org/

Bohle, U. (2007). *Das Wort ergreifen – das Wort übergeben: Explorative Studie zur Rolle redebegleitender Gesten in der Organisation des Sprecherwechsels*. Berlin: Weidler.

Boland, J. E. (2005). Cognitive mechanisms and syntactic theory. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 23–42). Mahwah, NJ: Erlbaum.

Boomer, D. S. (1965). Hesitations and grammatical encoding. *Language & Speech, 8*, 148–158.

Borge, V. (1990). The best of Victor Borge. VCR Distributed by GMG Great Neck, NY.

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language & Speech, 44*, 123–147.

ten Bosch, L., Oostdijk, N. H. J., & de Ruiter, J. P. (2004). Durational aspects of turn-taking in spontaneous face-to-face and telephone dialogues. In P. Sojka, I. Kopeček, & K. Pala (Eds.), *Text, speech and dialogue: 7th international conference*, Brno, Czech Republic (pp. 563–570). Berlin: Springer.

Bourdieu, P. (1991). *Language and symbolic power*. Cambridge, MA: Harvard University Press.

Brady, P. T. (1965). A technique for investigating on-off patterns of speech. *Bell System Technical Journal, 44*, 1–22.

Brady, P. T. (1968). A statistical analysis of on-off patterns in 16 conversations. *Bell System Technical Journal, 47*, 3–91.

Bres, J. (1995). Hóu! Haa! Yrrââ: Interjection, exclamation, actualisation. *Faits de langues, 6*, 81–91.

Bringhurst, R. (1999). *A story as sharp as a knife: The classical Haida mythtellers and their world*. Vancouver, BC: Douglas & McIntyre; Lincoln, NE: University of Nebraska Press.

Brodine, R. (1986). Getting the applause. In R. M. B. Bosinelli (Ed.), *U.S. presidential election 1984: An interdisciplinary approach to the analysis of political discourse* (pp. 171–211). Bologna: Pitagoria Editrice.

Brown, E., & Miron, M. S. (1971). Lexical and syntactic predictors of the distribution of pause time in reading. *Journal of Verbal Learning & Verbal Behavior, 10*, 658–667.

Brown, G. (1995). *Speakers, listeners and communication: Explorations in discourse analysis*. Cambridge: Cambridge University Press.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Bruce, G. (1992). Comments. In J. Svartvik (Ed.), *Directions in corpus linguistics: Proceedings of the Nobel Symposium 82, Stockholm, August 4–8, 1991* (pp. 145–147). New York: Mouton de Gruyter.

Bruthiaux, P. (1993). Knowing when to stop: Investigating the nature of punctuation. *Language & Communication, 13*, 27–43.

Buber, M. (1958). *Ich und du*. Heidelberg: Lambert Schneider.

Buber, M. (1967). Das Wort, das gesprochen wird. In *Sprache und Wirklichkeit: Essays* (pp. 7–19). München: Deutscher Taschenbuch Verlag.

Bühler, K. (1934/1982). *Sprachtheorie*. Stuttgart: Fischer.

Bühler, K. (1990). *Theory of language: The representational function of language* (D. F. Goodwin, Trans.). Amsterdam: Benjamins. (Original work published 1934).

Bull, P. (1986). The use of hand gesture in political speeches: Some case studies. *Journal of Language and Social Psychology, 5*, 103–118.

Burckhardt, A. (1998). Interjektionen: Begriff, Geschichte(n), Paraphrasen. In T. Harden & E. Hentschel (Eds.), *Particulae particularum: Festschrift für Harald Weydt* (pp. 43–73). Tübingen: Niemeyer.

Burger, H. (1980). Interjektionen – eine Randwortart? In H. Sitta (Ed.), *Ansätze zu einer pragmatischen Sprachgeschichte* (pp. 53–69). Tübingen: Niemeyer.
Butcher, A. (1981). *Aspects of the speech pause: Phonetic correlates and communicative functions*. Kiel, Germany: Institut für Phonetik der Universität Kiel.
Cameron, D. (1995). *Verbal hygiene*. London: Routledge.
Cameron, D. (2007). *The myth of Mars and Venus*. Oxford: Oxford University Press.
Carroll, D. W. (1986). *Psychology of language*. Monterey, CA: Brooks/Cole.
Carroll, D. W. (2004). *Psychology of language* (4th ed.). Belmont, CA: Wadsworth/Thomson.
Carroll, D. W. (2007). *Psychology of language* (5th ed.). Belmont, CA: Wadsworth/Thomson.
Carpenter, S., & O'Connell, D. C. (1988). More than meets the ear: Some variables affecting pause reports. *Language & Communication, 8*, 17–27.
Cassirer, E. (1944). *An essay on man: An introduction to a philosophy of human culture*. New Haven: Yale University Press.
Cattell, J. M. (1885). Über die Zeit der Erkennung und Benennung von Schriftzeichen, Bildern und Farben. *Philosophische Studien, 2,* 635–650.
Cattell, J. M. (1886). The time it takes to see and name objects. *Mind, 11*, 63–65.
Chafe, W. L. (1980a). The deployment of consciousness in the production of a narrative. In W. L. Chafe (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production* (pp. 9–50). Norwood, NJ: Ablex.
Chafe, W. L. (1980b). Some reasons for hesitating. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 169–180). The Hague: Mouton.
Chafe, W. (1986). Writing in the perspective of speaking. In C. R. Cooper & S. Greenbaum (Eds.), Studying writing: Linguistic approaches (pp. 12–39). Beverly Hills, CA: Sage.
Chafe, W. (2003a). Laughing while talking. In D. Tannen & J. E. Alatis (Eds.), *Georgetown University round table on languages and linguistics 2001. Linguistics, language, and the real world: Discourse and beyond* (pp. 36–49). Washington, DC: Georgetown University Press.
Chafe, W. (2003b). Importance of not being earnest. Paper presented at the meeting of the International Society for Humor Studies, Chicago, IL.
Chafe, W. (2007). *The importance of not being earnest: The feeling behind laughter and humor*. Amsterdam: John Benjamins.
Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
Clancy, T. (2000). *The bear and the dragon*. New York: G. P. Putnam's Sons.
Clark, H. H. (1985). Language use and language users. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 2, pp. 179–231). New York: Random House.
Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasly (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: American Psychological Association.
Clark, H. H., & Clark, E. V. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.
Clark, H. H., & Fox Tree, J. E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition, 84*, 73–111.
Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology, 37*, 201–242.
Clayman, S. E. (1992). Caveat orator: Audience disaffiliation in the 1988 presidential debates. *Quarterly Journal of Speech, 78*, 33–60.
Clayman, S. E. (1993). Booing: The anatomy of a disaffiliative response. *American Sociological Review, 58*, 110–130.
Clayman, S., & Heritage, J. (2002). *The news interview: Journalists and public figures on the air*. Cambridge: Cambridge University Press.

Clemmer, E. J., O'Connell, D. C., & Loui, W. (1979). Rhetorical pauses in oral reading. *Language & Speech, 22*, 397–405.
Clinton, B. (2004). *My life*. New York: Alfred A. Knopf.
Clinton, H. R. (2003). *Living history*. New York: Simon & Schuster.
Cohen, A. A. (1987). *The television news interview*. Newbury Park, CA: Sage.
Colapinto, J. (2007, April 16). Has a remote Amazonian tribe upended our understanding of language? *New Yorker, 83*, 118–137.
Collins cobuild advanced learner's English dictionary. (2003). (4th ed.). Glasgow: HarperCollins.
Cook, G. (1990). Transcribing infinity: Problems of context presentation. *Journal of Pragmatics, 14*, 1–24.
Coser, R. L. (1960). Laughter among colleagues. *Psychiatry, 23*, 81-95.
Couper-Kuhlen, E. (2001). Intonation and discourse: Current views from within. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *Handbook of discourse analysis* (pp. 13–34). Malden, MA: Blackwell.
Couper-Kuhlen, E., & Selting, M. (Eds.). (1996a). *Prosody in conversation: Interactional studies*. Cambridge: Cambridge University Press.
Couper-Kuhlen, E., & Selting, M. (1996b). Towards an interactional perspective on prosody and a prosodic perspective on interaction. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in conversation: Interactional studies* (pp. 11–56). Cambridge: Cambridge University Press.
Crocker, M. W. (2005). Rational models of comprehension: Addressing the performance paradox. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 363–380). Mahwah, NJ: Erlbaum.
Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge: Cambridge University Press.
Crystal, D. (1997). *The Cambridge encyclopedia of language* (2 nd ed.). Cambridge: Cambridge University Press.
Crystal, D., & Davy, D. (1969). *Investigating English style*. London: Longman.
Crystal, D., & Quirk, R. (1964). *Systems of prosodic and paralinguistic features in English*. London: Mouton.
Curtius, E. R. (1949/1993). Essay. In J. Ortega y Gasset (Ed.), *Über die Liebe: Meditationen* (pp. 205–247). München: Deutscher Taschenbuch Verlag.
Cutler, A. (Ed.). (2005). *Twenty-first century psycholinguistics: Four cornerstones*. Mahwah, NJ: Erlbaum.
Cutler, A., & Pearson, M. (1986). On the analysis of prosodic turn-taking cues. In C. Johns-Lewis (Ed.), *Intonation in discourse* (pp. 139–155). London: Croom Helm.
Dauenhauer, B. P. (1980). *Silence: The phenomenon and its ontological significance*. Bloomington, IN: Indiana University Press.
Deese, J. (1984). *Thought into speech: The psychology of a language*. Englewood Cliffs, NJ: Prentice-Hall.
Denes, P. B., & Pinson, E. N. (1963). *The speech chain: The physics and biology of spoken language*. Baltimore, MD: Bell Telephone Laboratories.
Derwing, B. L. (1979). Against autonomous linguistics. In T. A. Perry (Ed.), *Evidence and argumentation in linguistics* (pp. 163–189). Berlin: de Gruyter.
Diening, D., & Schuller, M. (2007, July 29). Hillary führt Listen mit den Namen ihrer Feinde. *Der Tagesspiegel*, p. S 1.
Dietrich, R. (2002). *Psycholinguistik*. Stuttgart: Metzler.
Dillon, G. L. (1976). Clause, pause, and punctuation in poetry. *Linguistics, 169*, 5–20.
Docherty, G., & Khattab, G. (2008). Sociophonetics and clinical linguistics. In M. J. Ball, M. R. Perkins, N. Müller, & S. Howard (Eds.), *The handbook of clinical linguistics* (pp. 603–625). Malden, MA: Blackwell.
Dorval, B. (Ed.). (1990). *Conversational organization and its development*. Norwood, NJ: Ablex.
Dreher, W. (1982). *Gesprächsanalyse: Macht als Kategorie männlichen Interaktionsverhaltens: Sprecherwechsel und Lachen*. Unpublished master's thesis, Free University Berlin, Berlin, Germany.

Du Bois, J. W. (1991). Transcription design principles for spoken discourse research. *Pragmatics, 1*, 71–106.

Duden Deutsche Rechtschreibung. (2004). Mannheim: Dudenverlag.

Dulany, D. E., & O′Connell, D. C. (1963). Does partial reinforcement dissociate verbal rules and the behavior they might be presumed to control? *Journal of Verbal Learning & Verbal Behavior, 2*, 361–372.

Duncan, S., & Fiske, D. W. (1977). *Face-to-face interaction: Research, methods, and theory*. Hillsdale, NJ: Erlbaum.

Duranti, A. (1991). Four properties of speech-in-interaction and the notion of translocutionary act. In J. Verschueren (Ed.), *Pragmatics at issue: Selected papers of the International Pragmatics Conference, Antwerp, August 17-22, 1987* (Vol. I, pp. 133–150). Amsterdam: Benjamins.

Duranti, A. (1997). *Linguistic anthropology*. Cambridge: Cambridge University Press.

Duranti, A. (2003). The voice of the audience in contemporary American political discourse. In D. Tannen & J. E. Alatis (Eds.), *Georgetown university round table on languages and linguistics 2001. Linguistics, language, and the real world: Discourse and beyond* (pp. 114–134). Washington, DC: Georgetown University Press.

Eastman, C. M. (1992). Swahili interjections: Blurring language-use/gesture-use boundaries. *Journal of Pragmatics, 18*, 273–287.

Ebbinghaus, H. (1885/1992). *Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Eco, U. (1986, August 29). Eine Palette von Grautönen. *Die Zeit*, p. 51.

Edmondson, M. S. (1987). Notes on laughter. *Anthropological Linguistics, 29*, 23–34.

Edwards, J. A. (1989). *Transcription and the new functionalism: A counterproposal to CHILDES' CHAT conventions* (Berkeley Cognitive Science Report No. 58), Berkeley: Institute of Cognitive Studies, University of California, Berkeley.

Edwards, J. A. (1993). Principles and contrasting systems of discourse transcription. In J. A. Edwards & M. D. Lampert (Eds.), *Talking data: Transcription and coding in discourse research* (pp. 3–31). Hillsdale, NJ: Erlbaum.

Ehlich, K. (1986). *Interjektionen*. Tübingen: Niemeyer.

Ehlich, K. (1993). HIAT: A transcription system for discourse data. In J. A. Edwards & M. D. Lampert (Eds.), *Talking data: Transcription and coding in discourse research* (pp. 123–148). Hillsdale, NJ: Erlbaum.

Ellis, A., & Beattie, G. (1986). *The psychology of language and communication*. New York: Guilford Press.

Enfield, N. J. (2007). Meanings of the unmarked: How 'default' person reference does more than just refer. In N. J. Enfield & T. Stivers (Eds.), *Person reference in interaction: Linguistic, cultural, and social perspectives* (pp. 97–120). Cambridge: Cambridge University Press.

Enfield, N. J., & Stivers, T. (Eds.). (2007). *Person reference in interaction: Linguistic, cultural, and social perspectives*. Cambridge: Cambridge University Press.

Ensik, T., & Sauer, C. (2003). *Framing and perspectivizing in discourse*. Amsterdam: Benjamins.

Erard, M. (2004, January 2). Just like, er, words, not, um, throwaways. *New York Times*, pp. A 13, A 15.

Erard, M. (2007). *Um... Slips, stumbles, and verbal blunders, and what they mean*. New York: Pantheon.

Ervin-Tripp, S. (1979). Children's verbal turn-taking. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 391–429). New York: Academic Press.

Evans, V., & Green, M. (2006). *Cognitive linguistics: An introduction*. Mahwah, NJ: Erlbaum.

Ferber, R. (1991). Slip of the tongue or slip of the ear? On the perception and transcription of naturalistic slips of the tongue. *Journal of Psycholinguistic Research, 20*, 105–122.

Field, J. (2003). *Psycholinguistics: A source book for students*. London: Routledge.

Field, J. (2004). *Psycholinguistics: The key concepts*. London: Routledge.
Firth, J. R. (1968). *Selected papers of J. R. Firth 1952–59* (F. R. Palmer, Ed.). London: Longman.
Fitch, W. T. (2005). Computation and cognition: Four distinctions and their implications. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 381–400). Mahwah, NJ: Erlbaum.
Flores d'Arcais, G. B., & Levelt, W. J. M. (1970). Preface. In G. B. Flores d'Arcais & W. J. M. Levelt (Eds.), *Advances in psycholinguistics* (pp. v–vii). Amsterdam: North-Holland.
Fodor, J. (2000). *The mind doesn't work that way*. Cambridge, MA: MIT Press.
Foppa, K. (1994). About the psychology of dialogue – even though psychologists are not interested in it. In E. Weigand (Ed.), *Concepts of dialogue: Considered from the perspective of different disciplines* (pp. 145–158). Tübingen: Niemeyer.
Foppa, K. (2002). Knowledge and perspective setting: What possible consequences on conversation do we have to expect? In C. F. Graumann & W. Kallmeyer (Eds.), *Perspective and perspectivation in discourse* (pp. 15–23). Amsterdam: Benjamins.
Foppa, K. (2003). Verständigung über das Verstehen: Der dialogische Dreischritt. In D. Weimer & M. Galliker (Eds.), *Sprachliche Kommunikation: Ansätze und Perspektiven* (pp. 49–53). Heidelberg: Asanger.
Fries, N. (1988). Interjektionen und Interjektionsphrasen. *Sprache und Pragmatik*, 17, Arbeitsberichte. Lund.
Fries, N. (1990). Zur Grammatik von Interjektionen. In E. Feldbusch, R. Pogarell, & C. Weiss (Eds.), *Neue Fragen der Linguistik: Akten des 25. Linguistischen Kolloquiums 1* (pp. 283–295). Paderborn/ Tübingen: Niemeyer.
Fries, N. (1992). Interjektionen, Interjektionsphrasen und Satzmodus. In I. Rosengren (Ed.), *Satz und Illokution I* (pp. 307–341). Tübingen: Niemeyer.
Frommer, J. G., & Ishikawa, W. (1980). Alors. . . euh. . . on parle français. *The French Review, 53*, 501–506.
Funkhouser, L. (1979a). Acoustical rhythms in cummings' "Buffalo Bill's." *Journal of Modern Literature, 7*, 219–242.
Funkhouser, L. (1979b). Acoustical rhythms in Randall Jarrell's *The Death of the Ball Turret Gunner*. *Poetics, 8*, 381–403.
Funkhouser, L., & O'Connell, D. C. (1978). Temporal aspects of poetry readings by authors and adults. *Bulletin of the Psychonomic Society, 12*, 390–392.
Gadamer, H.-G. (1975). *Truth and method* (G. Barden & J. Cumming, Trans. and Eds.). New York: Seabury Press.
Gallmann, P. (1996). Interpunktion (Syngrapheme). In H. Günther & O. Ludwig (Eds.), *Schrift und Schriftlichkeit/Writing and its use: Ein interdisziplinäres Handbuch/An interdisciplinary handbook of international research* (pp. 1456–1467). Berlin: de Gruyter.
Garman, M. (1990). *Psycholinguistics*. Cambridge: Cambridge University Press.
Garrod, S. (1999). The challenge of dialogue for theories of language processing. In S. Garrod & M. J. Pickering (Eds.), *Language processing* (pp. 389–415). Hove, UK: Psychology Press.
Garrod, S., & Pickering, M. J. (1999). Issues in language processing. In S. Garrod & M. J. Pickering (Eds.), *Language processing* (pp. 1–11). Hove, UK: Psychology Press.
Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences, 8*, 9–11.
Garvey, C., & Berninger, G. (1981). Timing and turn taking in children's conversations. *Discourse Processes, 4*, 27–57.
Gaus, G. (2004). *Widersprüche: Erinnerungen eines linken Konservativen*. München: Propyläen.
Gee, J. P. (1986). Units in the production of narrative discourse. *Discourse Processes, 9*, 391–422.
Geier, M. (2006). *Worüber kluge Menschen lachen: Kleine Philosophie des Humors*. Reinbek: Rowohlt.

Gergen, K. (1994). *Realities and relationships: Soundings in social construction*. Cambridge, MA: Harvard University Press.
Gill, A. M., & Whedbee, K. (1997). Rhetoric. In T. A. van Dijk (Ed.), *Discourse as structure and process. Discourse studies: A multidisciplinary introduction* (Vol. 1, pp. 157–184). London: Sage.
Glenn, P. J. (1991/1992). Current speaker initiation of two-party shared laughter. *Research on Language and Social Interaction, 25*, 139–162.
Glenn, P. (2003). *Laughter in interaction*. Cambridge: Cambridge University Press.
Glinz, H. (1959). Die Leistung der Sprache für zwei Menschen. In H. Gipper (Ed.), *Sprache, Schlüssel zur Welt: Festschrift für Leo Weisgerber* (pp. 87–105). Düsseldorf: Schwann.
Goffman, E. (1959). *The presentation of self in everyday life*. New York: Anchor Press.
Goffman, E. (1964). The neglected situation. In J. J. Gumperz & D. Hymes (Eds.), The ethnography of communication. *American Anthropologist, 66* [Special Issue], 133–136.
Goffman, E. (1967). *Interaction ritual: Essays on face-to-face behaviour*. Garden City, NY: Doubleday.
Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.
Goldman-Eisler, F. (1961). Hesitation and information in speech. In C. Cherry (Ed.), *Information theory* (pp. 162–174). London: Butterworth.
Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous spoken discourse*. London: Academic Press.
Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language & Speech, 15*, 103–113.
Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 97–121). New York: Irvington.
Goody, J. (1987). *The interface between the written and the oral*. Cambridge: Cambridge University Press.
Grabher, G. M., & Jessner, U. (1996). Introduction. In G. M. Grabher & U. Jessner (Eds.), *Semantics of silences in linguistics and literature* (pp. xi–xvii). Heidelberg: C. Winter.
Grady, K., & Potter, J. (1985). Speaking and clapping: A comparison of Foot and Thatcher's oratory. *Language & Communication, 5*, 173–183.
Graesser, A. C., Hernsbacher, M. A., & Goldman, S. R. (Eds.). (2003). *Handbook of discourse processes*. Mahwah, NJ: Erlbaum.
Graumann, C. F. (1984). Wundt – Mead – Bühler. In C. F. Graumann & T. Herrmann (Eds.), *Karl Bühlers Axiomatik* (pp. 217–247). Frankfurt/Main: Vittorio Klostermann.
Graumann, C. F. (1987). Sprache als Medium sozialer Interaktion. In W. Maiers & M. Markard (Eds.), *Kritische Psychologie als Subjektwissenschaft: Klaus Holzkamp zum 60. Geburtstag* (pp. 57–65). Frankfurt: Campus.
Graumann, C. F. (1989). Perspective setting and taking in verbal interaction. In R. Dietrich & C. F. Graumann (Eds.), *Language processing in social context* (pp. 95–122). Amsterdam: Elsevier.
Graumann, C. F. (2006). Die Verbindung und Wechselwirkung der Individuen im Gemeinschaftsleben. In G. Jüttemann (Ed.), *Wilhelm Wundts anderes Erbe: Ein Missverständnis löst sich auf* (pp. 52–68). Göttingen: Vandenhoeck & Ruprecht.
Graumann, C. F., & Kallmeyer, W. (2002). Perspective and perspectivation in discourse: An introduction. In C. F. Graumann & W. Kallmeyer (Eds.), *Perspective and perspectivation in discourse* (pp. 1–11). Amsterdam: Benjamins.
Greene, G. (1950). *The third man*. Harmondsworth, UK: Penguin.
Greene, G. (1962). *Der dritte Mann*. Wien: Paul Zsolnay.
Greene, G. (1984). *The third man* (rev. ed.). London: Lorrimer.
Gregory, J. C. (1924). *The nature of laughter*. London: Kegan Paul, Trench, Trubner & Co.
Greig, J. Y. T. (1923/1969). *The psychology of laughter and comedy*. London: Allen & Unwin.
Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58). New York: Seminar Press.

Grisham, J. (2000). *The brethren*. New York: Random House.
Grosjean, F., & Collins, M. (1979). Breathing, pausing, and reading. *Phonetica, 36*, 98–114.
Grosjean, F., Grosjean, L., & Lane, H. (1979). The patterns of silence: Performance structures in sentence production. *Cognitive Psychology, 11*, 58–81.
Grossman, R. (June 10, 2007). Shaking language to the core. *Chicago Tribune*, Metro Section, pp. 1, 4.
Gülich, E. (1970). *Makrosyntax der Gliederungssignale im gesprochenen Französisch*. München: Fink.
Gumperz, J. J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.
Gumperz, J. J. (1992). Contextualization revisited. In P. Auer & A. di Luzio (Eds.), *The contextualization of language* (pp. 39–53). Amsterdam: Benjamins.
Gutfleisch-Rieck, J., Klein, W., Speck, A., & Spranz-Fogasy, T. (1989). *Transkriptionsvereinbarungen für den Sonderforschungsbereich 245, Sprechen und Sprachverstehen im sozialen Kontext*. Heidelberg, Germany: Universitäten Heidelberg/Mannheim.
Habermas, J. (1998). *On the pragmatics of communication* (M. Cooke, Trans. & Ed.). Cambridge, MA: MIT Press.
Hagoort, P. (2005). Broca's complex as the unification space for language. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 157–172). Mahwah, NJ: Erlbaum.
Hagtvet, B. E., & Wold, A. H. (2003). On the dialogical basis of meaning: Inquiries into Ragnar Rommetveit's writings on language, thought, and communication. *Mind, Culture, and Activity: An International Journal, 10*, 186–104.
Hahn, E. (1949). An analysis of the delivery of the speech of first grade children. *Quarterly Journal of Speech, 35*, 338–343.
Hall, G. S. & Allin, A. (1897). The psychology of tickling, laughing, and the comic. *American Journal of Psychology, 9*, 1–44.
Hall, K., Bucholtz, M., & Moonwomon, B. (Eds.). (1992). *Locating power: Proceedings of the Second Berkeley Women and Language Conference, April 4 and 5, 1992* (Vols. 1 and 2). Berkeley, CA: Berkeley Women and Language Group, University of California.
Halliday, M. A. K. (1989). *Spoken and written language*. Oxford: Oxford University Press.
Handel, S. (1989). *Listening: An introduction to the perception of auditory events*. Cambridge, MA: The MIT Press.
Hansen, M.-B. M. (1998). *The function of discourse particles: A study with special reference to spoken standard French*. Amsterdam: Benjamins.
Hanssen, F. (2007, July 23). Hände hoch! Künstlerlohn und Akt der Meinungsbildung: Eine kleine Kulturgeschichte des Theaterapplauses. *Der Tagesspiegel*, p. 29.
Harley, T. A. (2001). *The psychology of language: From data to theory* (2nd ed.). East Sussex, UK: Psychology Press.
Harley, T. A. (2008). *The psychology of language: From data to theory* (3rd ed.). East Sussex, UK: Psychology Press.
Harris, R. (1980). *The language-makers*. London: Duckworth.
Harris, R. (1981). *The language myth*. London: Duckworth.
Hartmann, C. O. (1980). *Free verse: An essay on prosody*. Princeton, NJ: Princeton University Press.
Havelock, E. A. (1986). *The muse learns to write*. New Haven, CT: Yale University Press.
Hecker, E. (1873). *Die Physiologie und Psychologie des Lachens und des Komischen*. Berlin: F. Dummler.
Hellinger, M. (1990). *Kontrastive feministische Linguistik: Mechanismen sprachlicher Diskriminierung im Englischen und Deutschen*. Ismaning, Germany: Max Hueber.
Heritage, J., & Greatbatch, D. (1986). Generating applause: A study of rhetoric and response at party political conferences. *American Journal of Sociology, 92*, 110–157.
Herrmann, T. (1985). *Allgemeine Sprachtheorie: Grundlagen und Probleme*. München: Urban & Schwarzenberg.

Herrmann, T. (2005). *Sprache verwenden: Funktionen – Evolution – Prozesse*. Stuttgart: Kohlhammer.
Herrmann, T., & Grabowski, J. (1994). *Sprechen: Psychologie der Sprachproduktion*. Heidelberg: Spektrum Akademischer Verlag.
Hieke, A. E., Kowal, S., & O'Connell, D. C. (1983). The trouble with "articulatory" pauses. *Language & Speech, 26*, 203–214.
Hilton, J. (1953). Calculated spontaneity. In *Oxford book of English talk* (pp. 399–404). Oxford: Clarendon Press.
Holtgraves, T. (2002). *Language as social action: Social psychology and language use*. Mahwah, NJ: Erlbaum.
Hörmann, H. (1967). *Psychologie der Sprache*. Berlin: Springer.
Hörmann, H. (1976). *Meinen und Verstehen: Grundzüge einer psychologischen Semantik*. Frankfurt/Main: Suhrkamp.
Hörmann, H. (1981). *To mean – to understand: Problems of psychological semantics* (D. A. Jankowski, Trans.). Berlin: Springer.
Hörmann, H. (1986). *Meaning and context: An introduction to the psychology of language*. (R. E. Innis, Ed.). New York: Plenum Press.
Honan, P. (1960). Eighteenth and nineteenth century English punctuation theory. *English Studies, 41*, 92–102.
Hopper, R. (1992). *Telephone conversation*. Bloomington, IN: Indiana University Press.
Hosseini, K. (2003). *The kite runner*. New York: Riverhead.
von Humboldt, W. (1841–1852/1969). *Werke in fünf Bänden*. Bd. III: Schriften zur Sprachphilosophie. Stuttgart: Cotta.
Hustvedt, S. (2006). *A plea for eros: Essays*. New York: Picador.
Hutchby, I., & Wooffitt, R. (1998). *Conversation analysis: Principles, practices and applications*. Cambridge, UK: Polity Press.
Ideforss, H. (1928). *De primäre Interjektionera i Nysvenskan, I, Primära Impulsioner och Imperationer*. Lund: Gleerup.
Innis, H. A. (1951). *The bias of communication*. Toronto: University of Toronto Press.
Innis, R. E. (1986). Introduction. In H. Hörmann, *Meaning and context: An introduction to the psychology of language* (pp. 3–29). New York: Plenum Press.
Jaffe, J., & Feldstein, S. (1970). *Rhythms of dialogue*. New York: Academic Press.
Jahandarie, K. (1999). *Spoken and written discourse: A multi-disciplinary perspective*. Stamford, CT: Ablex.
James, D., & Clarke, F. (1993). Women, men, and interruptions: A critical review. In D. Tannen (Ed.), *Gender and conversational interaction* (pp. 231–280). New York: Oxford University Press.
James, W. (1891/1981). *The principles of psychology*. Cambridge, MA: Harvard University Press.
Jamieson, K. H. (1988). *Eloquence in an electronic age: The transformation of political speechmaking*. New York: Oxford University Press.
Jarrell, R. (1969). *The complete poems*. New York: Farrar, Straus & Giroux.
Jaworski, A. (1993). *The power of silence: Social and pragmatic perspectives*. Newbury Park: Sage.
Jefferson, G. (1973). A case of precision timing in ordinary conversation: Overlapped tag-positioned address terms in closing sequences. *Semiotica, 9*, 47–96.
Jefferson, G. (1979). A technique for inviting laughter and its subsequent acceptance declination. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 79–96). New York: Irvington.
Jefferson, G. (1984). *Notes on the sequential organization of laughter in conversation: Onset sensitivity in invitations to laugh*. Paper presented at the American Anthropological Association Convention, Mexico City.
Jefferson, G. (1989). Preliminary notes on a possible metric which provides for a 'standard maximum' silence of approximately one second in conversation. In D. Roger & P. Bull

(Eds.), *Conversation: An interdisciplinary perspective* (pp. 166–196). Clevedon, UK: Multilingual Matters.
Jefferson, G., Sacks, H., & Schegloff, E. (1984). On laughter in pursuit of intimacy. *Working Paper Università di Urbino*, no. 135, serie C.
Jenniches, K. M. (1969). Der Beifall als Kommunikationsmuster im Theater. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 21*, 569–584.
Jespersen, O. (1922). *Language: Its nature, development and origin*. London: Allen & Unwin.
Johnson, M. L. (1986). *Applying the breaks: A review of the poetic line*. Unpublished manuscript, University of Kansas, Department of English, Lawrence, KS.
Joyce, J. (1922/1960). *Ulysses*. London: The Bodley Head.
Kainz, F. (1969). *Psychologie der Sprache: Fünfter Band, II. Teil. Psychologie der Einzelsprachen II*. Stuttgart: Ferdinand Enke.
Karčevski, S. (1941). Introduction à l'étude de l'interjection. *Cahiers F. de Saussure, 1*, 57–75.
Käsermann, M.-L., & Foppa, K. (2003). Sprachproduktion im Gespräch. In T. Herrmann & J. Grabowski (Eds.), *Sprachproduktion* (pp. 767–799). Göttingen: Hogrefe.
Kasl, S. V., & Mahl, G. F. (1965). The relationship of disturbances and hesitations in spontaneous speech to anxiety. *Journal of Personality and Social Psychology, 1*, 425–433.
Keller, C., & Vornbäumen, A. (2007, April 8). Karl, lass mal die Heiligkeit weg. *Der Tagesspiegel*, p. S 1.
Keppler, A. (1987). Der Verlauf von Klatschgesprächen. *Zeitschrift für Soziologie, 16*, 288–302.
Kern, F., & Selting, M. (2006). Einheitenkonstruktion im Türkendeutschen: Grammatische und prosodische Aspekte. *Zeitschrift für Sprachwissenschaft, 25*, 239–272.
Kerstan, K., & Honsel, G. (1992). Die Jenniger-Rede: Zum Zitierverhalten deutscher Politiker und Journalisten. *Zeitschrift für Semiotik, 3*, 247–252.
Kidd, S. M. (2002). *The secret life of bees*. New York: Penguin.
Kitzinger, C. (1998). Inaccuracies in quoting from data transcripts: *Or* inaccuracy in quotations from data transcripts. *Discourse & Society, 9*, 136–143.
Kleemann, F. (1980). Der Gebrauch der Interjektionen bei Wilhelm Busch. *Sprachpflege, 20*, 8–11.
von Kleist, H. (1806/undated). Über die allmähliche Verfertigung der Gedanken beim Reden, an R[ühle] v[on] L[ilienstern]. In K. F. Reinking (Ed.), *Heinrich von Kleist, Sämtliche Werke* (pp. 975–980). Wiesbaden: R. Löwit.
Klewitz, G., & Couper-Kuhlen, E. (1999). Quote-unquote? The role of prosody in the contextualization of reported speech sequences. *Pragmatics, 9*, 459–485.
Kluge, F. (1902). Notschreie. *Zeitschrift für deutsche Wortforschung, 2*, 47–49.
Knobloch, C. (1984). *Sprachpsychologie*. Tübingen: Niemeyer.
Knobloch, C. (2003). Geschichte der Psycholinguistik. In G. Rickheit, T. Herrmann, & W. Deutsch (Eds.), *Psycholinguistik/Psycholinguistics: Ein internationales Handbuch/An international handbook* (pp. 15–33). Berlin: de Gruyter.
Koch, P., & Oesterreicher, W. (1994). Schriftlichkeit und Sprache. In H. Günther & O. Ludwig (Eds.), *Schrift und Schriftlichkeit. Writing and its use. Ein interdisziplinäres Handbuch internationaler Forschung. An interdisciplinary handbook of international research* (pp. 587–604). Berlin: de Gruyter.
Kohler, K. J. (1995). *Einführung in die Phonetik des Deutschen* (2 nd ed.). Berlin: Erich Schmidt.
Kohonen, S. (2004). Turn-taking in conversation: Overlaps and interruptions in intercultural talk. *Association for French Language Studies, 10*, 15–31.
Korda, A., Selznik, D. O. (Producers), & Reed, C. (Director). (1949). *The third man* [Motion picture]. (Available from Home Vision Cinema, 4411 N. Ravenswood Avenue, Chicago, IL 60640-5802).
Kotthoff, H. (1996). Vom Lächeln der Mona Lisa zum Lachen der Hyänen. In H. Kotthoff (Ed.), *Das Gelächter der Geschlechter: Humor und Macht in Gesprächen von Frauen und Männern* (2 nd ed., pp. 121–163). Konstanz: Universitätsverlag Konstanz.

Kowal, S. (1989). Verzögerungen in Nacherzählungen und Simultanerzählungen eines Films. In S. Kowal & R. Posner (Eds.), *Zeitliche und inhaltliche Aspekte der Textproduktion* (pp. 108–124). Arbeitspapiere zur Linguistik, 20, Berlin: Technische Universität Berlin, Institut für Linguistik

Kowal, S. (1991). *Über die zeitliche Organisation des Sprechens in der Öffentlichkeit: Pausen, Sprechtempo und Verzögerungen in Interviews und Reden von Politikern*. Bern: Huber.

Kowal, S. (1998). Review of the book *Einführung in die feministische Sprachwissenschaft*. *Archiv für das Studium der neueren Sprachen und Literaturen, 235*, 362–364.

Kowal, S., Barth, H.-C., Egemann, H., Galusic, G., Kögel, C., Lippold, N., Pfeil, A., & O'Connell, D. C. (1998). Unterbrechungen in Medieninterviews: Geschlechtstypisches Gesprächsverhalten? *Germanistische Linguistik, 139–140*, 279–299.

Kowal, S., Bassett, M., & O'Connell, D. C. (1985). The spontaneity of media interviews. *Journal of Psycholinguistic Research, 14*, 1–18.

Kowal, S., & O' Connell, D. C. (1987). Some temporal aspects of stories told while or after watching a film. *Bulletin of thePsychonomic Society*, *25*, 364–366.

Kowal, S., & O'Connell, D. C. (1995). Notation und Transkription in der Gesprächsforschung. *KODIKAS/CODE, 18*, 113–138.

Kowal, S., & O'Connell, D. C. (1997). Theoretical ideals and their violation: Princess Diana and Martin Bashir in the BBC interview. *Pragmatics*, 7, 309–323.

Kowal, S., & O'Connell, D. C. (2000). Psycholinguistische Aspekte der Transkription: Zur Notation von Pausen in Gesprächstranskripten. *Linguistische Berichte, 183*, 353–378.

Kowal, S., & O'Connell, D. C. (2003a). Datenerhebung und Transkription. In G. Rickheit, T. Herrmann, & W. Deutsch (Eds.), *Psycholinguistik/Psycholinguistics: Ein internationales Handbuch/An international handbook* (pp. 92–106). Berlin: de Gruyter.

Kowal, S., & O'Connell, D. C. (2003b). Die Transkription mündlicher Äußerungen. In T. Herrmann & J. Grabowski (Eds.), *Sprachproduktion* (Vol. 1, pp. 101–120). Göttingen: Hogrefe.

Kowal, S., & O'Connell, D. C. (Eds.). (2004a). Interjektionen [Special issue]. *Zeitschrift für Semiotik, 26* (1–2), 3–127.

Kowal, S., & O'Connell, D. C. (2004b). Interjektionen im Gespräch. *Zeitschrift für Semiotik, 26*, 85–100.

Kowal, S., O'Connell, D. C., Forbush, K., Higgins, M., Clarke, L., & D'Anna, K. (1997). Interplay of literacy and orality in inaugural rhetoric. *Journal of Psycholinguistic Research, 26*, 1–31.

Kowal, S., O'Connell, D. C., & Sabin, E. J. (1975). Development of temporal patterning and vocal hesitations in spontaneous narratives. *Journal of Psycholinguistic Research, 4*, 195–207.

Kucharczik, K. (1996). A review of A. Redder & K. Ehlich (Eds.), *Gesprochene Sprache: Transkripte und Tondokumente*. Tübingen: Niemeyer. *Leuvense Bijdragen*, *85*, 435–442.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago, IL: The University of Chicago Press.

Kuo, S. (2001). Generating applause and laughter: A study of rhetoric and response in the 1998 Taipei mayoral debates. *Concentric: Studies in English Literature and Linguistics, 27*, 189–215.

Kurzon, D. (1996). The White House speeches: Semantic and paralinguistic strategies for eliciting applause. *Text, 16*, 199–224.

Kurzon, D. (1997). *Discourse of silence*. Amsterdam: Benjamins.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Lakoff, R. T. (2001). The rhetoric of the extraordinary moment: The concession and acceptance speeches of Al Gore and George W. Bush in the 2000 presidential election. *Pragmatics, 11*, 309–327.

Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (Eds.). (2007). *Handbook of latent semantic analysis*. Mahwah, NJ: Erlbaum.
Langacker, R. (1987). *Foundations of cognitive grammar, Vol. 1: Theoretical prerequisites*. Stanford: Stanford University Press.
van Langenhove, L. & Harré, R. (1993). Positioning and autobiography: Telling your life. In N. Coupland & J. Nussbaum (Eds.), *Discourse and life-span development* (Vol. 1, pp. 81–99). Newbury Park: Sage.
Langenmayr, A. (1997). *Sprachpsychologie: Ein Lehrbuch*. Göttingen: Hogrefe.
Lazarus, M. (1879/1986). *Über Gespräche*. Berlin: Henssel.
Lehiste, I. (1984, June). *Rhythm in spoken sentences and read poetry*. Paper presented at the International Phonology Meeting, Eisenstadt, Austria.
Lenoski, D. S. (2001). The ubiquitous Catholic novel. *Perspective, 3*, 4–15.
Lenz, F. (1988). Turn-taking in technical meetings: Zu einem Aspekt der Interaktion in mündlicher Fachkommunikation. *Linguistische Berichte, 114*, 134–162.
Levelt, W. J. M. (1981). The speaker's linearization problem. *Philosophical Transactions Royal Society London B295*, 305–315.
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.
Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
Lévi-Strauss, C. (1964–1971). *Mythologiques* (4 vols.). Paris: Plon.
Lindsay, J. S. (1988). *I can't believe my ears: A study of errors in transcription*. Unpublished manuscript.
Lindsay, J., & O'Connell, D. C. (1995). How do transcribers deal with audio recordings of spoken discourse? *Journal of Psycholinguistic Research, 24*, 101–115.
Linell, P. (1982). *The written language bias in linguistics*. Linköping, Sweden: University of Linköping.
Linell, P. (1990). The power of dialogue dynamics. In I. Marková & K. Foppa (Eds.), *The dynamics of dialogue* (pp. 147–177). New York: Springer.
Linell, P. (1995). Troubles with mutuality: Towards a dialogical theory of misunderstanding and miscommunication. In I. Marková, C. F. Graumann, & K. Foppa (Eds.), *Mutualities in dialogue* (pp. 176–213). Cambridge: Cambridge University Press.
Linell, P. (1998). *Approaching dialogue: Talk, interaction and contexts in dialogical perspectives*. Amsterdam: Benjamins.
Linell, P. (2005). *The written language bias in linguistics: Its nature, origins and transformations*. London: Routledge.
Linell, P., & Rommetveit, R. (1998). The many forms and facets of morality in dialogue: Epilogue for the special issue. *Research on Language and Social Interaction, 31*, 465–473.
Livingstone, S., & Lunt, P. (1994). *Talk on television: Audience participation and public debate*. London: Routledge.
Lord, A. B. (1991). *Epic singers and oral tradition*. Ithaca, NY: Cornell University Press.
Lounsbury, F. G. (1954). Transitional probability, linguistic structure, and systems of habit-family hierarchies. *Journal of Abnormal and Social Psychology, 49,* Supplement, edited by C. E. Osgood & T. A. Sebeok, 93–101.
Luckmann, T. (1990). Social communication, dialogue and conversation. In I. Markovà & K. Foppa (Eds.), *The dynamics of dialogue* (pp. 45–61). New York: Springer.
Ludlum, R. (2000). *The Prometheus deception*. New York: St. Martin's Press.
Lueken, G.-L. (1996). Dialog und Verständlichkeit. In E. Hasselberg, L. Martienssen, & F. Radtke (Eds.), *Der Dialogbegriff am Ende des 20. Jahrhunderts* (pp. 84–91). Berlin: Hegel Institut.
Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word, 15*, 19–44.
MacWhinney, B. (1995). *The CHILDES project: Tools for analyzing talk* (2 nd ed.). Hillsdale, NJ: Erlbaum.

Mahl, G. F. (1956). Normal disturbances in spontaneous speech: General quantitative aspects. *American Psychologist, 11,* 390. (abstract) cf. ms.
Mahl, G. F. (1958). On the use of 'ah' in spontaneous speech: Quantitative, developmental, characterological, situational, and linguistic aspects. *American Psychologist, 13*, 349 (abstract) cf. ms.
Mahl, G. F. (1959). Measuring the patient's anxiety during interviews from 'expressive' aspects of his speech. *Transactions of the New York Academy of Sciences, 21*, 249–257.
Malinowski, B. (1923). The problem of meaning in primitive languages. In C. K. Ogden & I. A. Richards (Eds.), *The meaning of meaning* (pp. 296–336). New York: Harcourt, Brace & World.
Martin, R. A. (2006). *The psychology of humor: An integrative approach*. San Diego, CA: Elsevier Academic Press.
Maynard, D. W., Houtkoop-Steenstra, H., Schaeffer, N. C., & van der Zouwen, J. (Eds.). (2002). *Standardization and tacit knowledge: Interaction and practice in the survey interview*. New York: Wiley.
McCarthy, D. (1954). Language development in children. In L. Carmichael (Ed.), *Manual of child psychology* (pp. 492–630). New York: Wiley.
McClay, W. M. (2007, February 2). Twilight of sociology. *The Wallstreet Journal*, p. W 13.
McIlvenny, P. (1996). Heckling in Hyde Park: Verbal audience participation in popular public discourse. *Language in Society, 25*, 27–60.
McKendrick, J. (2003). Interpretations of gender differences in conversational interruptions. <http://www.teilhard.com/writings/nve%20-%20interruption.pdf>
McLuhan, M. (1962). *The Gutenberg galaxy: The making of typographic man*. Toronto: University of Toronto Press.
McQuail, D. (1997). *Audience analysis*. Thousand Oaks, CA: Sage.
McTear, M. (1985). *Children's conversation*. Oxford: Basil Blackwell.
Mead, G. H. (1980). *Gesammelte Aufsätze,* Vol 1. Frankfurt/Main: Suhrkamp.
Meise, K. (1996). *Une forte absence: Schweigen in alltagsweltlicher und literarischer Kommunikation*. Tübingen: Gunter Narr.
Menand, L. (2004, November 8). Ask not, tell not: Anatomy of an inaugural. *New Yorker, 80*, 110–119.
Merleau-Ponty, M. (1955). *Les aventures de la dialectique*. Paris: Gallimard.
Merriam-Webster's collegiate dictionary (11th ed., 2003). Springfield, MA: Merriam-Webster.
Miller, J., & Weinert, R. (1998). *Spontaneous spoken language: Syntax and discourse*. Oxford: Clarendon Press.
Mills, S. (2003). *Gender and politeness*. Cambridge: Cambridge University Press.
Montgomery, M. (2000). Televised talk: Face work, politeness and laughter in *The Mrs Merton Show*. In M. Coulthard, J. Cotterill, & F. Rock (Eds.), *Dialogue analysis VII: Working with dialogue* (pp. 121–135). Tübingen: Niemeyer.
Müller, C., & Posner, R. (Eds.). (2004). *The semantics and pragmatics of everyday gestures*. Berlin: Weidler.
Nemerov, H. (1977). *The collected poems of Howard Nemerov*. Chicago: University of Chicago Press.
Newman, S., & Mather, V. G. (1938). Analysis of spoken language of patients with affective disorders. *American Journal of Psychiatry, 94*, 913–942.
Noonan, P. (1990). *What I saw at the revolution: A political life in the Reagan era*. New York: Ivy Books.
Norman, D. A., & Rumelhart, D. E. (1975). *Explorations in cognition*. San Francisco: Freeman.
Norrick, N. R. (2001). Discourse and semantics. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 76–99). Malden, MA: Blackwell.
Norwine, A. C., & Murphy, O. J. (1938). Characteristic time intervals in telephone conversations. *Bell Systems Technical Journal, 17*, 281–291.

Nübling, D. (2001). Von *oh mein Jesus!* zu *oje!* Der Interjektionalisierungspfad von der sekundären zur primären Interjektion. *Deutsche Sprache, 1*, 20–45.
Nübling, D. (2004). Die prototypische Interjektion: Ein Definitionsvorschlag. *Zeitschrift für Semiotik, 26*, 11–45.
Nübling, D. (2005). Die nicht flektierbaren Wortarten. In *Duden: Die Grammatik* (pp. 573–640). Mannheim: Dudenverlag.
Ochs, E. (1979). Transcription as theory. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 43–72). New York: Academic Press.
O'Connell, D. C. (1980). Toward an empirical rhetoric: Some comparisons of expressiveness in poetry readings by authors, English professors, and drama professors. *Archiv für Psychologie, 133*, 117–128.
O'Connell, D. C. (1982). Performance characteristics of poetry: Some cross-linguistic comparisons. *Psychological Research, 44*, 381–392.
O'Connell, D. C. (1988). *Critical essays on language use and psychology*. New York: Springer.
O'Connell, D. C. (1992). Some intentions regarding speaking. A review of *Speaking: From intention to articulation,* by W. J. M. Levelt. *Journal of Psycholinguistic Research, 21*, 59–65.
O'Connell, D. C., & Kowal, S. (1972). Cross-linguistic pause and rate phenomena in adults and adolescents. *Journal of Psycholinguistic Research, 1*, 155–164.
O'Connell, D. C., & Kowal, S. (1983). Pausology. In W. A. Sedelow & S. Y. Sedelow (Eds.), *Computers in language research* (Vol. 2, pp. 221–301). Berlin: de Gruyter.
O'Connell, D. C., & Kowal, S. (1986). Use of punctuation for pausing: Oral readings by German radio homilists. *Psychological Research, 48*, 93–98.
O'Connell, D. C., & Kowal, S. (1990a). A note on time, timing, and transcriptions thereof. *Georgetown Journal of Languages and Linguistics, 1*, 203–208.
O'Connell, D. C., & Kowal, S. (1990b). Some sources of error in the transcription of real time in spoken discourse. *Georgetown Journal of Languages and Linguistics, 1*, 453–466.
O'Connell, D. C., & Kowal, S. (1994). The transcriber as language user. In G. Bartelt (Ed.), *The dynamics of language processes: Essays in honor of Hans W. Dechert* (pp. 119–142). Tübingen: Gunter Narr.
O'Connell, D. C., & Kowal, S. (1995a). Basic principles of transcription. In J. A. Smith, R. Harré, & L. van Langenhove (Eds.), *Rethinking methods in psychology* (pp. 93–105). London: Sage.
O'Connell, D. C., & Kowal, S. (1995b). Transcription systems for spoken discourse. In J. Verschueren, J.-O. Östman, & J. Blommaert (Eds.), *Handbook of pragmatics* (pp. 646–656). Amsterdam: Benjamins.
O'Connell, D. C., & Kowal, S. (1998). Orality and literacy in public discourse: An interview of Hannah Arendt. *Journal of Pragmatics, 30*, 543–564.
O'Connell, D. C., & Kowal, S. (1999). Transcription and the issue of standardization. *Journal of Psycholinguistic Research, 28*, 103–120.
O'Connell, D. C., & Kowal, S. (2000). Are transcripts reproducible? *Pragmatics, 10*, 247–269.
O'Connell, D. C., & Kowal, S. (2003). Psycholinguistics: A half century of monologism. *American Journal of Psychology, 116*, 191–212.
O'Connell, D. C., & Kowal, S. (2004a). Hillary Clinton's laughter in media interviews. *Pragmatics, 14*, 463–478.
O'Connell, D. C., & Kowal, S. (2004b). The history of research on the filled pause as evidence of *The written language bias in linguistics* (Linell, 1982). *Journal of Psycholinguistic Research, 33*, 459–474.
O'Connell, D. C., & Kowal, S. (2005a). Laughter in Bill Clinton's *My life* (2004) interviews. *Pragmatics*, *15*, 275–299.
O'Connell, D. C., & Kowal, S. (2005b). *Uh* and *um* revisited: Are they interjections for signalling delay? *Journal of Psycholinguistic Research, 34*, 555–576.

O'Connell, D. C., & Kowal, S. (2005c). Where do interjections come from? A psycholinguistic analysis of Shaw's *Pygmalion*. *Journal of Psycholinguistic Research, 34*, 497–514.
O'Connell, D. C., & Kowal, S. (2006a). Laughter in the film *The third man*. *Pragmatics, 16*, 305–327.
O'Connell, D. C., & Kowal, S. (2006b). The research status of Clayman and Heritage's (2002) *The news interview*. *Journal of Psycholinguistic Research, 35*, 147–165.
O'Connell, D. C., & Kowal, S. (2006c). Verbal integrity. *Journal of Social Distress and the Homeless, 15*, 207–227.
O'Connell, D. C., Kowal, S., & Ageneau, C. (2005). Interjections in interviews. *Journal of Psycholinguistic Research, 34*, 153–171.
O'Connell, D. C., Kowal, S., & Dill, E. J., III. (2004). Dialogicality in TV news interviews. *Journal of Pragmatics, 36*, 185–205.
O'Connell, D. C., Kowal, S., & Hörmann, H. (1969). Semantic determinants of pauses. *Psychologische Forschung, 33*, 50–67.
O'Connell, D. C., Kowal, S., & Hörmann, H. (1970). Semantic determinants of pauses. In G. B. Flores d'Arcais & W. T. M. Levelt (Eds.), *Advances in psycholinguistics* (pp. 218–219). Amsterdam: North-Holland.
O'Connell, D. C., Kowal, S., & Kaltenbacher, E. (1990). Turn-taking: A critical analysis of the research tradition. *Journal of Psycholinguistic Research, 19*, 345–373.
O'Connell, D. C., Kowal, S., & King, S. P. (2007). Interjections in literary readings and artistic performance. *Pragmatics, 17*, 1–22.
Oehl, W. (1933). *Das Lallwort in der Sprachschöpfung*. Freiburg (Breisgau): St. Paulus-Druckerei.
Ogden, C. K., & Richards, I. A. (1923). *The meaning of meaning*. New York: Harcourt, Brace & World.
Okamoto, D. G., Rashotte, L. S., & Smith-Lovin, L. (2002). Measuring interruption: Syntactic and contextual methods of coding conversation. *Social Psychology Quarterly, 65*, 38–55.
Oliveira, M. (2002). The role of pause occurrence and pause duration in the signalling of narrative structure. Retrieved from http://www.iltec.pt/pdf/wpapers/2002-moliveira-pauses.pdf
Olson, D. R. (1994). *The world on paper: The conceptual and cognitive implications of writing and reading*. Cambridge: Cambridge University Press.
Ong, W. J. (1967). *The presence of the word: Some prolegomena for cultural and religious history*. Minneapolis: University of Minnesota Press.
Ong, W. J. (1982). *Orality and literacy: The technologizing of the word*. London: Methuen.
Oreström, B. (1983). *Turn-taking in English conversation*. Lund: Liber.
Osgood, C. E., & Sebeok, T. A. (Eds.). (1954). Psycholinguistics: A survey of theory and research problems. *Journal of Abnormal and Social Psychology, 49*, Supplement. *Oxford English Dictionary*. (1989, 2nd ed.). Oxford: Oxford University Press.
Page, N. (1988). *Speech in the English novel*. Atlantic Highlands, NJ: Humanities Press International.
Pamuk, O. (1997). *The new life* (Güneli Gün, Trans.). London: Faber and Faber.
Panksepp, J . (2005). Beyond a joke: From animal laughter to human joy? *Science, 308*, 62–63.
Panksepp, J., & Burgdorf, J. (2003). "Laughing" rats and the evolutionary antecedents of human joy? *Physiology & Behavior, 79*, 533–547.
Papaioannou, E. (2003). *Applause for inaugurals and state of the union addresses: A study of applause as a communicative audience response from a social-dialogical perspective*. Unpublished doctoral dissertation, Georgetown University, Washington, DC.
Parkes, M. B. (1993). *Pause and effect: An introduction to the history of punctuation in the West*. Berkeley, CA: University of California Press.
Parry, M. (1971). *The making of Homeric verse*. Oxford: Clarendon Press.
Partington, A. (2006). *The linguistics of laughter: A corpus-assisted study of laughter-talk*. London: Routledge.
Pascal, G. (Producer), Asquith, A., & Howard, L. (Directors). (1938). Pygmalion [Motion picture]. (Available as VHS or DVD from http://www.amazon.com)

Pearce, W. B. (2002). Toward a national conversation about public issues. In W. F. Eadie & P. E. Nelson (Eds.), *The changing conversation in America: Lectures from the Smithsonian* (pp. 13–38). Thousand Oaks, CA: Sage.

Perfetti, C. C., & Papi, M. B. (1985). Towards a cognitive typology of pause phenomena. *Communication & Cognition, 4*, 339–351.

Philips, S. U. (1976). Some sources of cultural variability in the regulation of talk. *Language in Society, 5*, 81–95.

Pickering, M. G., & Garrod, S. (2005). Establishing and using routines during dialogue: Implications for psychology and linguistics. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 85–101). Mahwah, NJ: Erlbaum.

Piddington, R. (1933). *The psychology of laughter: A study in social adaptation*. London: Figurehead.

Pike, K. L. (1943). *Phonetics: A critical analysis of phonetic theory and a technic for the practical description of sounds*. Ann Arbor, MI: University of Michigan Press.

Poeppel, D., & Embick, D. (2005). Defining the relation between linguistics and neuroscience. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 103–118). Mahwah, NJ: Erlbaum.

Pompino-Marschall, B., Kowal, S., & O'Connell, D. C. (2007, August 4). *Some phonetic notes on emotional expressions: Laughter, interjections and weeping*. Paper presented at an interdisciplinary workshop on The Phonetics of Laughter, Saarbrücken, Germany.

Posner, R. (Ed). (1990). *Warnungen an die ferne Zukunft: Atommüll als Kommunikationsproblem*. München: Raben Verlag.

Potter, J. (1996). *Representing reality: Discourse, rhetoric and social construction*. London: Sage.

Power, R. J. D., & Dal Martello, M. F. (1986). Some criticisms of Sacks, Schegloff, and Jefferson on turn taking. *Semiotica, 58*, 29–40.

Provine, R. R. (1993). Laughter punctuates speech: Linguistic, social and gender contexts of laughter. *Ethology, 95*, 291–298.

Provine, R. R. (2000). *Laughter: A scientific investigation*. New York: Penguin.

Provine, R. R. (2000/2001). The laughing species. *Natural History, 109*, 72–77.

Provine, R. R. (2004). Laughing, tickling, and the evolution of speech and self. *Current Directions in Psychological Science, 13*, 215–218.

Psathas, G., & Anderson, T. (1990). The 'practices' of transcription in conversation analysis. *Semiotica, 78*, 75–99.

Quasthoff, U. M. (1995). Oral communication: Theoretical differentiation and integration of an empirical field. In U. M. Quasthoff (Ed.), *Aspects of oral communication* (pp. 3–16). Berlin: de Gruyter.

Quintilian, M. F. (ca. 95/1958). *The institution oratoria* (Vol. 4, H. E. Burler, Trans.). Cambridge, MA: Harvard University Press.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.

Raible, W. (2002). Literacy and orality. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of social and behavioral sciences* (pp. 8967–8971). Amsterdam: Elsevier.

Raschka, C. (1993). *YO! YES?* New York: Orchard Books.

Raschka, C. (1997). *HEY! JA?* München: Hanser.

Rasoloson, J. N. (1994). *Interjektionen im Kontrast: Am Beispiel der deutschen, madagassischen, englischen und französischen Sprache*. Frankfurt/Main: Peter Lang.

Reagan, R. (1985). Second inaugural address. *Congressional Record, 131*, 372–374.

Reagan, R. (1988). Inaugural address, January 21, 1985. *Public Papers of the Presidents of the United States* (Book I – January 1 to June 28, 1985, pp. 55–58). Washington, DC: United States Government Printing Office.

Redder, A., & Ehlich, K. (Eds.). (1994). *Gesprochene Sprache: Transkripte und Tondokumente*. Tübingen: Niemeyer.

Reddy, M. (1979). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought* (pp. 284–324). Cambridge: Cambridge University Press.

Reisigl, M. (1999). *Sekundäre Interjektionen*. Frankfurt/Main: Peter Lang.

Repp, B. H. (1986). *The sound of two hands clapping: An exploratory study*. (Status Report on Speech Research SR-86/87). New Haven, CT: Haskins Laboratories.

Reyna, V. F. (1993). Theory and reality in psycholinguistics. *Psychological Science, 4*, 15 and 19–23.

Rhodes, R. A. (1992). Interjections. In W. Bright (Ed.), *International encyclopedia of linguistics* (Vol. 2, p. 222). New York: Oxford University Press.

Rich, F. (2007, September 30). Is Hillary Clinton the New Old Al Gore? *New York Times*, **OP-ED,** p. WK 12.

Rickheit, G., Sichelschmidt, L., & Strohner, H. (2002). *Psycholinguistik*. Tübingen: Stauffenburg.

Ripper, H. (1998). *Der Große Kommunikator: Die Medienstrategie Ronald Reagans im Kontext der US-Präsidenten*. Freiburg (Breisgau): Alber.

Rommetveit, R. (1974). *On message structure: A framework for the study of language and communication*. London: Wiley.

Rommetveit, R. (1979). Deep structure of sentences versus message structure: Some critical remarks on current paradigms, and suggestions for an alternative approach. In R. Rommetveit & R. M. Blakar (Eds.), *Studies of language, thought and verbal communication* (pp. 17–34). London: Academic Press. (Originally published in *Norwegian Journal of Linguistics,* 1972, *26*, 3–22).

Rommetveit, R. (1990). On axiomatic features of a dialogical approach to language and mind. In I. Marková & K. Foppa (Eds.), *The dynamics of dialogue* (pp. 83–104). New York: Springer.

Rommetveit, R. (1998a). On divergent perspectives and controversial issues in studies of language and mind. In M. Janse and A. Verlinden (Eds.), *Productivity and creativity* (pp. 179–189). Berlin: Mouton de Gruyter.

Rommetveit, R. (1998b). Intersubjective attunement and linguistically mediated meaning in discourse. In S. Bråten (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 354–371). Cambridge: Cambridge University Press.

Rommetveit, R. (1999). *Ordinary language, common sense, psychologic, and dialogical psychology: An invitation to dialogues about a dialogically founded psychology of the second person*. Unpublished manuscript, University of Oslo: Department of Psychology.

Rose, R. L. (1998). *The communicative value of filled pauses in spontaneous speech*. Unpublished doctoral dissertation, University of Birmingham, UK.

Rubenstein, H., & Aborn, M. (1960). Psycholinguistics. *Annual Review of Psychology, 11*, 291–322.

Sabin, E. J., & O'Connell, D. C. (2006, February 15). The microarchitecture of modern psycholinguistics: [Review of the book *Twenty-first century psycholinguistics: Four cornerstones*]. *PsycCRITIQUES: Contemporary Psychology – APA Review of Books, 51,* Article 14. Retrieved February 16, 2006 from the PsycCRITIQUES database.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*, 696–735.

Sahar, A. D., Brenninkmeyer, S. M., & O'Connell, D. C. (1997). Maya Angelou's inaugural poem. *Journal of Psycholinguistic Research, 26*, 449–463.

Samel, I. (1995). *Einführung in die feministische Sprachwissenschaft*. Berlin: Erich Schmidt.

Sapir, E. (1921). *Language: An introduction to the study of speech*. New York: Harcourt, Brace & World.

Saville-Troike, M. (1985). The place of silence in an integrated theory of communication. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 3–18). Norwood, NJ: Ablex.

Schachter, S., Christenfeld, N., Ravina, B., & Bilous, F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology, 20*, 362–367.
Schäfer, A. (2007, February 26). Roman eines Vaterlosen. *Der Tagesspiegel*, p. 24.
Schegloff, E. A. (1979). Identification and recognition in telephone conversation openings. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 23–78). New York: Irvington.
Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis 1*. Cambridge: Cambridge University Press.
Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica, 8*, 289–327.
Schenkein, J. N., (1972). Towards an analysis of natural conversation and the sense of *heheh*. *Semiotica*, 6, 344-377.
Schenkein, J. (1978). Explanation of transcript notation. In J. Schenkein (Ed.). *Studies in the organization of conversational interaction* (pp. xi–xvi). New York: Academic Press.
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.). (2003). *Handbook of discourse analysis*. Malden, MA: Blackwell.
Schmidt, C. (2004). Interjektionen im Zweitsprachenerwerb. *Zeitschrift für Semiotik, 26*, 101–112.
Schneider, W. (1959). *Stilistische deutsche Grammatik* (pp. 355–363). Basel: Herder.
Schudson, M. (1978). The ideal of conversation in the mass media. *Communication Research, 5*, 320–329.
Schwyzer, E. (1924). Griechische Interjektionen und griechische Buchstabennamen auf –a. *Zeitschrift für vergleichende Sprachforschung, 58*, 170–205.
Scollon, R., & Scollon, S. (1995). Somatic communication: How useful is 'orality' for the characterization of speech events and cultures? In U. M. Quasthoff (Ed.), *Aspects of oral communication* (pp. 19–29). Berlin: de Gruyter.
Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.
Seibt, G. (2002). Der Einspruch des Körpers: Philosophien des Lachens von Platon bis Plessner – und zurück. *Merkur, 56* (9/10), 751–762.
Selting, M. (2001). Probleme der Transkription verbalen und paraverbalen/prosodischen Verhaltens. In K. Brinker, G. Antos, W. Heinemann, & S. F. Sager (Eds.), *Text- und Gesprächslinguistik/Linguistics of text and conversation: Ein internationales Handbuch zeitgenössischer Forschung/An international handbook of contemporary research* (pp. 1059–1068). Berlin: de Gruyter.
Selting, M., Auer, P., Barden, B., Bergmann, J., Couper-Kuhlen, E., Günthner, S., Meier, C., Quasthoff, U., Schlobinski, P., & Uhmann, S. (1998). Gesprächsanalytisches Transkriptionssystem (GAT). *Linguistische Berichte, 173*, 91–122.
Sennet, R. (1977). *The fall of public man*. New York: Alfred A. Knopf.
Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.
Shaw, G. B. (1916/1969). *Pygmalion. A romance in five acts and my fair lady: A musical play in two acts*. New York: Signet.
Sheridan, T. (1787). *Courses of lecture on elocution*. London: J. Dodsley, Pall Mall.
Shotter, J. (1993). *Conversational realities: Constructing life through language*. London: Sage.
Siegman, A. W. (1979). Cognition and hesitation in speech. In A. W. Siegman & S. Feldstein (Eds.), *Of speech and time: Temporal speech patterns in interpersonal contexts* (pp. 151–178). Hillsdale, NJ: Erlbaum.
Sinclair, J. (1995). From theory to practice. In G. Leech, G. Myers, & J. Thomas (Eds.), *Spoken English on computer: Transcription, mark-up, and application* (pp. 99–109). Harlow, England: Longman.

Skinner, B. F. (1948). *Walden two*. Toronto: Macmillan.
Smith, F. (1982). *Writing and the writer*. New York: Holt, Rinehart and Winston.
Spinos, A.-M. R., O'Connell, D. C., & Kowal, S. (2002). An empirical investigation of pause notation. *Pragmatics, 12*, 1–9.
Steiner, G. (1977). A happy few leap before looking. In M. A. Sperber (Ed.), *Arthur Koestler: A collection of critical essays* (pp. 145–150). Englewood Cliffs, NJ: Prentice-Hall.
Stivers, T., Enfield, N. J., & Levinson, S. C. (2007). Person reference in interaction. In N. J. Enfield & T. Stivers (Eds.), *Person reference in interaction: Linguistic, cultural, and social perspectives* (pp. 1–20). Cambridge: Cambridge University Press.
Stuckenberg, A., & O'Connell, D. C. (1988). The long and the short of it: Reports of pause occurrence and duration in speech. *Journal of Psycholinguistic Research, 17*, 19–28.
Suleiman, C. (2000). *Pronoun use in television interviews: Social interaction and the Middle East peace process*. Unpublished doctoral dissertation. Georgetown University, Washington, DC.
Suleiman, C., Lucas, R., Blum-Kulka, S., Kampf, Z., & Liebes, T. (2001, September). *Talking to friend and foe: TV interviewing in conflict*. Paper presented at the Ross Priory Broadcast Talk Annual Meeting, Glasgow, Scotland.
Suleiman, C., & O'Connell, D. C. (2003). Perspective in the discourse of war: The case of Colin Powell. *Pragmatics, 13*, 401–422.
Suleiman, C., & O'Connell, D. C. (2007). Bill Clinton on the Middle East: Perspective in media interviews. *Studies in Language & Capitalism, 2*, 75–100. Retrieved March 7, 2007, from http://languageandcapitalism.info
Suleiman, C., & O'Connell, D. C. (2008). Gender differences in the media interviews of Bill Clinton and Hillary Clinton. *Journal of Psycholinguistic Research, 37*, 33–48.
Suleiman, C., O'Connell, D. C., & Kowal, S. (2002). "If you and I, if we, in this later day, lose that sacred fire. . .": Perspective in political interviews. *Journal of Psycholinguistic Research, 31*, 281–299.
Svartvik, J., & Quirk, R. (1980). *A corpus of English conversation*. Lund: Gleerup.
Swets, B., Ferreira, F., & Altmann, E. M. (2006, November 18). *"Where was I?": A psycholinguistic investigation of interruptions to language production*. Paper presented at the 47th annual meeting of the Psychonomic Society, Houston, TX.
Szawara, J., & O'Connell, D. C. (1977). Temporal reflections of spontaneity in homilies. *Bulletin of the Psychonomic Society, 9*, 360–362.
Tannen, D. (1983). When is an overlap not an interruption? In R. J. Di Pietro (Ed.), *The first Delaware symposium on language studies* (pp. 119–129). Newark: University of Delaware Press.
Tannen, D. (1986). *That's not what I meant! How conversational style makes or breaks your relations with others*. New York: William Morrow.
Tannen, D., & Saville-Troike, M. (Eds.). (1985). *Perspectives on silence*. Norwood, NJ: Ablex.
Tartter, V. C. (1998). *Language and its normal processing*. Thousand Oaks, CA: Sage.
Taylor, T. J. (1997). *Theorizing language*. Amsterdam: Pergamon.
Taylor, T. J., & Cameron, D. (1987). *Analysing conversation: Rules and units in the structure of talk*. Oxford: Pergamon.
Tesnière, L. (1936). Sur la classification des interjections. In A. Šestak & A. Dokoupil (Eds.), *Mélanges dédiés à la mémoire de Prokop M. Haškovec par ses amis et ses élèves* (pp. 343–353). Brno, Czech Republic: Globus.
Thorndike, E. L. (1948). The psychology of punctuation. *American Journal of Psychology, 61*, 222–228.
Tierney, J. (2007, March 13). What's so funny? Well, maybe nothing. *New York Times, Science Times*, pp. D 1, D 6.
Timmermann, B., & Baker, F. (2002). *Der dritte Mann: Auf den Spuren eines Filmklassikers*. Wien: Czernin.

Tolson, A. (2006). *Media talk: Spoken discourse on TV and radio*. Edinburgh: Edinburgh University Press.
Tosi, O. (1965). *A method for acoustic segmentation of continuous sound into pauses and signals and measurement of segment durations*. Unpublished doctoral dissertation, Ohio State University, Columbus, Ohio.
Tracy, K. (2002). *Everyday talk: Building and reflecting identities*. New York: Guilford.
Traxler, M., & Gernsbacher, M. (Eds.). (2006). *Handbook of psycholinguistics* (2nd ed.). New York: Academic Press.
Trouvain, J. (2003). Segmenting phonetic units in laughter. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th international congress of phonetic sciences* (pp. 2793–2796). Barcelona: Universitat Autònoma de Barcelona.
Tufte, E. (2003). *The cognitive style of power point: Pitching out corrupts within* (2nd ed.). Chesire, CT: Graphics Press.
Turner, F., & Pöppel, E. (1983). The neural lyre: Poetic meter, the brain, and time. *Poetry, 142*, 273–309.
Upchurch, C. M., & O'Connell, D. C. (2000). Typical Clinton: Brazen it out. *Journal of Psycholinguistic Research, 29*, 423–431.
Uppstad, P. H., & Solheim, O. J. (2007). Aspects of fluency in writing. *Journal of Psycholinguistic Research, 36*, 79–87.
Urban, G. (1996). Entextualization, replication, and power. In M. Silverstein & G. Urban (Eds.), *Natural histories of discourse* (pp. 21–44). Chicago: University of Chicago Press.
Van de Water, D. A., & O'Connell, D. C. (1986). From page to program: Some typographical and temporal variables in radio homilies. *Journal of Psycholinguistic Research, 15*, 525–538.
Védénina, L. G. (1973). La transmission par la ponctuation des rapports du code oral avec le code écrit. *Langue Française, 19*, 33–40.
Verhaar, J. W. M. (1963). *Some relations between perception, speech and thought: A contribution towards the phenomenology of speech*. Assen, Netherlands: Van Gorcum.
Voloshinov, V. N. (1973). *Marxism and the philosophy of language* (L. Matejka & I. R. Titunik, Trans.). New York: Seminar Press.
Vygotsky, L. L. (1934/1962). *Thought and language* (G. Vakar & E. Hanfmann, Trans.). Cambridge, MA: The MIT Press.
Wagner, L., Kako, E., Amick, E., Carrigan, E., & Liu, K. (2005). Children's use of pointing to anchor reference during story-telling. In A. Brugos, M. R. Clark-Cotton, & S. Ha (Eds.), *Proceedings of the 29th annual Boston University conference on language development* (pp. 639–650). Somerville, MA: Cascadilla Press.
Walker, A. D. (1986). The verbatim record: The myth and the reality. In S. Fisher & D. Todd (Eds.), *Discourse and institutional authority: Medicine, education, and law* (pp. 205–222). Norwood, NJ: Ablex.
Walker, J. (1811). *Elements of elocution*. Philadelphia: Ferrett and Walton.
Waller, R. H. W. (1980). Graphic aspects of complex texts: Typography as macro-punctuation. In P. A. Kolers & M. E. Wrolstad (Eds.), *Processing of visible language* (Vol. 2, pp. 241–253). New York: Plenum Press.
Warner, J. L., & Cukor, G. (Producers). (1964/1986). My fair lady [Motion picture]. (Available from CBS/FOX VIDEO, 3900 Seven Mile Road, Livonia, MI 48152)
Washburn, R. W. (1929). A study of the smiling and laughing of infants in the first year of life. *Genetic Psychology Monographs, 6*, 397–535.
Weingarten, R. (n.d.). *Interpunktion im Schreibprozeß*. Retrieved December 7, 2006, from http://www.ruediger-weingarten.de/Texte/Interpunktion.pdf
Werner, M. (2006). *Am Hang*. Frankfurt/Main: Fischer Taschenbuch.
Wertsch, J. V. (2003). Ragnar Rommetveit: His work and influence. *Mind, Culture, and Activity: An International Journal, 10*, 183–185.
West, D. M. (1984). Cheers and jeers: Candidate presentations and audience reactions in the 1980 presidential campaign. *American Politics Quarterly, 12*, 23–50.

Whitney, P. (1998). *The psychology of language*. Boston: Houghton Mifflin.
Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin and Review, 12*, 957–968.
Wilson, T. P., & Zimmerman, D. H. (1986). The structure of silence between turns in two-party conversation. *Discourse Processes, 9*, 375–390.
Winteler, J. (1892). *Naturlaute und Sprache: Ausführungen zu W. Wackernagels* Voces variae animalium. Aarau, Switzerland: H. R. Sauerländer.
Wunderlich, D. (1972). Sprechakte. In U. Maas & D. Wunderlich (Eds.), *Pragmatik und sprachliches Handeln: Mit einer Kritik am Funkkolleg "Sprache"* (pp. 69–188). Frankfurt/Main: Athenäum.
Wundt, W. (1900–1920). *Völkerpsychologie: Eine Untersuchung der Entwicklungsgesetze von Sprache, Mythos und Sitte* (10 vols.). Leipzig: Engelmann.
Wundt, W. (1911). *Völkerpsychologie: Eine Untersuchung der Entwicklungsgesetze von Sprache, Mythos und Sitte. Vol. 1: Die Sprache* (3rd ed.). Leipzig: Engelmann.
Yakovleva, E. (2004). *Deutsche und russische Gespräche: Ein Beitrag zur interkulturellen Pragmatik*. Tübingen: Niemeyer.
Yang, C. (2004). Interjektionen im Sprachvergleich: Deutsch versus Chinesisch. *Zeitschrift für Semiotik, 26*, 47–70.
Zimmermann, K. (1983). Überlegungen zu einer Theorie des Schweigens. In I. Rosengren (Ed.), *Sprache und Pragmatik: Lunder Symposium 1982* (pp. 37–45). Stockholm: Almqvist & Wiksell International.
Zimmermann, D. H., & West, C. (1975). Sex roles, interruptions and silences in conversation. In B. Thorne & M. Hanley (Eds.), *Language and sex: Difference and dominance* (pp. 105–129). Rowley, MA: Newbury House.
Zuo, Y. (2002). *The golden silence: A pragmatic study on silence in dyadic English conversation*. München: LINCOM Europa.

Name Index

Parentheses around initials indicate occasional inconsistent use thereof.

Subject Index

N.B. Foreign words and phrases as well as ordinary English words and phrases used with specialized meanings are printed herein in italic font.