INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Prosody, syntax, and discourse pragmatics: Assessing information flow in German conversation

Schuetze-Coburn, Stephan Mark, Ph.D.

University of California, Los Angeles, 1994
Prosody, Syntax, and Discourse Pragmatics:
Assessing Information Flow in German Conversation

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Linguistics

by

Stephan Mark Schuetze-Coburn

1994
The dissertation of Stephan Mark Schuetze-Coburn is approved.

Robert S. Kirsner

Peter Ladetoged

Terence H. Wilbur

John W. Du Bois, Committee Co-chair

Pamela Munro, Committee Co-chair

University of California, Los Angeles

1994
To
Mnemosyne
TABLE OF CONTENTS

Dedication.................................................................................................................. iii
Table of Contents ........................................................................................................ iv
List of Tables .............................................................................................................. vii
List of Figures ............................................................................................................. viii
Acknowledgments ..................................................................................................... x
Vita ............................................................................................................................... xiv
Abstract .................................................................................................................... xv

1. The Study of Discourse ......................................................................................... 1
1.1 Discourse and discourse data ........................................................................... 3
1.2 Previous studies of German discourse ............................................................. 9
1.3 The aims of this investigation .......................................................................... 15
1.4 The corpora and their representation ............................................................... 17
1.4.1 Primary corpus ............................................................................................. 18
1.4.2 Secondary corpus ......................................................................................... 22
1.4.3 Transcription ................................................................................................. 23
1.5 Turn taking ......................................................................................................... 35
1.6 Turns in the primary corpus ............................................................................. 47
1.7 Summary ............................................................................................................. 49

2. Discourse Prosody ............................................................................................... 50
2.1 Prosody, intonation, and discourse ................................................................. 51
2.2 Prosodic units .................................................................................................... 69
2.2.1 The basis of prosodic segmentation ............................................................ 72
4.2 Information flow........................................................................................................277
4.2.1 Given and new information..................................................................................278
4.2.2 Activation state.........................................................................................................281
4.2.3 Identifiability.............................................................................................................289
4.2.4 Other discourse properties......................................................................................292
4.3 Probing fundamental notions of information flow.........................................................295
4.3.1 Concept and referent...............................................................................................296
4.3.2 The determination of activation state.......................................................................312
4.4 The distribution of new mentions and referential tracking...........................................333
4.5 Evaluating hypotheses about information flow............................................................336
4.5.1 New information and transitive subjects.................................................................337
4.5.2 New information and grammatical subjects............................................................342
4.5.3 New information and clausal arguments...............................................................345
4.5.4 New information and intonation units...................................................................349
4.6 Summary.....................................................................................................................353

5. An Analysis of Conversational Discourse.....................................................................355
5.1 The conversational story ‘Dirges’..............................................................................360
5.2 Prosodic analysis of ‘Dirges’......................................................................................365
5.3 Clausal analysis of ‘Dirges’........................................................................................383
5.4 Discourse-pragmatic analysis of ‘Dirges’.................................................................391
5.5 Summary.....................................................................................................................407

6. Summary and Prospects..............................................................................................409

References.......................................................................................................................416
Appendix..........................................................................................................................450
LIST OF TABLES

Table 1.1. Speech events, speakers, and stories in the primary corpus......19
Table 1.2. Reduced forms of indefinite and definite articles..................32
Table 1.3. Forms of personal pronouns........................................34
Table 1.4. Examples of other reduced forms...................................34
Table 1.5. Turn status of conversational utterances............................45
Table 1.6. Frequency of turns and backchannels by speaker...............48
Table 2.1. Types & sources of data in German intonation studies..........62
Table 2.2. Some prosodic units and hierarchical organization..............85
Table 2.3. Descriptive statistics for the distribution of contour types.....163
Table 3.1. General syntactic characterization of IUs in the primary
    corpus subset..................................................................182
Table 3.2. General syntactic characterization of IUs in the primary
    corpus subset, allowing for partial alignment..............................184
Table 3.3. Frequency of three clause types.....................................257
Table 3.4. Frequency of three clause types with respect to position of
    the finite verb................................................................259
Table 3.5. Frequency of intermediate clause boundaries in proper
    clauses........................................................................261
Table 3.6. Frequency of intermediate clause boundaries with respect
    to clause type................................................................262
Table 3.7. Frequency of peripheral elements in clause constructions......263
Table 3.8. Frequency of clause constructions in terms of IUs per
    clause...........................................................................268
Table 4.1. Interaction of activation state and identifiability in tracking
    referents........................................................................331
Table 4.2. Referential function for new nominal mentions in corpus.....334
LIST OF FIGURES

Figure 2.1. F0 reset (step up) at an IU boundary ........................................ 98
Figure 2.2. F0 reset (step down) at an IU boundary ..................................... 99
Figure 2.3. Example of a stop closure with a long duration ....................... 123
Figure 2.4. Waveforms of corpus excerpt (12) ........................................... 131
Figure 2.5. Coordinate transcription of a corpus excerpt ............................ 153
Figure 2.6. Types of IU entries in the corpus database ............................... 156
Figure 2.7. Frequency of IUs in the primary corpus subset as a function of syllables per IU ................................................................. 159
Figure 2.8. IUs types in the primary corpus subset .................................... 160
Figure 2.9. Frequency of IUs in the primary corpus subset as a function of syllables per IU, by backchannel vs. turn proper ................................................................. 161
Figure 2.10. Frequency of IUs in the primary corpus subset as a function of syllables per IU, by contour type ................................................. 162
Figure 3.1. Waveform of an intonation unit ............................................. 187
Figure 3.2. Dependency (a) and constituency (b) of a conversational excerpt ..................................................................................... 190
Figure 3.3. Types of clause entries in the corpus database ....................... 250
Figure 3.4. Frequency of clauses in the basic clausal subset .................... 252
Figure 3.5. Frequency of clauses in the basic clausal subset as a function of words per clause, with clause equivalent constructions as a separate category ........................................ 253
Figure 3.6. Verbal characteristics of clauses in the primary clause subset ..................................................................................... 254
Figure 3.7. Frequency of proper clauses as a function of words per clause, with contributions of basic clause types indicated ........................................ 257
Figure 3.8. Frequency of IUs in the primary corpus subset as a function of words per IU, with backchannels and clause equivalents as a separate categories .............................................265

Figure 3.9. Frequency of IUs in the primary corpus subset in syllables per IU..................................................................266

Figure 3.10. Frequency of IUs in the primary corpus subset in syllables per IU.................................................................267

Figure 5.1. F0 tracings for IUs 134-144.................................................................366

Figure 5.2. F0 tracings for IUs 145-151.................................................................369

Figure 5.3. F0 tracings for IUs 152-156.................................................................373

Figure 5.4. F0 tracings for IUs 157-161.................................................................376

Figure 5.5. F0 tracings for IUs 163 & 167-8..........................................................378

Figure 5.6. F0 tracings for IUs 169-176.................................................................380
ACKNOWLEDGMENTS

It would be platitudinous for me to say that a work of this size—and duration—necessarily owes a great deal to a great number of people. Nevertheless, it is very true, and I wish to express my gratitude to some of the individuals who have helped shape my thinking over the years and who have offered me exceptional opportunities. Many others, I regret, must remain unnamed here.

The most direct and rewarding influence, of course, has come from my exceptional committee members. My intellectual debt to Jack Du Bois is incalculable. It is primarily to him that I owe my understanding of discourse, transcription, and the value (and limitations) of discourse pragmatics. I thank him for his willingness to devote countless hours at a time to what for me has been very fruitful dialog; his ideas from our talks will be recognizable throughout this dissertation.

I am very pleased to say that all chapters here have benefited enormously from the careful comments of Pam Munro, who also deserves much credit for repeatedly getting the ball rolling. If ever a doctoral candidate has had the good fortune to have every snippet thoroughly scrutinized by a sage advisor, I have had this pleasure. She has, however, judiciously allowed me to learn the hard way, so that she is not to be taken to task for any remaining infelicities. My study of Muskogean languages under her tutelage has vastly broadened my understanding of linguistic structure, and it is with her, too, that I first learned solid field research tech-
niques, as well as the patience that an eager linguist must often have in working with members of different language communities.

Warmest thanks are due to Peter Ladefoged and Sandy Thompson for much support during my graduate career and for serving on my committee. Both individuals have been exemplary inspirational models for a willing student. It was Peter who got me thinking seriously about phonetics in the first place, who encouraged me to pursue projects in the Phonetics Lab, and who provided opportunities on more than one occasion for me to put my knowledge to practical use. Sandy encouraged me to view old problems in new ways, but at the same time made sure I did not get caught up in the quagmire of the unsolvable. I thank her, too, for always taking time out for me from an exceedingly busy schedule. I regret that I have not always been able to follow her advice.

I also want to thank Bob Kirsner and Terence Wilbur, my outside committee members, very much for their comments and assistance. All too often it is too easy to turn inward when one concentrates on a subject so intently. Their unique perspectives have helped to broaden my thinking in a number of crucial respects.

Numerous other teachers and colleagues have played significant roles in making this dissertation what it is. Several symposia and workshops I had the privilege of attending provided a critical mass of brilliant minds on whom I could test a number of my ideas. I would like especially to thank the participants at the Workshop on Information Flow, UC Santa Barbara, 29-30 April 1989; at the Workshop on Reference, UC Santa Barbara, 7-8 Oc-
tober 1989; and at the Workshop on Prosody in Natural Speech, University of Pennsylvania, 5-12 August 1992. Thanks as well to Mark Liberman and Cindie McLemore for their assistance in the Phonetics Lab at the University of Pennsylvania, and to Felicia Hurewitz, also in the Linguistics Department at Penn, for digitizing and pitch tracking the conversational excerpt analyzed in Chapter 5.

Special notes of thanks go to Dieter Metzing and Peter Schlobinski. Both wielded instrumental influence in the early stages of this project. Dieter did his best to see that my stay in Bielefeld was productive; Peter provided welcomed advice and a sorely needed context to discuss all the right issues. My work with him has been highly rewarding.

Other individuals with whom I have had the pleasure to discuss topics contained herein include Jules Archer, Dani Byrd, Wally Chafe, Susanna Cumming, Candy and Chuck Goodwin, Shoichi Iwasaki, Peter Lutzeier, Marianne Mithun, Yoshi Ono, Marian Shapley, and Betsy Weber. I hope they will not be disappointed with the use I made of our interactions.

Most of the conversational discourse that I have spent hundreds—if not thousands—of hours listening to was recorded in 1987-8 in Bielefeld, Germany, during a stay funded by the Deutscher Akademischer Austauschdienst. This generous support got my dissertation off to a fine running start. Additional conversational material was supplied by Reinhard Fiehler. Many thanks to him, and also to Roland Posner and Sabine Kowal, who permitted me to use their German 'Pear Film' recordings.
I thank Christel Bueb, Matthias Dinkelmeyer, Marian Flaherty, Reinhard Kaiser, Marlene Marlow, Thilo Metzing, Albert Nienhaus, Beate Ritz-Barr, Silvia Rode, Andrej Schlobinski, and Nicola Wolf for their advice, comments, and help with the transcriptions. The ears of Hartmut Kreft were especially acute in comprehending passages that others could not quite figure out. In a work that depends on the accurate perception of nearly every syllable, his is a gift for which I am very grateful.

Penultimately, this dissertation would in all likelihood not have been written had it not been for the unwitting influence of Margaret Schuetze and Gilbert Wright, for the support of Irene and Winston Roeth, for advice at just the right moments from George Settera and Karl Zimmer, and for the wonderful creations of H. Dresner, S. Jobs & Wozniak, and A. Oberwelland, not necessarily always in that order.

Finally, for your Hiobsgeduld, thank you thank you, M.
VITA

1981
B.A., Linguistics
University of California, Berkeley
Berkeley, California

1984-1986
Teaching Assistant
Department of Linguistics
University of California, Los Angeles

1987
M.A., Linguistics
University of California, Los Angeles
Los Angeles, California

1988
C.Phil., Linguistics
University of California, Los Angeles
Los Angeles, California

1989
Teaching Fellow
Department of Linguistics
University of California, Los Angeles

PUBLICATIONS


ABSTRACT OF THE DISSERTATION

Prosody, Syntax, and Discourse Pragmatics:
Assessing Information Flow in German Conversation

by

Stephan Mark Schuetze-Coburn
Doctor of Philosophy
in Linguistics
University of California, Los Angeles, 1994
Professor John W. Du Bois, Co-chair
Professor Pamela Munro, Co-chair

In this dissertation I discuss methodological and theoretical problems relating to the prosodic and syntactic segmentation of conversational discourse, and to the determination of information flow. An auditorily defined prosodic phrase, the 'intonation unit', is proposed as the basic unit of analysis in the domain of prosody. I discuss and exemplify in detail the prosodic features that define this unit. In the domain of syntax, I propose a clausal unit, the 'clause construction', as the most relevant to the analysis of discourse. Clause constructions incorporate the prosodic structure of the discourse, thereby capturing a temporal aspect of language production. In the domain of discourse pragmatics, I clarify certain notions in Chafe's in-
formation flow model, which leads to a full account of problematic data. I propose the strict separation of referent-concepts and class-concepts, each of which bears its own activation status.

The theoretical framework I articulate is then applied to a corpus of ordinary conversational German, thereby exemplifying the notion of whole-text analysis. The separate, but parallel, analyses of basic prosodic and syntactic structure and of information flow allow for an examination of the interaction of the three represented dimensions. In this context I evaluate four previously proposed hypotheses in light of the German conversational data. All four, Du Bois' 'Given A Constraint' and 'One New Argument Constraint' and Chafe's 'Light Subject Constraint' and 'One New Idea Constraint', are strongly supported.
1. The Study of Discourse

Sie selbst [die Sprache] ist kein Werk (Ergon), sondern eine Thätigkeit (Energie). ... Sie ist nämlich die sich ewig wiederholende ARBEIT DES GEISTES, den ARTICULIRTEN LAUT zum Ausdruck des GEDANKEN fähig zu machen. Unmittelbar und streng genommen, ist dies die Definition des jedesmaligen SPRECHENS; aber in wahren und wesentlichen Sinne kann man auch nur gleichsam die Totalität dieses Sprechens als die Sprache ansehen. Denn in dem zerstreuten Chaos von Wörtern und Regeln, welches wir wohl eine Sprache zu nennen pflegen, ist nur das durch jenes Sprechen hervorgebrachte EINZELNE vorhanden .... Gerade das Höchste und Feinste läßt sich an jenen getrennten Elementen nicht erkennen, und kann nur, was um so mehr beweist, daß die eigentliche Sprache in dem Acte ihres wirklichen Her vorbringens liegt, in der VERBUNDENEN REDE wahrgenommen oder geahndet werden.

Language is not a product (Ergon), but a process (Energie). ... It is the constantly repeating WORK OF THE MIND to express THOUGHTS by means of ARTICULATED SOUND. Taken literally, this is the definition of an act of SPEAKING, but in a valid and essential sense, it is only the totality of this speaking, as it were, that can be viewed as language. In the dispersed chaos of words and rules which we are in the habit of calling language, only the SINGLE ELEMENT produced by this speaking is present .... That which is highest and finest cannot be recognized from these isolated elements, but can only be perceived or surmised in CONNECTED DISCOURSE. This proves all the more that actual language lies in the act of its production.

Wilhelm von Humboldt

Über die Verschiedenheit des menschlichen Sprachbaues, 1836, p. LVII.

Interest in the study of discourse has come from a multitude of quarters. From as diverse practices as Conversation Analysis (e.g. Goodwin, Sacks, Schegloff), Ethnography of Speaking (e.g. Gumperz, Hymes), Form-Content Analysis (e.g. Diver, Kirsner, Reid), and Systemics (e.g. Coulthard, Halliday, Sinclair) comes a committed attention to language in ordinary use. All recognize, each in their own way, just as von Humboldt did over one hundred fifty years ago, that language cannot be understood
apart from discourse, that meaning is embedded in parole, in practical action, and that form, too, is motivated by the composite aggregate of linguistic tokens, by the totality of speaking. And now more than ever, at a time of rapidly shifting fashions, when yesterday's theories seem to combust almost as spontaneously as they were generated, and the prospects of generating something more enduring are remote, the empirical challenges that discourse has to offer not only provide a fresh source of inspiration for the linguist leery (and weary) of the chaos of isolated words and rules, but also the opportunity to found a new, more promising approach to the inner form of language, to the higher and finer.

It is in the spirit of viewing language as discourse that the focus for this dissertation is ordinary conversation. It has been observed that what people know about how they actually use language is quite limited (e.g. Silverstein 1981), thus it is incumbent on the linguist—of whatever persuasion—to explore the vast sea of language that we all sail every day. It is in the spirit of viewing language as a process that this study attempts an analysis of conversation in three dynamic linguistic dimensions. Prosody by its very nature is a dynamic phenomenon. Ever-changing, but always present in speech, it provides a much-needed means—both for its users and for the analyst—of apprehending all structure in discourse, and thus it will serve as a stepping stone to other dimensions in this study. Syntax is commonly reified—like language itself—but in the steady flow of the rhythmic current of discourse, it, too, takes on a more fluid shape. The discourse pragmatics of information flow, as the appellative implies, is also inherently dynamic. While our minds may slow at times, they never
fully drop to absolute zero. These three dimensions are indeed coordinated, but their interplay and the trace patterns that thereby emerge show us some of the details of the workings of the mind.

In this first chapter of this dissertation, I will touch on a number of preliminary issues. To begin with, I will explain what I hold the most important descriptive variables of discourse to be, and what kind of discourse will be the subject of this study. In the second section I will briefly examine previous studies of German discourse, measuring them against these variables. Next, I will outline the specific goals of this dissertation, as well as its organization. In the fourth section I will describe my corpora of spoken German and explain the system of transcription used in the examples in all of the following chapters. Thereupon I will discuss turn taking, which I view as a discourse category intermediate to transcription and analysis, and present my notational system. In the final section I will outline the basic turn structure of my corpus of conversational German discourse.

1.1 Discourse and discourse data

An important question that should be answered at the outset is: What is discourse? Conceived of broadly, this is a question that probably does not have an answer, at least not a useful one that I can outline here. With a broad scope, it is all too reminiscent of the Arthurian quest for an answer to the question: What is a text? Conceived of narrowly, however, it seems quite possible, indeed necessary, to provide an answer. In the context of this study, a narrow conception of ‘discourse’ answers the alternate
question: What kind of language samples will count as 'discourse data'? That is, by specifying what I am including in the corpus database, I am delimiting one type of discourse from other types of language use. For my purposes, discourse consists of natural, unscripted, spontaneous, connected speech. In other words, in this dissertation I will be concerned exclusively with a highly specific kind of spoken language.

In order to understand what I mean exactly by natural, unscripted, spontaneous, connected speech, an elaboration of each descriptive term is called for. The first, and in some sense, most important, qualification has to do with 'naturalness'. There are—perhaps surprising to a discourse analyst—various views as to what constitutes 'natural' speech. On the one hand, a distinction can be made between 'natural' vs. 'synthetic' speech, i.e. produced by humans vs. produced by machine. For researchers in speech synthesis and acoustic analysis, this is the primary dimension in which 'natural' has significance. On the other hand, there is a much more narrow distinction to be made between 'natural' vs. 'prompted' speech. Natural in this sense implies that the speech event is not set up for the express purposes of collecting data, but would have taken place anyway had the recorder or camera not been present (cf. Du Bois, Schuetze-Coburn, Paolino & Cumming 1992: 9).\footnote{This is not to claim that the outcome of a speech event in the context of a recording is exactly the same were the recording not made. The issue of the 'observer's paradox' (Labov 1972) is a complex one; for a discussion with a conversational example of the effect, see Henne & Rehbock (1982: 49-52); cf. also Stubbs (1983: 224-7).} While many linguists ascribe to
the first view, it is the second distinction that I mean here.\textsuperscript{2} Thus, elicited narratives, moderated discussions, and formal interviews (and other types of highly structured speech events) organized and managed to meet some analytic goal constitute discourse of a kind that differs substantially from the one that will be in focus here.

The second and third qualifications have to do with the nature of the preparation that underlies the speech. Natural speech that has not been set down in written form prior to the speech event is ‘unscripted’. Natural unscripted speech that has not been prepared mentally prior to the speech event itself is ‘spontaneous’. (While ‘spontaneous’ precludes ‘scripted’ for all practical purposes in the domain of spoken language, it does not for all domains; I will use both terms to emphasize the lack of written preparation.) Unlike the distinction between natural and prompted, the spontaneity of speech falls along a continuum: there different degrees of ‘forethought’ and ‘organizational preparation’ (Ochs 1979). It must be acknowledged, then, that natural unscripted speech can only be RELATIVELY spontaneous. What I consider to be the most important factor in this regard is the amount of preparation for a given speech event prior to the initiation of that event. One important implication of this interpretation of relativity is that ‘spontaneous’ does not imply

\textsuperscript{2} As an example of the prevalence of the broader conception, in the proceedings of a 1992 workshop on ‘Prosody in Natural Speech’ (Institute for Research in Cognitive Science, University of Pennsylvania), 15 reports are based on speech which is ‘natural’ in the sense of ‘not synthetic’. Only 12 reports include speech which would qualify as natural under the narrower definition.
'never before uttered'. The telling of a story in one natural setting, for example, does not negate the spontaneity of its later retelling in another natural setting. Restricting the application of this parameter has great practical value, for it is, in general, impossible to know which sequences in any given recording of natural speech are recycled and which are truly novel.\(^3\) In certain instances, there may be direct indications of a lack of spontaneity in expression (e.g. comments to this effect by the interlocutors), but such overt clues cannot be expected.

The final qualification has to do with the accompanying context. 'Connected' speech is speech that includes enough linguistic and extralinguistic context to make an interpretation of the utterances possible. Basically, this requires the preservation of larger discourse units (see below). Being connected thus precludes excising bits of text, in particular, 'sentences', and analyzing them without further regard to their production, as is sometimes done in acoustic studies. A particularly striking example of the use of natural, spontaneous, NONCONNECTED speech is found in a study by Lieberman, Katz, Jongman, Zimmerman & Miller (1985). Lieberman et al. investigate the occurrence and measurement of declination in American English by analyzing a corpus of short, simple declarative sentences extracted from recordings of conversational speech. Needless to say, these bits of spontaneous speech do not constitute discourse of the kind I am interested in.

\(^3\) There is an important theoretical issue lurking here, which I cannot go into at present, namely the extent of 'lexicalization'; see Pawley & Syder (1983) and Pawley (1986) for the view that much speech is formulaic.
For all practical purposes, the use of discourse data in a linguistic analysis necessitates the establishment of a corpus of a relatively large number of utterances which have been properly transcribed and coded. While the mention of 'corpus' once evoked a negative image among linguists, who generally shunned corpus-based studies, this bias is no longer justified, as the contributions in Bergen Holtz & Schaeder (1979), Aarts & Meijs (1984; 1986; 1990), Meijs (1987), Kytö, Ihalainen & Rissanen (1988), and Edwards & Lampert (1993) amply demonstrate. Corpus linguistics has over the past decade developed a sophisticated set of approaches to the study of language: there is no insistence here on a mechanistic analytic approach, as was the case, for example, in Harris (1952).

Under the view expressed in this dissertation, one of the benefits of a corpus of discourse data is that it reflects a type of language use that edited, normalized language or introspection cannot provide. This last statement perhaps requires some amplification, for it is just this claim that is made by those who reject corpus-based approaches. The standard argument runs as follows: any given corpus will exemplify only a subset of the possible linguistic structures that are of potential interest. The use of intuition (or direct elicitation) is thus necessary to fill in the gaps. Thus, any analysis based solely on a corpus will be deficient to some degree. The fallacy in this line of reasoning is two-fold. First, the same limitations apply to the use of intuition and elicitation. The linguist's use of intuition is simply one (highly specialized) form of language contextualization.

---

4 Here and throughout I use 'utterance' in a pretheoretical, nontechnical way simply to mean any verbal contribution to the discourse by a speaker.
elicitation is another). Because the extent to which one can visualize different contexts for the use of linguistic expressions is limited—it is no small part of the training of a syntactician to expand one's skill at doing so—the use of intuition (or elicitation) invariably results in different subsets of possible linguistic structures. The salient difference in this regard between these two data types, then, is that discourse is common to the human experience, while the construction of linguistic examples is exclusively practiced by a small group of academics.

Second, although there is much rhetoric on accounting for the full range of data, systematically constructed paradigms, as Gibbon (1988: 6-7) points out, are rarely encountered. The practice of introspection and the use of highly selective data go hand in hand. A different situation exists in the study of discourse. Here, the analyst is ultimately responsible for the entire data sample, in its various dimensions (cf. Du Bois & Schuetze-Coburn 1993: 221-2). This requirement for an integrated 'whole-text analysis' commits the linguist to examining, describing, and explaining every feature as given, without the luxury of a mental rug under which one can sweep the grammatical crumbs that drop from one's theories.

---
5 For example, Ulvestad (1979: 92-9) discusses five topics in German syntax where the use of introspective and corpus-based methods has led to different descriptions.
6 See Bungarten (1979: 32-3) for a list of specific problems associated with intuition as the source of primary data. Bausch (1979a: 81) has a similar list for elicitation.
7 Thus, I believe the real reason that corpus-based studies are rejected is that it is much more troublesome to obtain exactly the one datum that one wants—and with discourse in the corpus, it is nearly impossible (cf. Ulvestad 1979: 100).
1.2 Previous studies of German discourse

Although genuine interest in spoken German as a linguistic topic extends back at least a century (Wunderlich 1894; Behaghel 1927 [1899]), and isolated earlier works exist (e.g. Baumgardner 1959), its study—apart from strictly dialectological and phonetic issues—is usually thought of as having been initiated not even thirty years ago with the publication of Leska (1965) and Zimmermann (1965), the first linguistic works that made use of a corpus of recorded, unscripted, connected speech. Over these three decades, research into spoken German has seen a broad array of questions posed and a variety of spoken language used as primary material. The latter is of special interest here, for if we restrict ourselves to a conception of discourse as ‘natural, unscripted, spontaneous, connected speech’, we would find in a review of the literature on spoken German that there are relatively few monographs using discourse as primary material. Most of the early works—and many of the more recent ones—fail to satisfy one or more of the required criteria. This is not to imply that linguists did not value this type of discourse data or were unaware of the lack of efforts to collect it. Early on, Rupp (1965: 20) states unequivocally:

Wir haben also Material über die gesprochene Sprache in Hülle und Fülle. Aber, und da kommt das Große Aber: wir besitzen trotz aller Aufnahmen bis heute kaum authentische und wissenschaftlich auswertbares Quellenmaterial für die gesprochene

---

8 See Kaempfert (1972, 1973); Raettig (1973); Betten (1977, 1978); also Rath (1979: 25-33) and Lappé (1983: 16-23), for complete surveys. There, in addition, are several unpublished Staatsarbeiten that survey the literature listed in these reports.
Sprache, wie sie immer und überall vollzieht. Gesprochene Sprache vollzieht sich nämlich zuerst und zum großen Teil im spontanen Gespräch zwischen zwei oder mehr Menschen. ...
Spontane Gespräche, wie sie sich in jedem Augenblick vollziehen, unreflektiert vollziehen, sind auf Platten und Bändern kaum zu finden.

(We have spoken language material in great quantities. But—and this is a big but—despite all our recordings we hardly have any authentic, usable materials with spoken language as it occurs everywhere all the time. Above all, spoken language occurs for the most part in spontaneous conversation between two or more persons. ... Spontaneous conversation, as it occurs every moment—without reflection—is hardly to be found on our records and tape recordings.)

In pointing out the problems in using only lectures, sermons, and news broadcasts on the one hand, or elicited monologs and interviews on the other, Rupp is clearly making a case for collecting and studying unscripted, spontaneous, connected speech in a naturally occurring context. Steger (1967), too, states the desideratum clearly, listing four, much cited (and debated), criteria as to what counts as ‘gesprochen’ (‘spoken’).

Still, nearly all the major linguistic works on spoken German in the 1960s and 70s cannot claim to have used this kind of discourse data, and the major corpus assembled during this period contains relatively little of it. Leska (1965) and Grundmann (1975) elicited monologs in order to be able to concentrate on ‘zusammenhängende Äußerungen’ (‘connected ut-
terances'). Wackernagel-Jolles (1971) used a mixed corpus of professional news interviews and recordings of university seminars, where a high degree of preparation must be suspected, as well as elicited narratives. Schulz (1973) is based on the written (and edited!) versions of a set of interviews with townspeople in a Ruhr locality made by a journalist. The Freiburger corpus of the Institut für deutsche Sprache (see Berens 1979) has been the source material for numerous spoken language studies, e.g. Dittmann (1976), Schoenthal (1976), Bausch (1979b), Sennekamp (1979), and Winkler (1979). Rath (1979) and Schank & Schoenthal (1983) also make extensive use of the Freiburg corpus. While some of this material undoubtedly qualifies as the discourse type I am targeting, the majority of it consists of reports, interviews, and discussions recorded from radio and television broadcasts, where it must be questioned how unscripted and spontaneous the speech is, as well as to what degree editing has taken place for the program. There is also another reservation to be expressed with such material. While speech events designed for public broadcasts are indeed ‘natural’—they would have taken place regardless of the interests of the researcher—there is an inherent aspect of such performances, which I believe is dangerously close to the ‘prompted’ feature I set as the polar opposite of ‘natural’. Many other studies have been based on other spoken language corpora, which are not discourse exhibiting the four qualities I outlined above. As examples may be mentioned Caroli (1977), Lappé (1983), Royé (1983), Linke (1985), Geißner (1988), Liedtke (1990), and Jin (1992). This is not to belittle the contributions that these

---

9 Originally for a television broadcast, published as Bottroper Protokolle (1968).
works have made to the understanding of spoken German. The studies I have mentioned in fact deserve close readings. I wish only to illustrate that there still is a dire need for investigations based on natural, unscripted, spontaneous, connected discourse.

Among the works on spoken German, there are a few full-length studies based on discourse of the type that is the focus of investigation here. To be singled out are Zimmermann (1965), Weiss (1975), Quasthoff (1980), and Selting (1987). In addition, it can be added that Ramge (1978) and Henne & Rehbock (1982) contain conversational discourse that is evaluated in the context of methodological presentations. Brons-Albert (1984) provides a corpus of telephone conversations, though without any analysis.

Undoubtedly the first to record and analyze a body of ordinary conversational German is Zimmermann (1965). Using his own native Basel dialect as the basis for his study, Zimmermann focuses on what might be termed ‘psycho-interactional’ aspects of discourse. His main interest is the relationship between thought and speech in the context of expression in dialog. Despite the subtitle of his monograph, he does not undertake a systematic syntactic analysis of the corpus—indeed, one of his points is that ‘[d]er mündliche Rede ... ist mit grammatikalischen Regeln nicht beizukommen’ (88) (‘spoken discourse does not respond to grammatical rules’).10 Neither is the prosodic structure of the utterances dealt with sys-

10 Because Zimmermann’s analysis is not based on colloquial Standard German, but on a form of Swiss-German, a comparison with the present study would be difficult anyway.
tematically. Yet his attention to phenomena such as ellipsis, grammatically anomalous constructions (‘anakolouthon’), parentheticals, and particles—topics which have subsequently been much studied (e.g. Weydt 1969; Rath 1975, 1976, 1979; Betten 1976)—did help to usher in a way of thinking about spoken language that was very different from that of traditional views.

Weiss (1975), on the other hand, does present a comprehensive syntactic analysis of a corpus of (Austrian) German discourse. In comparing speech that was produced in three different situations, he shows, among other things, that under certain conditions, the syntactic structure of utterances may be more or less oriented to written norms. The down side to Weiss’ study is that his segmentation of the corpus is purely syntactic. Although he had originally planned a prosodic segmentation as well, for ‘[d]em Material zu entsprechen scheint eine Segmentierung nach akustischen Kriterien ...’ (16-17) (‘a segmentation following acoustic criteria would seem appropriate for the material’), the attempts at an application of von Essen (1964) and Isačenko & Schädlich (1966) were not successful. He then has to resort to a sentence-oriented analysis, following Bausch (1971).

Quasthoff (1980) and Selting (1987) both reflect developments that began in the mid-1970s. Following a period which was characterized by an overemphasis on sentence syntax to the exclusion of other domains, linguists working with discourse turned to other problems. For Quasthoff, the problem is the structure and function of narratives in conversation,
and the development of a model that describes the process of narrative construction. The majority of her study is theoretically oriented, rather than analytic, although she does provide an analysis of the ‘historical present’ for her corpus, and she does discuss the occurrence and use of ‘hesitation phenomena’, ‘linkers’, and ‘lexical boundary markers’. For Selting, the problem is how communication is interactionally achieved, and especially how communication difficulties are signaled and resolved. She examines repair phenomena from a linguistic, primarily a prosodic, perspective. She shows that by attending to linguistic forms in discourse, a more coherent, differentiated typology of repair arises. The main significance of her work for the present study is the unequivocal demonstration of the vital role that is played by prosody in discourse, and that any study of spoken language which does not systematically represent significant aspects of prosodic structure must be viewed, at best, as highly provisional.

In sum, spoken German has been a major focus of research over the last thirty years, with specific morphosyntactic and interactional questions having been answered. Much of this work, however, has not been based on conversational discourse, but on speech whose naturalness and spontaneity is dubious or absent. Moreover, all individual studies of spoken German known to me have concentrated on answering questions in only one linguistic domain. In studies focusing on syntactic issues, the prosodic structure of the speech has been either disregarded or not represented systematically. In studies focusing on prosodic issues, the syntactic composition of the utterances has been of little concern. The few studies
exploring informational issues using a corpus of spoken language are based on experimentally derived speech.

1.3 The aims of this investigation

This dissertation has two aims. My ultimate goals are to contrast the basic prosodic structure found in conversational discourse with the corresponding syntactic structure and information flow, and to describe the patterns that result from linguistic interaction in these three dimensions. Specifically, I seek to outline the distribution of 'new nominal mentions' in my corpus of ordinary German conversation, and to evaluate this distribution with respect to several hypotheses that have been put forth regarding the interaction of prosodic, grammatical, and discourse-pragmatic categories. My second aim is related to the first. In order to achieve these goals, a number of fundamental questions must be answered. That is, I must first address basic methodological issues necessary for the linguistic analysis of conversation. Discussion of these issues and concrete proposals resulting from the discussion constitute a significant part of each of three central chapters: Chapter 2 on discourse prosody, Chapter 3 on discourse syntax, and Chapter 4 on discourse pragmatics.

Concerning the analysis of prosodic structure in conversational discourse, the following must be asked:

- What are the issues involved in a segmentation of discourse?
- Which prosodic units have relevance to discourse?
- What constitutes a basic segmentation?
- Which prosodic categories are relevant to a basic segmentation?
- How can the basic prosodic structure be represented systematically?
These questions will be answered in Chapter 2, where I also outline the results of the auditory analysis of the corpus utterances in terms of 'intonation units', the phrasal unit that forms the basis for all subsequent analysis.

Concerning the analysis of syntactic structure in conversational discourse, these questions must be asked:

- Which syntactic units are relevant to discourse?
- What is the role of 'sentence' in a syntactic description?
- How is basic syntax to be viewed?
- In what way is the realization of syntax a process?
- Can interactional aspects of discourse be incorporated?
- How can basic syntactic structure be represented?

I will answer this set of questions in Chapter 3. Following a profile of the 'clause constructions' in the corpus, this syntactic unit is contrasted with the prosodic structure of the preceding chapter.

Concerning a discourse-pragmatic analysis of conversational discourse, the following questions must be asked:

- How is information flow to be conceived?
- Which pragmatic categories are relevant to information flow?
- By what procedures is information flow determined?
- What is the status of referents in the information flow model?

After answering these questions in Chapter 4, it becomes possible to detail the interaction of information flow with the prosodic and syntactic structures. As stated above, I focus in particular on the expression of new referents. Then, a number of universal constraints that have been proposed in the literature relating information flow with prosodic structure and
grammar can be evaluated in light of the German conversational discourse.

Following the methodological considerations in Chapters 2-4 and the presentation and discussion of the summary results for the conversational corpus, I offer in Chapter 5 a detailed example of the application of the techniques developed in these earlier chapters by analyzing one of the conversational exchanges. Finally, in Chapter 6 I summarize my findings and discuss the outlook for future research.

Before turning to a discussion of discourse prosody in Chapter 2, I describe in the following section the corpora used in this study and the transcription procedures. In §1.5 I discuss turn taking and my system for representing turn structure. In §1.6 I report on this turn structure for my conversational corpus.

1.4 The corpora and their representation

The observations and analyses contained in this dissertation have come about through the consideration of a wide range of data. I have made use of grammars of Standard German and regional dialects, descriptions and explanations in the linguistics literature, constructed examples, the intuitions of native speakers, written texts, anecdotal remarks, and notes about something said in a particular context. Central, however, has been the role played by recordings and transcriptions of natural, unscripted, spontaneous, connected speech, especially conversational discourse. In other words, the analyses reported in Chapters 2-5 are informed
by various sources, but they based on a primary corpus containing only extended excerpts from naturally occurring conversation.

1.4.1 Primary corpus

The primary corpus used in this study consists of extended excerpts of six conversations recorded in and around the cities of Bielefeld and Bonn, Germany, in 1987-8. Except for two nonnative German speakers in ‘Doppelpkopf’ (simply named Z1 and Z2, as their contributions are treated separately from the others), all participants are native speakers of a northern variety of standard German.11 The conversations took place between friends and acquaintances, in each case in the home of one of the participants. In all, the recording was an incidental fact of the interaction, and was made as unobtrusively as possible.12 Table 1.1 below provides a summary of features of the conversational samples, listing the speech events, the (fictional) names of the participants, the names of coherent segments (‘stories’), the main speaker, and their duration. As the titles of some of the speech events indicate, the occasion for much of the conversation was

---

11 The relationship between standard and nonstandard varieties of language is viewed very differently in Germanic and Anglo-American traditions. I will largely follow Barbour (1987), who identifies five points in continuum:
   a) Formal standard German (*Hochsprache*), largely written, prescriptively defined
   b) Colloquial standard German (*standardnahe Umgangssprache*), prestige pronunciation
   c) Colloquial standard German with local accent
   d) Vernacular speech (*dialektliche Umgangssprache*), local accent
   e) Traditional dialect (*Dialekt*)

The speech of all participants ranges between types (b) and (c).

12 See Chapter 2 of Du Bois et al. (1992) on techniques for making a good recording, and Chapter 25 for the appropriate equipment.
<table>
<thead>
<tr>
<th>Speech event</th>
<th>Participants</th>
<th>Story</th>
<th>Main speaker</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skat</td>
<td>Käthi</td>
<td>Aldi drunk</td>
<td>Käthi</td>
<td>1' 57&quot;</td>
</tr>
<tr>
<td></td>
<td>Tamina Wolf</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morning visit</td>
<td>Käthi</td>
<td>Gas stove</td>
<td>Käthi</td>
<td>59&quot;</td>
</tr>
<tr>
<td></td>
<td>Tamina</td>
<td>Washing machine</td>
<td>Tamina</td>
<td>2' 11&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Refrigerator</td>
<td>Tamina</td>
<td>1' 04&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Propane tank</td>
<td>Tamina</td>
<td>1' 12&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5' 26&quot;</td>
</tr>
<tr>
<td>Uli's visit</td>
<td>Anna</td>
<td>Gorges du Verdun</td>
<td>Uli</td>
<td>2' 18&quot;</td>
</tr>
<tr>
<td></td>
<td>Uli</td>
<td>Ambassador</td>
<td>Anna</td>
<td>4' 52&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vacation travel</td>
<td>Anna</td>
<td>3' 43&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Airport chaos</td>
<td>Uli</td>
<td>1' 42&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12' 35&quot;</td>
</tr>
<tr>
<td>Henrikas's visit</td>
<td>Günther</td>
<td>Getting by</td>
<td>Henrikas</td>
<td>5' 48&quot;</td>
</tr>
<tr>
<td></td>
<td>Henrikas</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ingo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afternoon tea</td>
<td>Elena</td>
<td>(Project)</td>
<td>(Sabine)</td>
<td>22&quot;</td>
</tr>
<tr>
<td></td>
<td>Sabine</td>
<td>Best kid</td>
<td>Elena</td>
<td>1' 48&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dirges</td>
<td>Elena</td>
<td>55&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cello lessons</td>
<td>Elena</td>
<td>1' 13&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4' 18&quot;</td>
</tr>
<tr>
<td>Doppelkopf a.</td>
<td>Diana</td>
<td>Visa for India</td>
<td>Diana</td>
<td>2' 01&quot;</td>
</tr>
<tr>
<td></td>
<td>Jörg</td>
<td>Weavers' fair</td>
<td>Marlies</td>
<td>2' 06&quot;</td>
</tr>
<tr>
<td></td>
<td>Marlies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Z1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Z2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>= 6 speech events</td>
<td>12 participants</td>
<td>15 stories</td>
<td>8 main speakers</td>
<td>34' 11&quot;</td>
</tr>
</tbody>
</table>

Table 1.1. Speech events, speakers, and stories in the primary corpus.
the visit of a friend: Käthi at Tamina’s, Uli at Anna’s, Henrika at Günther’s, and Sabine at Elena’s. Two of the speech events (‘Skit’ and ‘Doppelkopf’) took place in the context of evening card games.

Of the twelve German participants, eight related ‘stories’ (see below) that are included in the corpus, and therefore are considered ‘main speakers’ (see §1.5). The corpus contains a total of fifteen stories by the main speakers (the final portion of ‘Project’ is included in the corpus, but is not counted as a separate story). The division of the speech event selections into stories was done largely for practical reasons, but it also provides a useful heuristic for a characterization of the selections. In three of the four longer selections, there were definite transition points in the conversation marked by speaker and topic change, so that a division into stories seemed natural. In the selection from ‘Henrika’s visit’, no such points were evident, so it is considered to form one story, though this is clearly less satisfactory in this case. Similarly, the three shorter selections consisted of one story each.13 Effort has been made to include excerpts from all the stories in the course of presenting examples. The entire story ‘Dirges’ is given in Chapter 5 in conjunction with a detailed exemplification of the prosodic, syntactic, and discourse-pragmatic analyses.

Upon inspection of some of the corpus excerpts, my use of the term ‘story’ will perhaps seem somewhat surprising. I am not using it in a strict sense of ‘narrative’, where a definite formal internal structure (e.g. Labov

13 Note that ‘Visa for India’ and ‘Weavers’ fair’ are separate, noncontiguous selections from ‘Doppelkopf’, thus the labels ‘Doppelkopf a’ and ‘Doppelkopf b’ in the table.
& Waletzky 1967) or a minimal set of semantic and syntactic features (e.g. Quasthoff 1980) is required to be present. In natural conversation, the story structure must often be abstracted out of the talk, rather than existing on the surface (Polanyi 1985: 194), and thus I am applying the term in a broader, more holistic way. Note that this is a conception of story somewhat different from that discussed in Polanyi (1989: 20-6). The story ‘Dirges’ discussed in Chapter 5 provides a good example of how few formal characteristics need occur. Although a particular ‘storyworld’ is indeed described, the events and states are verbalized for the most part in the subjunctive mood, presenting a hypothetical situation. There is some minimal temporal ordering (e.g. as evidenced by the use of und dann ‘and then’ in line 168), but the story is largely devoid of ‘... discrete moments at which instantaneous occurrences are reported to take place’ (Polanyi 1989: 16). However, the discourse does seem to have a clear ‘point’ in relation to the rest of the conversation (cf. lines 132-4 with 174-6), an important feature of Polanyi’s ‘story’ narrative subtype.

Finally, the primary corpus contains exchanges that are not random slices from the conversations, but well motivated selections. Coherent story-like portions (including accompanying pre-story and post-story talk) from the speech events were chosen in which there was a dominant speaker. For each of the linguistic dimensions under investigation, it was judged advantageous to ensure that substantial contributions by individ-

14 Cf. also Gűlich & Quasthoff (1985: 170-2), who list four ‘frequently used definitional criteria’ with the concluding comment: ‘All these definitional criteria retain their validity only so long as they are couched in relatively vague terms’ (172).
ual participants were included in the corpus for each speech event. The main prosodic questions that are to be answered in Chapter 2 have to do with the segmentation of utterances. A corpus consisting of mostly short, highly interactional utterances would, however, impart a de facto segmentation (cf. Schuetze-Coburn, Shapley & Weber 1991: 217-8) by virtue of the frequent change in speakers. Thus, the desired interaction of prosodic cues would not be observable apart from the interactional structure. Nevertheless, the selections are still quite interactional in nature (see §1.6). Because the prosodic structure provides the fundament for subsequent analyses, the main syntactic questions to be addressed in Chapter 3 would also not be fully answered. The motivation for the discourse-pragmatic dimension is different. In Chapter 4 I will be focusing on the introduction of new referents. Conversational stories offer the clearest forum for an examination of referentiality in discourse.

1.4.2 Secondary corpus

A second corpus consisting of 20 experimentally elicited 'Pear film' narratives complemented the primary corpus.¹⁵ Though not qualifying as 'discourse' as defined above, these elicitations constituted unscripted, spontaneous, connected speech, whose properties were useful in realizing the goals of the study for a number of reasons. Due to a minimization of listener interaction and to the task the speakers were asked to do (retell the story in the film they had just seen, i.e. much like an oral version of

---

¹⁵ I thank Sabine Kowal and Prof. Roland Posner, Technische Universität Berlin, for the use of these data. See Chafe (1980) for details on the 'Pear film'.
Nacherzählung), the narratives were produced as long stretches of speech without interruption. Due to the structure of the film, there were numerous referents to describe in a short amount of time. These two facts produced a relatively high frequency of certain syntactic structures (e.g. relative clauses) which were used to explore clausal relationships in the development of an account of 'clause construction' that I outline in Chapter 3. Crucially, the semantic context was severely limited, as is typical for experimental situations, which facilitated the evaluation of the information flow model described in Chapter 4. In relating the events of the film, the speakers could only assume general 'mutual knowledge' that would be expected of any educated German listener. Furthermore, there were twenty instances of speakers reporting on the same stimulus, so that a range of formulations ('packaging', Chafe 1976) could be compared. Thus, the secondary corpus played a significant role in the development of my syntactic and discourse-pragmatic analytic frameworks; however, apart from using a few short excerpts to succinctly illustrate points in Chapter 3, no part of this corpus shows up directly in any analysis. That is, no systematic analysis of the German Pear film narratives was undertaken, and none of these data were included in the statistical compilations.

1.4.3 Transcription

Any systematic analysis of discourse requires a systematic representation of its features. The system of transcription I have use for both the primary and secondary corpora is largely the standard detailed in Du Bois,
Schuetze-Coburn, Paolino & Cumming (1992), which combines insights and conventions from a number of (mostly semicodified) traditions. An outline of this system can be found in Du Bois, Schuetze-Coburn, Cumming & Paolino (1993). An earlier version was published as Du Bois, Cumming & Schuetze-Coburn (1988). My major deviations from Du Bois et al. (1992) have to do with the representation of certain prosodic categories; in §2.3 I describe these fully, so they need not be treated here. Deviations from Du Bois et al. (1992) in other categories are minor and should cause no difficulty for those familiar with this standard. (Appendix 1 contains a summary of all transcription conventions used in the present work.)

One transcription matter, however, does need to be addressed that is often taken for granted: the orthographic representation. Segmental features can have an important bearing on the prosodic, syntactic, and discourse-pragmatic structures of the discourse, and their representation should be carefully considered. Du Bois et al. (1992) recommend standard orthography in rendering the lexical forms, because ‘... a discourse transcription does not generally seek to represent every variation in pronunciation’ (59). There are, however, a number of systematic departures from this principle that I have employed in my transcriptions, which I will outline here.

Edwards (1989), following Preston (1985), points out the possible inadequacies of nonstandard spellings, specifically drawing attention to the risks of introducing (a) spelling inconsistencies, (b) meaning ambiguities,
(c) un informativeness, (d) false phonetics, and (e) poor readability, into the transcription when impressionistic spelling is used. Any deviation from standard orthography, then, should in general be well motivated.

Most transcriptions of spoken German depart from current standard orthographic practice in one way or another. Many use so-called 'kleinschreibung', that is, lower-case spelling of nouns, instead of the standard capitalization (e.g. Rhode & Roßdeutscher 1973; Weiss 1975; Caroli 1977; Ramge 1978; Jäger (1979); Schaeffer 1979; Qua thoff 1980; Henne & Rehbock 1982; Schank & Schoenthal 1983; Geißner 1988; Brinker & Sager 1989; Keim & Schwitalla 1989; cf. also Pheby 1975), though to varying degrees.\textsuperscript{16} That is, in some transcriptions, lower-case letters are used consistently throughout, while in others, names, pronouns of respect, abbreviations, and the first word of the utterance may retain their capitalization. There seem to be two reasons for the use of kleinschreibung in transcriptions. First, the prescribed capitalization of all nouns is viewed as a peculiarity of written Standard German; in speech one does not hear upper vs. lower case.\textsuperscript{17} Second, capitalization is an iconic, and thus important, orthographic device for representing prominence of various kinds. While the first reason might be true, it is the second that provides an appropriate motivation for the deviation. The use of capitaliza-

\textsuperscript{16} Actually, behind the term 'kleinschreibung' lies a complex set of spelling reform issues which are currently being debated. See Mentrup (1992) for details.

\textsuperscript{17} This is similar to the argument put forward for not using ordinary punctuation marks—one does not hear commas and periods, but pauses and pitch contours (see e.g. Rhode & Roßdeutscher 1973: 30).
tion to systematically indicate accented syllables, for example, would justify the use of kleinschreibung. As the transcription system that I am using does not make use of capitalization to represent prosodic features, there is no compelling reason to adopt kleinschreibung here.

Another common departure from standard orthography is in the representation of regional dialects. Regional dialects possess widely recognized phonological and lexical variants which have achieved some degree of standardization in written representations. For example, the phrase Nichts Genaues weiß man 'Nothing certain is known' spoken as [mɪʃt jənauəst ves man niç] and transcribed Nischt Jenauet weß man niç would instantly be recognized as a Berlin dialect. There generally is, then, a conventional systematicity in the rendering of regional dialects, even if all phonetic differences are not captured. The inclusion of such variants in a discourse transcription obviously gives important information about the background of the participants and their attitude towards the interaction. There is a creative aspect at play here as well. Speakers who have some degree of control over a regional dialect can vary their use of standard and regional forms for particular rhetorical effects or for

---

18 Selting (1989) uses upper-case letters for 'primär akzentuierte Silben' ('primarily accented syllables'), but retains standard capitalization as well, perhaps with mixed success, as in later work (e.g. Selting 1992) she applies kleinschreibung consistently.

19 I am using the term 'regional dialect' here to encompass both Barbour's 'traditional dialect' and 'vernacular speech' (see fn. 12).

20 Example from Schlobinski (1984: 103).

21 For examples of the transcription of regional dialects in discourse studies, see Zimmermann (1965) and Lappé (1983).
other reasons (see e.g. Schlobinski 1988). Preservation of these variant forms in the transcription can be crucial to understanding the discourse, if nothing else.

In the case of the current corpus, as the speech of all participants is a variety of ‘standard, educated’ German, the question of dialect representation is less important. However, slight regional features for a majority of the speakers were identifiable—vowel quality differences, intonation patterns, and pronunciations of specific lexical items—and the possibility of their representation must be considered.\footnote{Also identifiable were regional expressions and syntactic constructions, but these are not at issue in the segmental representation.} As most of the recordings were made in and around Bielefeld, the prominent regional characteristics are Westphalian (Westfälisch), though one speaker was from the Ruhr and one lived a number of years in Berlin, and so these speakers show characteristics of speech of those areas instead. All lexical items and expressions used by speakers have, of course, been retained. Regionally influenced pronunciations were also included in the original transcriptions, but because the impact of these morphophonemic variants on the discourse was judged to be minor, it did not seem particularly important to include them in the excerpts.\footnote{Probably the most common variant is \textit{sacht} [zəkt] or [zəxt] ‘says’ in place of Standard German \textit{sagt} [zəkt].}

Much more important than the regional influences in the corpus are the divergences from Standard German pronunciation due to the nature of the speech events themselves. The events are informal and the speech
is correspondingly colloquial ('Umgangssprache'), so that it everywhere exhibits the segmental reductions that can be expected in rapid, casually spoken, connected utterances (see Bresson 1982; Dedenbach 1987; Kohler 1990 for details). A good deal of this type of reduction seems to be predictable; still, I argue much of it should be represented in a discourse transcription. One reason for including reduced forms is that speakers manipulate their realization. Selting (1983; 1985), for example, has shown that style shifting to and from a 'baseline' colloquial register can be used as a device to adopt a particular perspective or to refocus a topic in a discussion. Other strategic applications undoubtedly could be found, so that in general '... variation is not only a quantitative correlate of social, regional, or contextual parameters but is used by conversationalists as a device to signal social and interactive meaning ...' (Selting 1985: 195). 'Restoring' (i.e. replacing) actual segmental reductions with the corresponding full forms of the standard language would betray, at best, an overemphasis on the grammatical at the expense of the interactional aspects of the discourse.

A second reason to approximate colloquial segmental reductions in a transcription is the formal ambiguity that many reductions introduce, which cannot always be restored to the level of specificity the standard language would require. Neutralizations of at least the following normally important grammatical contrasts occur: definiteness (definite vs. indefinite articles), tense (present vs. past), case (nominative vs. accusative). An example of the interpretive problems segmental reduc-
tions can pose is given in (1).\textsuperscript{24}

(1) Corpus excerpt from ‘Cello lessons’ [Afternoon tea: 214-8].

<table>
<thead>
<tr>
<th>Line</th>
<th>English</th>
<th>German</th>
</tr>
</thead>
<tbody>
<tr>
<td>214</td>
<td>~da hät er bei so 'ner änder'n Frau /+ there has he by like a other woman</td>
<td></td>
</tr>
<tr>
<td>215</td>
<td>.. Céllounterricht [ bekòmm'n ] /+ cello.lessons received</td>
<td></td>
</tr>
<tr>
<td>216</td>
<td>(S): [ mhm ] /+</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>.. ~an 'er Musik .. schûle /+ at a/the music school</td>
<td></td>
</tr>
<tr>
<td>218</td>
<td>... ~an der Musikhôschule war das + at the music conservatory was that</td>
<td></td>
</tr>
</tbody>
</table>

214 [E]: so he (took) from another woman,  
215  cello lessons (),  
216  (S): mhm,  
217  at a/the music school,  
218  it was at the music conservatory;

In line 217 the preposition + article sequence \textit{an 'er}, pronounced [anɐ], occurs. When reduced, the indefinite article \textit{einer} will generally be realized as 'ner [nɐ] in such contexts, and the definite article \textit{der} will occur as 'er [ɐ], maintaining the distinction in definiteness (as is the case in line 214

\textsuperscript{24} Examples from the primary corpus are identified by story. All of my examples are numbered by line, each line corresponding exactly to one intonation unit (IU) in the talk (see §2.2.3). The speakers’ initials are given in the second column, indicating the turn structure (see §1.5 for details). Underneath the German text, a word-by-word gloss is given. A line-by-line translation follows each excerpt. Due to constituent order differences between English and German, it is occasionally necessary to move a constituent from one line in the translation to another. I have indicated this by enclosing the shifted constituent in parentheses and placing a pair of empty parentheses in the original line, as is done in lines 214-5 of (1). Also in the translations, small angled brackets ‘\&’ enclose words supplied by me to make the sense clearer.

29
with so 'ner). But because the preposition in this instance ends in an alveolar nasal, it is impossible to tell phonologically whether the source form is (ei)ner or (d)er. In many cases, contextual or semantic limitations will come into play, so that only one form will make sense or be possible.\(^{25}\) Yet in line 217, both articles seem plausible: a new situation is being introduced (hence the possible use of the indefinite article), but there may be only one music conservatory in question, as is implied by line 218 (hence the possible use of the definite article).\(^{26}\) In order to render this example in standard orthography, the transcriber would have to impart a grammatical distinction that the speaker has herself not articulated.

Apart from the interpretive difficulties, a rather peculiar situation exists with regard to the standard orthography of preposition + article sequences. There is inconsistency in the acceptance of reduced forms in the written language. A number of these contractions are accepted as standard, e.g. am (=an + dem) 'at/on the [MASC/NEUT.DAT.SG]', zur (=zu + der) 'to the [FEM.DAT.SG]', but others are considered nonstandard, e.g. an 'er (=an + der) 'at/on the [FEM.DAT.SG]' (as in the above example), and zu 'n (=zu + den) 'to the [DAT.PL]' (Duden 1984: 223), even though these forms also occur. Given the seemingly arbitrary distinction between standard and nonstandard (cf. Dedenbach 1987: 151-2), it does not make much sense to adhere to the written standard.

In order to represent the reductions, the question of possible forms

\(^{25}\) Kohler (1990: 73-4) makes this point, too, though he argues from constructed examples.

\(^{26}\) Other interpretations are also possible, and will be discussed in Chapter 4.
must be addressed. Dedenbach (1987: 42, 46) lists three to seven forms for each article. For example, for *eine* 'a' [NOM/ACC.F] four forms are given, from the unreduced [ainə] to [ainə] to [ənə] to the fully reduced [na]. What is important here, though, is not a detailed phonetic representation, but the occurrence of a systematic segmental reduction. For this reason, I have chosen to make a distinction only between ‘unreduced’ and ‘reduced’ forms (cf. Harweg 1989, who establishes a paradigmatic opposition between ‘strong’ and ‘weak’ forms of the definite article).

Table 1.2 below lists the reduced forms of the indefinite and definite articles that occur in the corpus, both in preposition + article sequences and elsewhere. To facilitate the identification of reduced forms in the transcription, the convention is used that an apostrophe (') always indicates some segmental reduction.\(^\text{27}\) Note that in some instances, a reduced article does have more than one representation listed. Where two reduced forms are given, this is of necessity due to the grammatical ambiguity discussed above that sometimes occurs with the reduction.

Another area where standard orthographic practice is problematic is in the representation of the pronouns. For each personal pronoun there is one standard form. In the spoken language, however, a case can be made for a richer system. In the corpus, three phonological levels can apparently be distinguished: ‘accented’, ‘unaccented’, and ‘reduced’. Apart from

\(^{27}\) Another option—preserving standard orthography—would be to mark the reduced form with a diacritic, e.g. unreduced *der* vs. reduced *dér*. I have opted instead for segmental spelling with an apostrophe, due to the problems of determining the full form discussed above. Also, this is common practice, and thus readily understandable.
<table>
<thead>
<tr>
<th></th>
<th>Unreduced</th>
<th>Reduced</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Indefinite</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[NOM.M/N]</td>
<td>ein</td>
<td>'n</td>
</tr>
<tr>
<td>[ACC.N]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ACC.M]</td>
<td>einen</td>
<td>ein'n/n</td>
</tr>
<tr>
<td>[DAT.M/N]</td>
<td>einem</td>
<td>'m</td>
</tr>
<tr>
<td>[NOM/ACC.F]</td>
<td>eine</td>
<td>'ne/e</td>
</tr>
<tr>
<td>[DAT.F]</td>
<td>einer</td>
<td>'ner/er</td>
</tr>
<tr>
<td><strong>Definite</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[NOM.SG.M]</td>
<td>der</td>
<td>'er</td>
</tr>
<tr>
<td>[DAT.SG.F]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ACC.SG.M]</td>
<td>den</td>
<td>d'n/n</td>
</tr>
<tr>
<td>[DAT.PL]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[DAT.SG.M/N]</td>
<td>dem</td>
<td>'m</td>
</tr>
<tr>
<td>[NOM/ACC.N]</td>
<td>das</td>
<td>d's/s</td>
</tr>
<tr>
<td>[NOM/ACC.F]</td>
<td>die</td>
<td>d'e/e</td>
</tr>
<tr>
<td>[NOM/ACC.PL]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.2. Reduced forms of indefinite and definite articles.

general treatments of segmental reduction and Dedenbach (1987), little attention has been devoted to pronominal variation in spoken Standard German.

Better coverage is given in descriptions of regional dialects, where perhaps the awareness (and acceptance) of spoken variants is greater. The case of Swiss-German provides an especially interesting contrast. In Switzerland, regional dialect use is nearly universal in most social spheres where spoken language is prominent (Barbour & Stevenson 1990: 214). For this reason, regional dialects there receive more or less equal—if not preferential—treatment vis-à-vis the Standard German (which is called *Schriftdeutsch*). In Marti’s (1985) *Berndeutsche-Grammatik*, for
example, we find exactly the type of coverage of the personal pronouns that might be expected in an adequate grammar of spoken colloquial German. Marti differentiates three levels, which he calls 'stressed' (betont), 'normal', and 'weakened' (geschwächt), with the remark:

Je nach der Stellung im Satz und der dem Pronomen zugedachten Wichtigkeit erscheint besonders das nachgestellte Personalpronomen in verschiedener Lautgestalt.... (92)

According to its position in the clause and its intended importance, a personal pronoun, especially when postposed, appears in different phonological forms.

Whether or not a grammatical distinction does exist for Standard German, pronominal variants have a great potential value for the analysis of discourse-pragmatic structure to be undertaken in Chapter 4, so that a systematic representation of the occurring variants can hardly be ignored.

Table 1.3 below lists representations of accented, unaccented, and reduced pronominal forms. Possible representations of forms that do not occur in my corpus are given in parentheses.

Beyond the reduction of articles and personal pronouns, certain reductions in other forms are noted in the transcriptions. Table 1.4 below lists a selection of these. Very common in colloquial speech are reductions of word-final -en and -e. Reductions of some final consonants, especially -t/-d, are also common. As above, an apostrophe is always used in a reduced form.
<table>
<thead>
<tr>
<th>Accented</th>
<th>Unaccented</th>
<th>Reduced</th>
</tr>
</thead>
<tbody>
<tr>
<td>'I' [NOM]</td>
<td>ich</td>
<td>ich</td>
</tr>
<tr>
<td>'we' [NOM]</td>
<td>wir</td>
<td>wir</td>
</tr>
<tr>
<td>'you' [NOM.SG]</td>
<td>dü</td>
<td>dü</td>
</tr>
<tr>
<td>'you' [NOM.PL]</td>
<td>ihr</td>
<td>ihr</td>
</tr>
<tr>
<td>'her' [DAT]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'he' [NOM]</td>
<td>er</td>
<td>er</td>
</tr>
<tr>
<td>'him' [ACC]</td>
<td>ihn</td>
<td>ihn</td>
</tr>
<tr>
<td>'him' [DAT]</td>
<td>ihm</td>
<td>ihm</td>
</tr>
<tr>
<td>'she' [NOM/ACC]</td>
<td>sie</td>
<td>sie</td>
</tr>
<tr>
<td>'they' [NOM/ACC]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'it' [NOM/ACC]</td>
<td>és</td>
<td>es</td>
</tr>
<tr>
<td>'them' [DAT]</td>
<td>ihnen</td>
<td>ihnen</td>
</tr>
<tr>
<td>'you' [RSP.NOM/ACC]</td>
<td>Sie</td>
<td>Sie</td>
</tr>
<tr>
<td>'you' [RSP.DAT]</td>
<td>Ihnen</td>
<td>Ihnen</td>
</tr>
</tbody>
</table>

Table 1.3. Forms of personal pronouns.

<table>
<thead>
<tr>
<th>[-an]</th>
<th>[-a]</th>
<th>[-t]</th>
<th>[-x]/[-x]</th>
<th>[-l]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbs</td>
<td>frag'n</td>
<td>hab'</td>
<td>is'</td>
<td></td>
</tr>
<tr>
<td>war'n</td>
<td>kern'</td>
<td>sin'</td>
<td>brauch'</td>
<td></td>
</tr>
<tr>
<td>Nouns</td>
<td>Brem'n</td>
<td></td>
<td>Arz'</td>
<td></td>
</tr>
<tr>
<td>Adjectives</td>
<td>schön'n</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particles</td>
<td>geg'n</td>
<td>solang'</td>
<td>un'</td>
<td>no'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.4. Examples of other reduced forms.
1.5 Turn taking

Fundamental to any conversational speech event is the interactional process generally known as 'turn taking'. For this reason alone, it is necessary here to provide a brief indication of what I take to be 'turn units' or 'speaker turns', or simply 'turns'. Yet there are good reasons for discussing turn taking in more detail, too. Of central importance is that the basic prosodic analysis of the conversational discourse into 'intonation units' that I describe in Chapter 2 assumes a prior analysis into 'speaker contributions'—not necessarily, it is noted, to be equated with turns, but a first step in that direction. The syntactic and discourse-pragmatic analyses also benefit from an awareness of turn structure, as will be explained below.

The basic observation is that a regular exchange of turns at speaking usually occurs among the participants in conversational speech (certainly true for my primary corpus). This orderly interactional process has been the subject of much attention, especially by ethnomethodologists. The classic paper is Sacks, Schegloff & Jefferson (1974), in which a number of other basic observations are outlined, and a 'simplest systematics' of turn taking is presented. (For a good reviews of approaches to turn and turn taking, see Goodwin 1981: 15-25; and Oreström 1983: 18-38.) Despite the formal simplicity of the proposed turn-taking rules, determining the turn structure of conversation is not at all straightforward. Goodwin (1981: 19-20) lists three reasons why:

- The boundaries of a turn are 'mutable'.
- Their locations are an issue for the participants (not just the analyst).
- 'Turn' cannot be defined independently of the exchange process.
In other words, from an outside vantage point, turns are problematic units: a definitive analysis does not exist. Furthermore, the structural (i.e. syntactic, lexical, prosodic) cues that are often appealed to are highly context dependent and may themselves be reevaluated by the participants in light of subsequent talk. What remains, then, is a situation-specific interpretative process which certainly may make use of such cues, but cannot be defined in terms of them. I will describe the process below, using an excerpt from a conversation with three participants (Henrika, Ingo, and Günther), given in (2) below.

The first step in an analysis of the turn structure must be a determination of each speaker's contribution to the conversation (cf. Du Bois et al. 1992: 92). In the second column of (2), the speaker of every line of text is signified by the speaker's initial, the only exception being that in line 75, X is used for a line which cannot be attributed to any one speaker.

As a second step, it is useful to try to characterize stretches of the interaction according to 'who has the floor' (cf. Edelsky 1981). The concept of 'floor' can be taken as a metaphor for the rights and responsibilities claimed by and awarded to the main speaker in a conversation. Because '[o]verwhelmingly, one party talks at a time' (Sacks et al. 1974: 700), it usually makes sense to think in terms of a 'main' speaker (vs. other participants, who may be 'secondary' speakers) at any given time, although it must be acknowledged that in some places a single main speaker may not

---

(2) Corpus excerpt from ‘Getting by’ [Henrika’s visit: 61-84].

61 H: .. und die is’ au=ch =vollkommen’ umgestellt word’n \•
  and she is also completely switched become

62 H: (H) [ .. auf --
  to

63 I: [ aber Henrika /+
  but H

64 I: .. --ich meine /+
  I mean

65 I: das mag ’n guter [ Arzt sein ] /+
  that may a good doctor be

66 H: [ @(Hx) ]

67 I: und [ Körner mög’n auch ] gesund sein \+
  and grains may also healthy be

68 H: [ (H) ]

69 I: --aber% --
  but

70 I: ... [ (H) willst ] <A du mir doch nicht A> im [ Ernst sa-gen ] \+
  want you to me certainly not in the seriousness say

71 H: [ mit dem% .. <mh> ] --
  with the

72 G: [ @(Hx) ]

61 H: and she completely switched over.

62 H: [to --

63 I: [but Henrika,

64 I: I mean,

65 I: he may be a good doctor,

66 H: (LAUGHS)

67 I: and whole grains may be healthy;

68 H: (INHALES)

69 I: but --

70 I: [you don’t seriously mean to tell me;

71 H: [with the --

72 G: (LAUGHS)
oder [z so das <si=nd> .. «F Körner F» ]/+ 
or thus that are grains

[ «F <das is'> P> eigentlich wölt @@ ] • 
that is actually wide

@= «.8»

[ «@ durch Körner @» /+
through grains

[ @= ((CONTINUES OVER NEXT 3 IUS))

diese Lėbenssitätuation \+
this life.situation

[.. wirklich auf ] die Rēihe kriegst \+
really on the row get

[ <X.. XX >]

.. [z ~das ist doch ] absurd •
that is certainly absurd

[ ja ja ja ] •
yes yes yes

[ z also ] •
so

[ P das weiß ich ] auch nicht P •
that know I also not

[or something that grains,
[that's pretty extreme.
(LAUGHS)

[by whole grains,
(LAUGHS)

this situation;
(you can) really get back in order;
(XXX)

[that's absurd.

[yeah yeah yeah.

[so.

[I don't know. 

38
be evident. In the conversation in question, Henrika has been speaking prior to line 61, with only an occasional *mhmm* from Ingo. Thus, up until line 63, she is the main speaker. At line 63, Ingo takes the floor, and until line 84, when Henrika takes the floor, he seems to be the main speaker. Between lines 63-84 both Henrika and Günther make themselves heard as secondary speakers. How do their contributions in this stretch of speech fit into the turn structure of the conversation?

Beyond determining who has the floor, there is a useful distinction to be made between turn and nonturn utterances. The short laugh by Henrika at line 66 overlapping Ingo’s speech, the similar laugh by Günther at line 72, his extended laughter starting at line 77, and Henrika’s three-syllable remark at line 82 can hardly be thought of as turns, since there is no indication that the speaker intended to claim the floor. The use of participant intent in interpreting turn structure will be amplified below, but it can be said that such interpretation must play a central role. Nonturn utterances have been variously described as ‘conventional signals of attention’ (Fries 1952: 53), ‘back-channel messages’ (Yngve 1970: 568), ‘Hörersignale’ (Schwitalla 1976: 86), ‘Rückmeldungspartikeln’ (Henne & Rehbock 1982: 26); ‘continuers’ (Schegloff 1982: 81), ‘reactive tokens’ (Thompson, Clancy, Suzuki & Tao 1993), among others. I will use the

---

29 In some conversations, the concept of ‘main speaker’ may make less sense. In particular, multi-party interaction involving more than three participants presents special concerns; see e.g. Edelsky (1981: 391-6).

30 Cf. Oreström (1983: 23): ‘There is unanimous agreement ... that certain brief, spontaneous reactions from the listener ... signaling continued attention, agreement, and various emotional reactions are not to be classified as turns ...’.
label ‘backchannel’ here, since this is a term that seems to be widely recognized for the type of utterance I wish to delineate. However, it must be stressed that there are many different conceptions of backchannel.

An absolute determination of speaker intention is, of course, impossible—we can not know what exactly is in the head of a speaker at any point in a conversation. Yet, in some sense what matters is not the intention per se, but the verbal actions and reactions of the participants, which can be derived from the record to a large extent. That is, it matters little that a participant (silently) wanted to speak, but (for whatever reason) chose not to: the speaker having the floor will have known nearly as little about this intention as the analyst. Only an expression of intent counts as far as turn taking goes. One obvious manifestation of an intent to claim the floor is a successful speaker change. At the beginning of lines 62-3, for example, both Henrika and Ingo are talking, but by the end of line 63, the floor has shifted from Henrika to Ingo. Ingo’s utterance at line 63 thus counts as (the initiation of) a turn.

Now, turn taking is a cooperative activity—an ‘interactionally managed system’ (Sacks et al. 1974: 725). The expressed intentions of a participant making a move for the floor constitutes just one aspect of a successful speaker change. The main speaker must also be willing to recognize the attempt. If an attempt is made and it is recognized as above, then a smooth transition between speakers takes place. If, however, an attempt is made, but the floor is not relinquished, if but briefly, no speaker change

---

31 Obviously, the record is limited to the auditory channel in the current study.
occurs. Still, in some sense a turn has been effected. I term this kind of turn 'secondary'. Depending on the response of the main speaker, the attempt may be considered blocked; it may be acknowledged, but not recognized; or it may simply be ignored. While there is probably not always a sharp distinction to be drawn among these subtypes, two instances of secondary turns are illustrated in excerpt (2). First, there is an example of an ignored turn at line 71. Henrika attempts to make a contribution to the conversation after Ingo pauses, but is ignored by him. Despite the pause at line 70, Ingo has not relinquished the floor, and he simply continues speaking. Second, there is a blocked turn later in the conversation, at line 74. There, Günther also makes a contribution to the conversation. In this case, he manages to complete a clause, but his attempt is apparently viewed by Ingo as a claim on the floor, and Ingo blocks any further contribution by substantially increasing the volume of his speech when uttering Körner. In both situations, Ingo continues to speak, so there is no speaker change. But at the same time, the attempts by Henrika and Günther constitute (potentially) substantive contributions to the conversation, and thus must be viewed as turns.\(^{32}\) Whether the turn was completed or not should be another question. In Henrika’s case (line 71) it clearly was not. But for Günther (line 74), although he is effectively shut out from speaking further by Ingo’s prosodic shift, what he did utter may be all he intended to say.

\(^{32}\) Not all overlapping speech is necessarily floor threatening. Cooperative secondary turns are common among some participants; see Tannen (1983; 1989) and Ford & Thompson (1993).
Given the importance of turn structure to the interpretation of conversational discourse, a system for its direct representation in the transcription would constitute a positive step towards integrating interactional aspects of speech into linguistic analyses. As each line of text is already associated (if possible) to a participant through the use of the speaker’s first initial, it is sensible to adopt a turn status convention that is tied to this label.

There are two distinctions that need to be coded. The first is the difference between turn and nonturn (i.e. backchannel) utterances. Turns are, in principle, indicated by the presence of a speaker label. Backchannels, then, can be easily indicated by attaching any distinguishing symbol to the speaker label (cf. Du Bois et al. 1992: 214); here, I will enclose all speaker labels connected with backchannels in parentheses, suggesting symbolically that these utterances should be ignored in the calculation of turns. The second distinction to be made is the difference between turn-initial and turn-medial utterances. Turn-medial utterances may be distinguished from turn-initial ones by simply not repeating the speaker label for each line that is attributed to the same speaker. That is, unmarked lines are to be attributed to the current speaker. Thus, the first few lines of (2) above would be as given in (2a) below.

According to the above-stated convention, line 62 is be interpreted as belonging to Henrika, and lines 64-5 and 67 to Ingo. Note that with all nonturns specially marked, the basic convention applies even in the case of an intervening backchannel (line 66). The line following line 66, which
carries no speaker label, should automatically be attributed to the current speaker, here Ingo.

(2a) 61  H: .. und die is' äu=ch =vøllkomm'n ùmgestellt word'n \•
           and she is also completely switched become

62 (H) [ .. ùuf --
           to

63 I: [ aber Henrika /+
          but H

64 .. ~ich mèine /+
    I mean

65 das màg 'n guter [ Ärzt sein ] /+
    that may a good doctor be

66 (H): [ @(Hx) ]

67 und [ z Kùrner mög'n auch ] gesùnd sein \+
    and grains may also healthy be

This standard representation breaks down in two situations, however, so that a systematic indication of speaker turns in the transcription requires an additional convention. The first case involves transcript and excerpt first lines. Consider (2b), also from (2) above:

(2b) 69 I: ~aber% --
            but

70 .. [(h)willst]\à du mir doch nicht\à im [ ërst sa=gen ] \+
     want you to. me certainly not in. the seriousness say

71 H: [ mit dem% .. <mh> ] --
     with the

72 (G): [ z @(Hx) ]

73 I: oder [ z so das <sl=nd> .. <F Kùrner F> ] /+
     or thus that are grains

Normally, the first line in a transcript or an excerpt must have a speaker label, as in (2a-b). By placing speaker labels at lines 61 in (2a) and at 69 in (2b), the implication is that new turns begin with those utterances. This is
clearly not true in line 69—Ingo’s utterance aber continues his turn from line 67. It is also not true in line 61. As mentioned, Henrika has been the main speaker prior to the excerpt, so her utterance in line 61 constitutes a continuation of her turn in the conversation as well. The second problematic case involves secondary turns. At line 71, Henrika’s utterance constitutes a secondary turn, as explained above. At line 73, it is necessary to repeat the initial I to indicate that this utterance is to be attributed to Ingo, not Henrika. Ingo, though, has not yet given up the floor—oder so das sind Körner does not begin a new turn. To cover these two situations, the turn-medial status of utterances requiring a speaker label is indicated by enclosing the label in square brackets. Thus, excerpt (2) should be as given in (2c), where the turn structure is fully indicated using this ‘speaker index notation’:

(2c) Corpus excerpt from (2) with turn structure indicated.

61 [H]: .. und die is’ äu=ch =völlkomm’n ümgestellt word’n •
62 (H): .. auf --
63 I: [ aber Henrika /+
64 .. ~ich mèine /+
65 das màg ’n guter [ Ärzt sein ] /+
66 (H): [ @(Hx) ]
67 und [2 Körner mög’n auch ] gesùnd sein \+
68 (H): [2 (H) ]
69 ~aber% --
70 .. [ (H)willst ] a du mir doch nicht A im [ Ërnstsa=gen ] /+
71 H: [ mit dem% .. <mh> ] --
72 (G): [2 @(Hx) ]
73 [I]: oder [3 so das <sì=nd> .. «F Körner f» ] /+
74 G: [3 «F <das is’> f» èigentlich wèit @@ ] •
Given the two parameters discussed above—attempt by a participant and recognition by the main speaker—there are then four possible turn statuses for conversational utterances, as summarized in Table 1.5.

Three of the configurations have already been discussed. The fourth configuration—no attempt made by the listener, but recognition granted by the speaker—does not occur in (2). While it completes the paradigm, it may seem to describe an unusual situation. Two reasonable interpretations, however, are possible. Under the rules governing turn construction proposed by Sacks et al. (1974), one mechanism is ‘current speaker selects next’ (704). That is, the main speaker may actively encourage a speaker

<table>
<thead>
<tr>
<th>Attempt by other participant</th>
<th>Recognition by main speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+ Turn ‘Speaker change’</td>
</tr>
<tr>
<td>-</td>
<td>- Turn ‘Secondary’</td>
</tr>
<tr>
<td>+</td>
<td>+ Turn ‘Unintended’</td>
</tr>
<tr>
<td>-</td>
<td>- Non-turn ‘Backchannel’</td>
</tr>
</tbody>
</table>

Table 1.5. Turn status of conversational utterances.
change at the next 'transition relevance place' (i.e. a point in the conversation where a speaker change may take place) by signaling (linguistically or otherwise) that a listener should speak. Such a situation seems to describe a request for information by the speaker, or for active participation in the discourse. A second interpretation involves the speaker granting the floor to a participant who has uttered what for the participant constituted a backchannel response. That is, no attempt to gain the floor was intended, but the speaker misinterpreted the intent of the response. This is, perhaps, a more unusual situation, but it apparently does occur. A plausible example is given in (3).

(3) Corpus excerpt from 'Best kid' [Afternoon tea: 92-99].

92 [E]: .. also total [ hûbsch ] \+
so totally pretty
93 (S): [ @ ]
94 .. @ [2 @ @ @]
95 S: [2 @ ècht @] / •
really
96 E: .. [3 jà \ •
yes
97 (S): [3 @
98 S: .. <L @ sùß @>L \ •
sweet
99 E: .. P total klàsse .. echt P \ •
completely classy really

92 [E]: so really great;
93 (S): (LAUGHS)
94 [2 (LAUGHS)
95 S: [2 really?
96 E: [3 yeah.
97 (S): [3 (LAUGHS)
98 S: nice.
99 E: really great.
In this excerpt, Elena has been the main speaker. At line 95 Sabine utters the expression *echt* 'really' with a high rise while Elena is chuckling. This use of *echt* could have been interpreted by Elena as a backchannel expressing Sabine's surprise or wonderment. But Elena treats it as a request for confirmation along the lines of 'is that really so?', as is evidenced by her response *ja* 'yes'. In other words, it may not have been the intention of Sabine to make a claim on the floor at line 95, but Elena, as the main speaker, chose to recognize the utterance as a turn, effectively transforming it into one.

In sum, using two interactional parameters—one oriented to the main speaker, one to the secondary speaker—it is possible to distinguish between (successful) speaker changes, nonturn utterances (backchannels), and two other types of turns. The basic turn structure resulting from an analysis that makes these distinctions can be represented in the transcription through the modification of the speaker initials used to identify the speech of the participants.

1.6 Turns in the primary corpus

Using the procedure outlined in the previous section, the contributions of each participant in the conversational discourse of the primary corpus were classified as 'turns' or 'backchannels'. The total number of turns and backchannels for each participant is given in Table 1.6 below. The most significant observation to make from this tabulation is that the speech contained in the primary corpus is highly interactional, despite the focus on stories and story-like portions of the conversations. For the three
<table>
<thead>
<tr>
<th>Speaker</th>
<th>Tums</th>
<th>Backchannels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna</td>
<td>67</td>
<td>18</td>
</tr>
<tr>
<td>Uli</td>
<td>65</td>
<td>81</td>
</tr>
<tr>
<td>Tamina</td>
<td>40</td>
<td>18</td>
</tr>
<tr>
<td>Käthi</td>
<td>32</td>
<td>16</td>
</tr>
<tr>
<td>Elena</td>
<td>26</td>
<td>10</td>
</tr>
<tr>
<td>Sabine</td>
<td>24</td>
<td>40</td>
</tr>
<tr>
<td>Marlies</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Henrik</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>Diana</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>Ingo</td>
<td>9</td>
<td>24</td>
</tr>
<tr>
<td>Günther</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>Jürg</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Wolf</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>X</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>(Z1)</td>
<td>(8)</td>
<td>(16)</td>
</tr>
<tr>
<td>(Z2)</td>
<td>(6)</td>
<td>(9)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>326</strong></td>
<td><strong>285</strong></td>
</tr>
</tbody>
</table>

Table 1.6. Frequency of turns and backchannels by speaker. Speakers listed in boldface are tellers of one or more stories (see Table 1.1).

speech events most heavily represented in the corpus, each participant (i.e. Anna vs. Uli; Tamina vs. Käthi; Elena vs. Sabine) contributes a comparable number of turns.\textsuperscript{33} That is, even though the excerpts selected speakers, rather than being random slices of the conversations, they are by were chosen in large part for their extended contributions by individual no means monologic. In addition, although no analysis in the following

\textsuperscript{33} However, the number of backchannels are only comparable for Tamina vs. Käthi. Unfortunately, this is not the place to go into the reasons for the interactional asymmetry between Anna vs. Uli, and Elena vs. Sabine.
chapters depends on the distribution of turns or backchannels among the participants, the proportion of each speaker's utterances to the corpus as a whole has been constantly kept in mind to guard against making generalizations on the basis of the speech of one (or two) individuals.

1.7 Summary

I began this chapter on the study of discourse by discussing properties of discourse, characterizing the discourse analyzed in this study as natural, unscripted, spontaneous, connected speech. I then discussed corpus-based linguistic analysis. In the second section the scope of previous studies of German discourse was examined. I stated my goals in the third section. My primary aim is a comparison of the basic prosodic and syntactic structures in conversational German in relation to the introduction of new referents. I pointed out that a prosodic, syntactic, and discourse-pragmatic analysis of conversation necessitated a substantial examination of basic methodological issues, and I outlined several fundamental questions to be asked for each dimension. In the next section I discussed the constitution and purpose of my two corpora, a body of extended selections from six ordinary conversations, and a set of elicited narratives. Issues relating to the representation of the speech were discussed next. I argued for certain deviations from standard orthography, notably in the transcription of phonologically reduced forms. In the fifth section I presented my interpretation of the turn structure of discourse and its representation in a discourse transcription. Finally, the number of turns and backchannels for each speaker in the conversational corpus was given.
2. Discourse Prosody

Der Begriff Takt ist sicher etwas, womit sich die Phonetik der Zukunft viel zu beschäftigen haben wird; aber bisher sind die Untersuchungen darüber noch kaum begonnen und eine brauchbare Definition wohl noch nicht aufgestellt.

The notion of rhythmic unit [lit., measure] is no doubt a topic that will much occupy the attention of future phoneticians; up til now, however, investigations of it have hardly begun, and a viable definition has not yet been proposed.

Otto Jespersen
Lehrbuch der Phonetik, 1904, p.203

In this chapter I examine the prosodic organization of the primary corpus of conversational stories in terms of phrasal units, based on a systematic auditory analysis. My approach is production-oriented in the sense that I focus on the temporal development of the stories primarily from the speaker’s point of view. However, as the analyst’s perspective is much like that of the listener, certain references to the perception of speech seem unavoidable.1 Terminological prerequisites are attended to at the outset, as the subfield of prosody is rife with competing definitions of basic concepts. The rationale for focusing on prosodic units as a first step in the analysis of spoken discourse is explained. Segmenting spoken language is discussed in terms of language type and prosody source. In the second section various

1 Not that there is any special need to avoid reference to the listener’s perspective. Speech production cannot be divorced from perception entirely, since ‘... production is sensitive to the perceptual needs of the listener as estimated by the speaker’ (Nootseboom 1991: 232; cf. also Nootseboom 1985).
approaches to segmentation are surveyed, and some of the principal phrasal units that have been proposed in the literature are reviewed. The 'intonation unit', which I use as my primary analytic unit in the current analysis, is introduced. In the next section features which play a role in the identification of intonation unit boundaries are examined, and their representation in a discourse transcription is discussed. In section four I outline the prosodic structure of the stories in terms of intonation units, and discuss the fundamental role of these units in explicating questions of discourse production. I then briefly explore the idea of treating the intonation unit as a prototype category. The nagging questions of competing analyses as well as the reliability and validity of the present auditory analysis is thereby addressed.

2.1 Prosody, intonation, and discourse

What is meant by 'discourse prosody'? My use of the term 'discourse' was clarified in the previous chapter. Remaining to be spelled out is the distinction between 'prosody' and 'intonation' drawn here, and the significance I attach to the notion of 'discourse prosody' itself. In the following discussion, reference will be made to works on English, as well as German, prosody, as virtually all theoretical advances in the description of German have had their origin in accounts of English (Girouard 1984: 38). Moreover, concrete descriptive knowledge about German prosody—as opposed to prescriptive generalizations mainly for pedagogical purposes—is in many respects sketchy compared with other major European languages (Klein 1980; Bannert 1983b: 9-10). While this situation may be somewhat disturbing, it
is also true that ‘German prosody closely resembles that of other West Germanic languages’ (Gibbon 1987: 2), so that insights about English may be fruitfully applied to German.

Researchers generally draw a distinction between prosody and intonation. On the one hand, ‘prosody’ may be used to refer to all meaningful auditory or acoustic aspects of the speech signal that are primarily nonsegmental in nature, while ‘intonation’ may serve as a cover term for a subset of those aspects which are systematically accounted for in a particular intonational model.2 ‘Prosody’ is thus the broader term, but the individual ‘suprasegmental’ features encompassed by the term ‘intonation’ will vary according to the specification of the system, as determined by language, data type, and descriptive methodology. For instance, Stock & Zacharias (1973: v), who have little need to develop a more differentiated account, take a very liberal position: ‘Unter dem Begriff Intonation fassen wir alle phonetischen Mittel zusammen, die neben den Lauten bzw. Lautketten

---

2 This is a useful simplification: ‘... segmental and suprasegmental structure are not cleanly separated by a single slice of the experimental knife. Suprasegmental features or properties are not merely overlaid on segmental organization; rather they are intertwined with it’ (Kent 1983: 58). However, from a phonological viewpoint, virtually any phonetic feature can be represented independently of individual segments. Depending on the framework (and degree of abstraction), features such as [nasal] and [voice] may be nonsegmental (cf. Robins 1973 [1957]: 274). Thus, in some descriptions it may be necessary to distinguish ‘prosodic’ further. By ‘meaningful’ I intend to encompass any speaker choice used to convey meaning. This is considerably more generous than the usual position, as stated by Fry (1968: 365): ‘Only those distinctions which have linguistic relevance are classed as prosodic features in a particular language’.
beim Zustandekommen eines Kommunikationseffektes ... Funktionen zu erfüllen haben' ('we take the concept intonation to include all intentional nonsegmental phonetic means which have communicative function'). In contrast, Isačenko & Schädlich (1966) choose a parsimonious description in their model of German intonation, based on perception tests using synthesized speech. They make use only of pitch accents ('Tonbrüche'), which include a directional component and a timing feature with respect to word stress.

On the other hand, 'intonation' may be said to encompass all aspects of the speech signal which involve fundamental frequency/pitch modulation. For instance, 't Hart, Collier & Cohen (1990: 2) define intonation as 'the ensemble of pitch variations in speech caused by the varying periodicity in the vibration of the vocal cords'. This definition includes not only the use of pitch to mark specific syllables in an utterance as especially prominent (the 'nucleas', 'tonic', 'Satzakzent'), but also aspects of the pitch contour which can be described as 'global'. Over the past dozen years or so, there has been some agreement—especially among researchers who make use of acoustic data—that this is an appropriate way to distinguish intonation from prosody. Since the analysis presented in this chapter makes reference to much more than pitch variation, I will to speak here of 'discourse prosody', instead of 'discourse intonation'.

---

3 Later (p. 6), it becomes clear that they mean principally four features: pitch (Sprechmelodie), loudness (Lautheit), duration (Sprechtempo), and timbre (Klangfarbe).

4 Intonation as distinct from prosody, as well as from the lexical use of tone, is sometimes specified as 'intonation in the narrow sense', e.g. Ehlich (1981).
The specific aspects of the sound continuum that are generally considered by all researchers to be prosodic include the following auditory parameters: ‘pitch’ (acoustically, fundamental frequency or $F_0$), ‘loudness’ (intensity), and ‘duration’ (time) (Lehiste 1970; Ladd & Cutler 1983: 1; Couper-Kuhlen 1986: 8; Johns-Lewis 1986b: xix). These parameters, however, are not necessarily unitary and their relevance in prosodic descriptions varies considerably. Crystal (1969: 140), for example, distinguishes two separate aspects of pitch—‘pitch direction’ (or ‘tone’) and ‘pitch range’. Also, two different aspects of duration—‘tempo’ and ‘rhythmicality’—play a role in his model. Additional phonetic features, of course, turn up in individual analyses, and some features may not be overtly represented. Again consulting Crystal (1969), we find that he includes ‘pause’ as a prosodic feature, in addition to ‘loudness’ and the four pitch and duration features. A seventh feature, ‘tension’ i.e. articulatory (muscular) effort, is held to be marginal; it straddles the border between prosodic and non-prosodic. On the other side of the spectrum, Bannert (1983b) minimally augments his model of German intonation (1983a) with a temporal component (inherent and context-dependent duration) to obtain a prosodic model.

The problem of deciding which phonetic aspects of the speech signal are relevant for a prosodic analysis is often phrased in terms of a ‘linguistic’–‘paralinguistic’ opposition (sometimes together with an ‘extralinguistic’ category as well), with paralinguistic (and extralinguistic) features being nonprosodic. However, it is obvious that such neat distinctions embody a certain arbitrariness, as extralinguistic functions may be ascribed to virtu-
ally any prosodic feature (cf. Crystal 1975: 95), and certain linguistic functions can be identified for so-called nonprosodic features. A case in point is the feature ‘tension’ mentioned above, where a strict division is simply impossible to maintain. The inherent gradient nature of most features allows for a categorial assignment associated with the choice of linguistic form, as well as for the interpretation of variation within the range of any one value for paralinguistic function. (Presumably speaker variation attributable to purely idiosyncratic sources is perceptually filtered, though this clearly has indexical value of a social or interpersonal nature; cf. Voice quality below.) The possibilities will become more apparent in the ensuing discussion, when we will see that certain so-called paralinguistic aspects of speech play at least an ancillary role in a prosodic analysis of discourse, so that reference to prosodic versus nonprosodic features does not seem to be particularly meaningful here. ‘Discourse prosody’ as conceived in this chapter, then, involves much more than attending to the one or two phonetic features typically granted special treatment in standard intonational analyses. In this connection, however, it is important to keep in mind that the aim of this chapter is not to develop a complete model of prosody in German (or other) discourse, but to construct a methodology which provides reliable and systematic access to the prosody for the purposes outlined above. As discussed earlier, any linguistic analysis incorporating spoken language can hardly be viewed as adequate without taking the prosody into consideration. The prosodic analysis of the current chapter is thus not to be taken simply as an end in itself, even though detailed phonetic anchoring of phonological entities is often lacking. More importantly, this
analysis forms the foundation for the grammatical and discourse-pragmatic analyses in subsequent chapters. This orientation is significant because in addition to the traditionally nonprosodic features that I will include in my prosodic analysis, there are some phonetic features usually considered prosodic which will receive proportionally little attention here.

In light of the goal of presenting a prosodic analysis of spontaneous speech which reflects the dynamic, process-oriented nature of discourse, it will be useful to concentrate on what may be in some sense basic units of production. Much of what has been written concerning language production is fairly speculative. Accordingly, a couple of cautionary remarks regarding production units are in order before embarking on a survey of the data. Several researchers have warned that searching for one particular unit to capture all aspects of language production is probably misguided. Levelt (1989: 23) points out that ‘[d]ifferent processing components have their own characteristic processing units, and these units may or may not be preserved in the articulatory pattern of speech’. Of course, the extreme complexity of speech makes this view quite plausible. Also plausible is the view that the degree of forward planning during speaking may vary somewhat as the speech conditions change (Beattie 1980), so that there can be ‘no single fundamental unit of encoding’ (142). On the other hand, it is equally plausible that much of the prosodic regularity in discourse has cognitive significance, and that explicating this regularity is the job of linguists working with discourse data. In any case, a certain abstraction from the raw data is necessary. Abercrombie’s (1973 [1954]: 2) remarks about this in transcribing dialect material are valid for any form of speech:
The real raw material, the utterances, cannot of course be handled directly, because they are unique events and they are complex events. Before we can say anything about them, we must be able to treat them as made up of constituent elements, each element being representative of a class.

Although Abercrombie is referring to segments here, an identical situation holds for levels higher up a phonological hierarchy. Establishing some basic unit which is derived from (and thus sensitive to) the temporal production of discourse is the minimum possible abstraction necessary in order to get a handle on natural, connected speech.

As I will illustrate in §2.3, the phonetic regularities inherently present in spoken language provide sufficient clues as to what kind of unit(s) should be established for the analysis of natural discourse. In general terms, this regularity includes recurrent patterns of prosodic features which help structure the discourse in various ways, whether they have a cognitive basis or not. In other words, the kind of unit that is of interest here is less the highly abstract mental processing unit for any specific component (or for a production model as a whole), and more the unit reflecting the local organization of connected speech which is based on some of the perceived properties of the acoustic signal. Of course, it is to be hoped that any observation about the surface phonetic structure made here will be consistent with realistic models of language processing, but it is not my intent to propose such a model. In any case, I believe it is possible to consider the form and function of units of speech as such. Certainly, from a surface production point of view, connected spontaneous speech is not ut-
tered in a uniformly continuous stream, but in distinct and identifiable quanta, a point which I will return to shortly. It is in this regard that it makes sense to inquire into the nature of production units and their role in the structuring of discourse. Thus, a central concern of ‘discourse prosody’ as represented here will be to describe the prosodic phrasing of natural, connected speech. A proper description entails an examination of the evidence motivating this phrasing, including the internal coherence of the resulting prosodic units, as well as phonetic cues to their boundaries. It may well be the case that not every unit is equally coherent, or that not every boundary percept has a clearly identifiable phonetic basis—given our current limited understanding of prosodic phenomena. And it cannot be expected that phonetic cues, when identified, correspond in any unique way to phrasal units. Nevertheless, it is reasonable to start with the hypothesis that every unit can be explained in terms of observable phonetic features. As Lehiste has repeatedly warned, ‘phonologists ignore phonetics at their own peril’ (1984: 96).

The view advocated here contrasts with approaches in which the phrasing often functions more as a theoretical construct to aid in the unity of a specific model. For example, in discussing the place of the phrasal unit ‘tone group’ in a hierarchy of phonological units, Pheby (1975: 50) makes it clear that tone group boundaries as he conceives them are not always rooted in phonetic substance:

Im Gegensatz zu den anderen Einheiten [Takt, Silbe, Phonem]
lässt sich die Tongruppengrenze auf phonetischer Grundlage
nenicht eindeutig festlegen. Sie kann durch eine Pause, einen aus-
gedehnten ("überlangen") Takt ... oder den Anfang eines neuen Vorlaufmusters ... signalisiert werden. Wenn diese Verhältnisse nicht bestehen, läßt sich die Tongruppengrenze bei fließendem Sprechen nicht aus dem Signal ableiten. ... Es ist in Ermangelung empirischer Evidenz notwendig, eine theoretische Entscheidung zu treffen ....

(In contrast to the other units [stress group, syllable, phoneme], a tone group boundary cannot be unambiguously determined on a phonetic basis. It may be signaled by a pause, a lengthened stress group, or the configuration of pitch at the start of a new group. When these conditions do not obtain, a tone group boundary cannot be derived from the signal in fluent speech. Lacking empirical evidence, it is necessary to make a theoretical decision.)

He goes on to say that one may make the assumption either that stress group and tone group boundaries always align, or that tone group boundaries align with syntactic structure. The decision is supposed to be arbitrary, from the point of view of the prosody.

Others have taken a more cautious position with regard to the value of phonetic features, while at the same time maintaining a phonological account of prosody. Ladd (1986: 314) states the basic conflict between adequate phonetic observation and phonological consistency in determining the phrasing of an utterance as follows:

[Intonational] P[hrase]s are supposed to be delimited by boundary phenomena of some intuitively definable sort, but they are also supposed to have well-defined internal phonological structure
and to match up with the syntax in well-defined ways. So any stretch of speech set off by audible boundaries is assumed to be an intonational phrase, and at the same time any stretch of speech identifiable on structural grounds as an intonational phrase is assumed to be set off by boundaries.

To address this competition between perspectives, Ladd argues for two levels of intonational phrasing, one which is governed by the location of nuclear tones, and one determined by other phonetic cues. By taking this intermediate position of sorts, he explains (1986: 316):

... we can mark boundaries where we ‘need’ them for a sensible description of the intonational phonology, while reserving a different sort of boundary mark for those boundaries that have conspicuously audible phonetic correlates.

Commendably, Ladd does not dismiss the importance of phonetic boundary signals. However, in retaining purely phonological boundaries for reasons of theoretical convenience, it seems that he forces intonation (in the narrow sense) to be strictly separated from prosody: unlike other suprasegmental features, pitch contours define their own prosodic units. As stated, I prefer to regard intonation as fully integrated into a prosodic model.

Since the notion of ‘unit’ in the fluid medium of spoken language is naturally intertwined with the issue of segmentation—as is amply evident from the above discussion—it will be necessary to consider more generally past approaches to the phrasing of utterances. While some linguists have long been interested in defining prosodic units, questions of segmentation
are often not even raised, with predictably negative consequences. This neglect can be traced to three factors: the abstractness of the linguistic model, the type of the language data used to explicate intonational or prosodic patterns, and the actual source of these prosodic patterns. Reliance on a model, of course, serves many useful purposes. But when a model acquires a certain inertia of its own, typically through increased abstraction, the segmentation process may no longer be applicable to naturally occurring utterances. Instead, model-specific expectations, which have been acquired by examining naturalistic language (e.g. constructed question–answer pairs) or in written forms upon which a prosodic pattern has been projected, will be used to determine prosodic breaks in actual natural discourse. Thus, directly related to the way generalizations about the prosody are expressed is the nature of the language used to describe and model prosody. The main problem here is that assumptions about the simple relationship between natural and naturalistic prosodic patterns are probably not valid. In a recent study, for example, it was found that intonation in discourse ‘... differs from intonation in semantically unrelated [i.e. constructed, context-free] sentences with respect to practically all F0 parameters investigated’ (Huber 1991: 190).

5 Cf. Ladd & Campbell (1991: 290): ‘Obviously, if the definition of e.g. “phrase” is open to debate, then this will affect the way “phrase boundaries” are marked in any given corpus or text sample...’.

6 Even in models containing variables to capture ‘nonlinguistic’ aspects of speech that affect phrasing (e.g. Bierwisch’s (1966) independent parameter p, which is said to vary with ‘speech style’, especially tempo), the underlying mechanisms are based on naturalistic data.
<table>
<thead>
<tr>
<th>Prosody source</th>
<th>Isolated syntactic unit</th>
<th>Short, connected</th>
<th>Text oriented</th>
</tr>
</thead>
</table>

Table 2.1. Types and sources of language data in German intonation studies.
Various characteristics of the 'raw data' which figure in an analysis can be discerned (cf. Gibbon 1988). Table 2.1 above presents a selection of works on German intonation and prosody according to the degree of context the language is embedded in and the source of the prosodic patterns under investigation. In the table, studies making use of so-called context-independent language samples are listed in the left-hand column ('isolated syntactic units'); those investigating language samples for which the natural context is included to the extent possible in an audio recording are listed in the right-hand column ('text oriented'). Some studies make use of an intermediate degree of contextualization by using constructed question-answer pairs or other semantically coherent sequences, or using short extracts from literary works; these are listed in the center column ('short, connected'). The prosodic patterns for any of these language types ('prosody source') may be assigned by the analyst intuitively ('introspection'), obtained by having native speakers read prepared material or written language aloud ('reading'), or derived from recordings of unscripted narrative, conversation, and so forth ('spontaneous speech'), which of course may vary in degree of spontaneity.

From this classification, it is apparent that there is a tendency to examine mainly language which has been restricted contextually and whose prosody has been produced under the influence of written forms. Intonation studies involving spontaneous discourse, while far from unknown, are more the exception.\(^7\) The significance of this for the issue of segmenta-

\(^7\) No works on German prosody were found using short excerpts of spontaneous speech as pri-
tion is that all speech types besides spontaneous discourse are more or less syntactically oriented. The stretches of speech assumed to be relevant for the explication of intonational patterns are thus grammatically, not prosodically, defined or influenced. Often, the simple 'sentence' is, by default, the domain used, and the need to identify large-scoped prosodic boundaries never surfaces. Thus, a serious limitation of many studies of prosody—typically unacknowledged—is the a priori imposition of a syntactic domain. The common assumption of the primacy of syntax can lead to problems of descriptive validity and false generalizations. For example, Schuetze-Coburn, Shapley & Weber (1991) show that taking a phrasal unit to be the usual domain of $F_0$ declination is largely due to the syntactic restrictions placed on the primary data; in natural, connected speech, declination often has a wider scope. Such observations lead one to wonder, as Gibbon (1988: 3) does, what syntax really has to do with prosody (and vice versa).\footnote{Bolinger (1972: 644) is quite emphatic: 'Whether one tries to set up prosodic rules for syntax or syntactic rules for prosody, the result is the same: two domains are confused which should be kept separate.'}

The relationship between prosody and syntax is often obscured, because terminology masks assumptions about the nature of the data, assumptions that researchers may themselves not be fully aware they are making. The presentation of German intonation in the influential work of von Essen (1956) is perhaps typical in this regard. Von Essen claims to describe the in-

\footnote{Bolinger (1972: 644) is quite emphatic: 'Whether one tries to set up prosodic rules for syntax or syntactic rules for prosody, the result is the same: two domains are confused which should be kept separate.'}
tonation of 'Aussprüche' ('utterances'), deriving his system of 'intonational rules' from transcriptions of recorded radio productions and private speech (63). In Part II of his treatise, he does include short excerpts of connected speech illustrating the system's application. However, no specifics are indicated as to the weight given the read texts as opposed to the (presumably) more spontaneous material. Furthermore, in Part I (the systematic outline) virtually every example given is a complete, isolated, grammatical sentence (complete with standard orthography and punctuation), creating the impression of—if not asserting—a one-to-one correspondence between syntax and prosody.

The motivation behind this kind of prosodic analysis often seems to be a desire to elaborate the 'meaning' of intonation in conjunction with grammatical structure. Under this view, intonation tacitly supplements the syntax of an utterance, functioning primarily to 'disambiguate' or to indicate grammatical contrasts not overtly present in the segmental material of the utterance. While linguists usually attribute some limited grammatical function to prosody, all too often this function is overemphasized.9 Wängler's (1974: 203) characterization is not unusual: 'Die Satzintonation

9 Perhaps this is due to the dominant role syntax has played in linguistics since the early sixties. It is interesting to note that while other subareas of linguistics have recently enjoyed an increase in attention, there has also been a renewed focus on the grammatical function of intonation (e.g. Batliner 1988; Goldbeck & Sendtimeier 1988; Oppenrieder 1988). True to form, these studies analyze the intonation of read, isolated sentences. For an alternate view, see Barry (1981: 336-7), who holds the position that differentiating sentence modes does NOT constitute a separate function of prosody, at least in English and German.
ist im Deutschen (a) Mittel zur Differenzierung verschiedener Satzkategorien, (b) emotioneller Ausdruck’ (‘German sentence intonation serves to differentiate sentence types and to express emotions’). The entrenched idea that one intonational pattern is ‘declarative’ and another ‘interrogative’ also stems directly from the concerns of this way of thinking. As Klein (1982) amply demonstrates, such a simplistic view is ‘völlig falsch’ (‘completely wrong’) (291). This is not to say that prosody is not used linguistically, only that there is a tendency to err by attributing more to the prosody directly than is normally necessary during actual speech events.¹⁰

Unfortunately for the discourse analyst, this situation has not changed much in recent years. How linguists think about prosody and the segmentation of running speech is still greatly influenced by analyses of intonation patterns in simple syntactic structures; what they know is still largely the accumulation of results from these studies. Nooteboom (1991: 235) states the situation thus: ‘There is an urgent need for studying the acoustic/

---

¹⁰ The problem is, of course, much more general than this. Naturally, most approaches devote some attention to intonational meaning, but there is no consensus as to its exact nature. For some, definite meanings—whether grammatical or attitudinal—can be ascribed to specific contour classes (cf. Liberman & Sag 1974), while others prefer to think of intonation as providing a set of semantic guidelines on interpretations from context (cf. McLemore 1991). Whatever the case may be, Gibbon’s (1983: 203) assessment of the situation in one particular instance is telling for all semantically based approaches:

‘Ironic’ intonation, for instance, would not stand up to serious definitions of irony: the role of wider than expected pitch bandwidth or slower rhythm and lengthening of accented syllables is simply to provide a prosodic pointer to the locus of an ironical incongruence between a locution and its context.
netic characteristics that distinguish spontaneous from prepared speech'. Furthermore, explicit models of intonation are developed using such structures, rather than first determining the natural domain of prosodic features on the basis of internal phonetic criteria. For instance, Gårding (1983: 11) states that the input to her model is '... a sentence, equipped with markings for lexical accents or tones, accents (tones) at phrase and sentence level, morphological and phrase boundaries, and the mode of sentence intonation'. The model of German intonation outlined in Bannert (1983a), which requires a semantically, syntactically, morphologically, phonologically, and prosodically specified input, is based on an analysis of the fundamental frequency of a set of 'sentences' that come across more as cleverly constructed tongue twisters than anything that would be found in a natural speech setting. Kohler's (1991) model is similarly sentence oriented. As can be expected, the rationale for this methodology seems to be due to simplistic underlying conceptions of language processing. For example, Sorensen & Cooper (1980: 400) explain: 'One of the main assumptions of our research is that a speaker formulates many aspects of the syntactic structure of an utterance before formulating its phonetic representation'. While this assumption is no doubt true to some degree, a leap of faith is made in its application: the primary object of investigation is, as usual, the well-formed syntactically complete 'sentence'. But such models are unlikely to be psychologically valid, even if they can successfully describe acceptable intonational patterns for isolated sentences or sentence strings, simply because speakers do not talk that way. Levelt's (1989) model of language processing is more realistic in this regard. While there is syntactic
input here too, processing is incremental, so that it is necessary to look ahead no more than ‘one or two successive lexical elements’ (411). This kind of model more closely reflects the characteristics of on-line production associated with spontaneous speech. On the other side of the coin, there are the global aspects of prosody. To date it has not been demonstrated that models based on isolated syntactic units can adequately describe all the prosodic patterns that arise in connected speech, especially when unscripted. In sum, Crystal’s (1969: 40) assessment of previous work on intonation still rings true today: ‘There is the hidden assumption that, having done an analytic survey of the basic functional “blocks” of intonation, the synthesis of these blocks into connected utterance is simple’.

Now, many researchers today are fully aware that certain prosodic patterns occur only in longer utterances and that such patterns play an important role, for example, in establishing coherence in connected speech (cf. Brindöpke, Gibbon, Langer & Pampel 1990: 138). Acoustic phenomena have been identified which are organized globally in longer utterance sequences (e.g. overall fundamental frequency declination, Thorsen 1985) or which result from an interaction of successive utterances (e.g. tonal coupling, Bruce 1982). One problem impeding progress in modeling these processes in a way that is convincing to discourse analysts, however, seems to reside in the data traditionally used in experimental work: there is a reliance on data that exhibits closer ties to written language than to speech (especially of an unplanned variety). While both Bruce (1982) and Thorsen (1985) are interested in textual effects of intonation, the ‘texts’ that constitute their primary data are read sequences composed of two or three sen-
tences (whose order may be alternated to provide a 'complete' text paradigm). Whatever the advantages of this methodology may be, we may make more progress heeding the words of Bolinger (1987: 24): 'It's time—where prosody is concerned—to shake ourselves free of print culture'.

In the following two sections, I discuss the rhythm of spoken language as it relates to the segmentation of speech, and I weigh evidence for positioning certain kinds of prosodic units. Various issues specific to prosodic segmentation are raised. The phonetic and nonphonetic features associated with prosodic units and their relevance to the specific unit adopted here—the intonation unit—are examined.

2.2 Prosodic units

Most linguistic research on intonation or prosody in discourse makes use of some basic unit about the size of what for the moment may be referred to generically as a 'prosodic phrase' (cf. also the 'intonation-group' of Cruttenden (1986: 35) as used in his constructed examples). Reference to a unit of speech approximately 1–2 seconds in duration (Chafe 1980b: 14; Altenberg 1987: 22) is useful because spoken language exhibits natural, quasi-rhythmic breaks in production, even when fluently articulated. Moreover, a prosodic unit of this length is important to language understanding. Helfrich (1985: 29) sums up this point as follows:

Der Perzeptionsprozeß scheint zwar einerseits Einheiten zu involvieren, die die Dauer eines Phonems oder einer Silbe wesentlich überschreiten, andererseits scheinen die sich über einen längeren Zeitraum erstreckenden semantisch-syntaktischen Ein-

(The perception process seems to involve units that are much larger than phonemes or syllables. But larger syntactic-semantic units cannot be considered as the basis for a perceptual analysis, because they do not have a close connection to the acoustic signal. It appears necessary to include units that have a longer duration but still stand in direct relation to the acoustic signal. Both criteria seem to be satisfied by units that are defined by the prosody of the utterance.)

Phoneticians have long been aware of this natural grouping, even if they have not always demonstrated its existence convincingly. Schubiger (1958: 9) writes: ‘Human speech is not a shapeless succession of sounds or words. The flow of utterance can be divided into sections, the elements of which are more closely connected with each other than with what precedes or follows’. In the psycholinguistic literature, the regular alternation of hesitant and fluent stretches of speech has been called ‘cognitive rhythm’ (Henderson, Goldman-Eisler & Skarbek 1966; Butterworth & Goldman-Eisler 1979). Linguists have also observed that ‘brief spurts’ of language are an obvious property of spontaneous speech (Chafe 1980b: 13). Significantly,
the resultant units have been associated with numerous functional and interactional properties. Crystal (1969: 204) anoints them ‘... the most readily perceivable, recurrent, maximal functional unit[s] to which linguistic meanings can be attached ...’. Considering their apparent fundamental nature and the attention that has been devoted to explicating their significance, a closer look at the way prosodic phrases have been analyzed seems easily warranted.

The basis for the coherence between groups of words that Schubiger and others refer to may be viewed in various ways, and the properties of the resulting prosodic units vary considerably as a result of this differing emphasis. The main problem is thus ‘how the observable fact that continuous speech is broken up into perceptible blocks or units is to be interpreted’ (Brazil 1985b: 15). Couper-Kuhlen (1986: 73-6), discusses numerous competing physiological, semantic/grammatical, and phonetic/phonological approaches to the prosodic phrase. As an alternate viewpoint, Altenberg (1987: 47) notes that the prosodic phrase may be analyzed from four complementary perspectives: cognitive, textual, prosodic, and grammatical. Below I will examine questions regarding the segmentation of prosodic phrases, addressing these perspectives. Although the most central question for my presentation concerns the grounds on which the segmentation of speech is determined, further questions concerning the systematic organization of the units may be raised. For example, is there a hierarchy of prosodic units? If so, can one level be considered ‘basic’? Are units linearly sequenced? While these and other questions cannot be answered definitely in the short space of this chapter, I will provide some motivation for the
particular choices made here.

2.2.1 The basis of prosodic segmentation

Above I mentioned the implicit syntactic orientation of the majority of intonation studies. In addition, many studies are explicitly syntactic, in that they assign a prosodic structure to a string of words directly from the syntax, rather than treating intonation as structurally independent, or only incidentally related. For instance, Bierwisch (1966)—who takes von Essen’s (1956) analysis as given—presents a set of rules for deriving intonation patterns in German within the domain of a prosodic ‘phrasal unit’ (‘Phrasierungseinheit’, or ‘PE’). The prosodic phrasing is, however, in effect syntactically based: ‘Grundlage für die Phrasierung sind die syntaktische Oberflächenstruktur und die davon abhängigen Akzentverhältnisse’ (‘The syntactic surface structure and the accent relationships, which are dependent on this structure, form the basis for the phrasing’) (127). Semantic aspects are deliberately given no weight, and contextual influence is kept to a bare minimum. Furthermore, as PEs are considered to be abstract linguistic units, it is expressly stated that no attempt is made to define them in phonetic terms.

Similarly, in many current phonological models, the prosodic phrasing of an utterance is normally determined by (although, it is usually emphasized, is not equivalent to) its syntactic structure.\footnote{While on the face of it phrasing (at the level under discussion) might be basically ‘free’ with respect to syntax (Selkirk 1981: 134), the constituent elements of this phrasing ‘refer di-}
prosodic phrases may be considered obligatory for displaced syntactic constituents, parentheticals, and nonrestrictive relative clauses, or at certain noun phrase and clause boundaries, as assumed in Pierrehumbert (1980: 20) or formalized in Nespor & Vogel's (1983: 125) rules for constructing prosodic categories. The assumption that syntax is prior and prosody is derivative can again be traced to the practice of 'obtaining' spoken language from the written, as is the case in reading aloud (most acoustic/instrumental studies) or in projecting an intonational pattern onto a (written) segmental string (the armchair approach). While using this type of data may perhaps be appropriate for some kind of studies, the prosodic analysis which results is too homogenous (abstract) to be applicable in a description of discourse prosody.

An interesting variation on the theme of appealing to the syntax of an utterance in order to identify the prosodic phrasing is found in Crystal (1975). One of the empirical claims put forth in his model of nonsegmental phonology is that tone-unit boundaries are actually determined by syntactic structure (15). Analyzing a corpus of spontaneous conversational speech

rectly to the syntax of the sentence' (126). Cf. also Pierrehumbert (1980: 19): '[How an utterance is broken up into intonation phrases] is really a problem in the relation of syntax and semantics'.

12 The problems encountered in identifying some of these structures in natural discourse, e.g. nonrestrictive relative clauses (cf. Fox & Thompson 1990: 297-8), are, of course, not discussed.

13 Other sources of syntax dominating prosody can be found. Halliday's (1967b) classification of tone units as 'marked' or 'unmarked' is also purely syntactic, being based on the correspondence of tone units with clause structure. However, this is a somewhat milder commingling than determining the prosodic boundaries on the basis of the constituent structure.
for which tone-unit divisions were previously indicated on phonetic
grounds, Crystal devises a set of post-hoc rules for boundary placement on
the basis of syntactic composition. Using these rules, over 99% of the
prosodic boundaries in his corpus are 'predictable'. This result, however, is
hard to reconcile with the observations of other researchers. In applying
Crystal's rules to a different kind of speech (prepared monolog), Altenberg
(1987) arrives at a lower success rate (which varies with grammatical cate-
gory). It is not so much that prosodic boundaries are really predictable in
any theoretical way, but that, in Altenberg's words: "... since rule failure
seldom produces quite unacceptable results, the system can be described as
surprisingly efficient" (195).\textsuperscript{14} Still, he does not hold that grammatical
structure actually determines prosodic structure, but merely correlates with
it sufficiently for practical use: '... speech segmentation, though basically a
cognitive and textual process, is most easily predicted in grammatical
terms' (48). In stark contrast, Jin (1990: 7-8)—arguing from French and
German task-oriented dialog—unequivocally considers syntactic units ill-
suited as the basis for establishing prosodic domains in spontaneous
speech.\textsuperscript{15} It is fair to say that the exact relationship between syntax and
prosody remains controversial. In any case, this issue will be addressed
more thoroughly in Chapter 3, when the prosodic and syntactic structures
of the current corpus are contrasted.

\textsuperscript{14} Note that the purpose of Altenberg's study is to facilitate automatic text-to-speech con-
version, i.e. to improve synthesized speech which is computer generated from written texts.

\textsuperscript{15} Of course, the type of syntactic boundary one means is crucial to the argument; see §3.1.
Other accounts of intonation, especially earlier ones, make use of semantic coherence, rather than syntactic constituency. Basic prosodic units in these studies are said to comprise 'sense groups' (Armstrong & Ward 1926; Kingdon 1958; Delattre, Poenack & Olsen 1965). The appeal to semantic aspects is usually rather general, however. Klinghardt & Klemm (1920: 32) explain the pairing of semantics and prosody in the following way:

Die einteilung eines satzes in sprechakte geschieht im interesse seiner sinnklarheit. Um diese zu fördern, werden die worte des sprechaktts gerade so und nicht anders ausgewählt und gruppiert, es sind SINNTAKTE.

(The division of a sentence into speech groups is for the sake of clarity of meaning. Accordingly, a sequence of words is selected and grouped just so, i.e. into SENSE GROUPS, not in any other way.)

The problem here, of course, is that ‘sense group’ is difficult—if not impossible—to define adequately. Interestingly, Klinghardt & Klemm make it clear that the division they are talking about is actually a function of the prosody:

Was ist es denn aber, das ... den sprechakt deutlich zu gehör bringt? Die antwort kann m.e. nur lauten: die tonbewegung, die intonation des schlusses eines jeden sprechaktts. Darum ... müssen [wir] vom INTONATORISCHEN sinntakte als der grundlage aller takteinteilung unserer rede sprechen (33).

(But what is it that makes speech groups perceivable? In my opinion, the answer can only be: the pitch, the intonation at the end of every group. Therefore, we should speak of the INTO-
NATIONAL sense group as the foundation of all segmentation of our speech.)

Thus, although Klinghardt & Klemm employ the term 'sense group', they actually mean a unit defined in terms of prosodic structure, suggesting that the term is best considered '... a misnomer for an intuitively perceived tone cum rhythm group' (Trim 1973 [1959]: 321). The reference to semantics may also be understood as an attempt to interpret the significance of a basic level of prosodic organization with respect to another structural level. However, the assertion of a correlation between semantic and prosodic structure requires the outlining of clear conceptions of such structure and a detailed empirical evaluation, which is still lacking.

A set of approaches to prosodic phrasing which may be compared with each other is presented in the works of Halliday and Chafe, both of whom unite prosodic and pragmatic dimensions of language. Halliday (1967b), taking a textual orientation, links the intonational system of English directly to the informational plane. His largest prosodic units, 'tone groups', are said to organize texts into 'information units'. Expressing an overtly cognitive orientation, Chafe (1980b) calls the prosodic units which result from his analysis of narrative monolog 'idea units'. The basic problem in both cases is indicated by the terminology: the assertion of equivalence seems to make one dimension superfluous, or at least subordinate.16 But

---

16 In later work, Chafe (1987; 1993a) adopts a prosodically oriented term, thus pulling back from the assumption of equivalence inherent in a cognitive/semantic label. Still, his starting point is information flow: 'When a speaker is speaking, [he] or she verbalizes one piece of temporarily active information after another. Each such piece is expressed in what I will
while the hypothesis that information structure is expressed primarily by prosodic structure may be true (at least in English), as we shall see in §4.3.1, the mapping of informational structure onto prosodic form is far from isomorphic, even when 'information' in only a very limited sense is meant. And certainly, the formal prosodic resources available to a speaker appear quite limited when a broader conception of information is considered. Nevertheless, the notion that prosodic phrases represent the domain of basic ideational units is an attractive hypothesis—one which I will examine more closely in Chapter 4—as it is tied closely to the significance that these units have for a process-oriented model of speaking. Although the relationship between prosodic units and ideation (as one component in a model of speech production) is impossible to observe directly, Laver (1970: 68-9) argues for a connection in terms of 'neurolinguistic pre-preparation'. Many others echo this view. Brazil (1985b: 12), for example, speculates that 'the speaker “plans” ... and the hearer “decodes”' prosodic phrases 'as a whole'. The difficulty here—beyond the problem of accessing the inner (cognitive) processes of speech production (and perception)—is that intonational phrasing is probably not the result of a single process, but instead—like other prosodic phenomena, e.g. pausing—is multiply determined.

Numerous approaches to prosody, of course, make use of prosodic units which are determined primarily or exclusively by phonetic/phonological factors—at least that is the claim—with little or no appeal to other analytic

dimensions. For many researchers, it is simply obvious that the prosody of an utterance plays a major, if not critical, role in the production and perception of phrasal units. (This is said to constitute the 'delimitation' function of prosody, cf. Barry 1981.) It is only natural that a basic prosodic unit should be defined in terms of features of that dimension! However, there is little consensus on the specific phonetic cues which serve this function. And disagreement on whether cues must actually be present in the acoustic signal (as opposed to being reflexes of linguistic expectations) is the rule rather than the exception. Consequently, a proliferation of prosodic units has resulted.

In some cases, the prosodic unit employed in an analysis is not specifically identified, but may simply be implied by the inherent nature of the phenomenon under review. In other cases, the specificity of the definition varies. Often, no more than a rough characterization is provided through references made to other 'similar' units in the literature. But in some instances, an attempt is made to give a workable definition. An early prosodic unit that was at one time in wide use is the 'phonemic clause'. In the structuralist tradition, this unit is defined as a 'minimal complete utterance' which contains one or more pitch phonemes and only one primary stress and ends in a 'terminal juncture' (Trager & Smith 1951: 49-50). While fatally vague in detail, as well as rigid in application, the phonemic clause did serve as a way to deal with the quantal properties of spoken language. Prosodic units of more recent currency are determined by features which may undoubtedly be related to those defining the phonemic clause, but with more attention to phonetic substance. The 'tone unit' of Crystal
(1969), for example, is identified mainly by establishing the location of a 'significant' pitch change which results in the perception of a nuclear tone, but also specifically by physical pauses and the 'segmental phonetic modifications' at boundaries between units (205-6). The 'intonation phrase' of Pierrehumbert (1980) is characterized by its detailed internal tonal structure, according to the proposed explicit phonological model: 'The well-formed tunes for an intonation phrase are comprised of one or more pitch accents followed by a phrase accent and then a boundary tone' (22). The two 'prosody frames' of Gibbon (1984) are described in terms of very specific articulatory and acoustic, as well as auditory, features.

While these units and others undoubtedly differ in subtle, if not fundamental, ways, one often encounters the underlying expectation that they are more or less equivalent. Pierrehumbert (1980: 64) asserts this in a particularly strong way:

Our intonation phrase corresponds to the domain of the intonational word in Liberman (1975), to the domain of the 'tone unit' in Crystal (1969), to the 'sense group' in Armstrong and Ward (1926) and Vanderslice and Ladefoged (1972), and to the 'tone group' in Ashby (1978) and Halliday (1967[b]). It appears to be the same as the 'breath group' in Lieberman (19[6]7).

Féry (1988: 41) makes a similar claim:

Die Intonationsphrase hat schon viele Bezeichnungen erhalten, je nachdem, welche ihrer Eigenschaften gerade betrachtet wird: Phrasierungseinheit bei Bierwisch (1966), rhetorisches Syntagma bei von Essen (1956), Tongruppe bei Pheby (1980) oder Fokus-
domäne bei Gussenhoven (1983) beziehen sich in den meisten Fällen auf dasselbe Phänomen.

(The intonation phrase has many names, each according to which of its properties is in focus: the phrasal unit of Bierwisch, the rhetorical syntagma of von Essen, the tone group of Pheby, and the focus domain of Gussenhoven refer in most cases to the same phenomenon.)

Cruttenden (1986: 35), too, implies equivalence among the various units which have been proposed. On the other hand, Couper-Kuhlen (1986: 76) is more cautious. Seeing that little or no attempt is ever actually made to examine the differences among these units, this caution seems advisable. Yet, despite the weakly motivated assumption that the ‘same phenomenon’ is being referred to in most cases, it is evident—at least for a study of discourse prosody—that the basic unit in the prosodic dimension must be anchored in a phonetic basis; otherwise, it cannot be considered a prosodic unit, in the sense employed here. In §2.4 I will advocate a solution to the dilemma that the existence of multiple prosodic phrases poses by suggesting that ALL phonetic cues which are thought to play a role in delimiting phrase boundaries should be considered in a prosodic unit model.

2.2.2 Prosodic hierarchies and unit sequencing

In addition to being characterized on the basis of its segmentation, a prosodic phrase in a given framework is also characterized by its relationship to the text and to other prosodic units which may be part of a phonological analysis. In each of the approaches to prosodic phrasing mentioned
above, it is assumed that speech is to be parsed exhaustively and more or less linearly. This view has been standard for quite some time. In one of the first handbooks of German intonation, Barker (1926: 1) states simply: ‘Speech can be broken up into melodic units called tone-groups’. Implied is a linear succession of units. More recently, Halliday (1967a: 9) has reaffirmed that ‘connected speech can ... be analysed into an unbroken succession of tone groups’. And while the internal structure of the prosodic phrases might vary from language to language, there is evidence for the universality of the basic units themselves. In the early part of this century, Klinghardt (1923) observes that the European languages which he investigated (especially English, French, and German, but also informally Czech, Finnish, Greek, and Welsh) all exhibit the same basic prosodic organization: speech may be segmented into ‘Sprechtakte’.17 Of course, these views are by no means universal, as is clear when a range of approaches is considered.

When segmenting a stretch of speech into prosodic phrases, it is cus-

---

17 The term ‘Sprechtakt’ signifies different units for different authors. Sievers (1901) uses the term to refer to a stretch of speech containing one stressed syllable and any number of unstressed ones, following Sweet’s (1876) ‘stress-group’. (Curiously, Sweet 1881 contains no mention of ‘stress-group’, only ‘breath-group’.) Klinghardt (1923) expressly distances himself from this use of the term (as well as from an interpretation as ‘breath-group’), preferring instead a comparison with the ‘groupes d’énonciation’ of Passy & Rambeau (1897). Note also that Barker (1926), who bases her handbook on Klinghardt’s description, translates ‘Sprechtakt’ as ‘tone-group’. In quotations from Klinghardt’s works, however, I have rendered ‘Sprechtakt’ more literally as ‘speech group’ to avoid obvious biases (cf. also ‘bar’ in Gibbon 1976).
tomary—as implied above—to consider all phrase boundaries as shared, i.e. a given boundary is simultaneously the final boundary of one phrase and the initial boundary of the following phrase. This practice has both a theoretical and a methodological rationale. Methodologically, it is convenient to assume that when one unit ends, another begins—one’s work is in effect cut in half. There is more involved here than convenience, however, which will be evident when we later consider the phonetic cues which signal phrase boundaries. For the moment, suffice it to say that while prosody extends over the entire utterance, boundary loci are identified by cues both at the end and at the beginning of phrases. In other words, phonetic (and other) properties at the periphery of the prosodic phrase—both beginning and end—figure most prominently in boundary identification.

For these reasons, the exhaustive segmentation of speech into prosodic units is fairly uncontroversial. Yet this aspect of prosodic analysis is heavily unit dependent. Two examples can be brought to bear on this observation. First, consider the views of Brazil, Coulthard & Johns (1980). Their prosodic phrase is the ‘tone unit’, which consists of three segments: an optional proclitic portion, the tonic portion, containing the prominent syllables of the phrase, and an optional enclitic portion. Since the entire intonational meaning of a prosodic phrase is ascribed to the tonic portion of the phrase, Brazil et al. do not consider the exact location of the resulting phrase unit boundaries very important: ‘the proclitic and enclitic segments contain no prominent syllables and are thus equally uninforming. It is therefore of no great significance which tone unit they are attached to’ (46).
Despite their lack of concern where the boundaries fall, they do segment their texts exhaustively: each word of a text is represented as belonging to one prosodic phrase or another, even if this feature appears to be more a by-product of their integrated prosodic notation rather than one actively designed. In other words, their notation could easily be altered so that proclitic and enclitic segments remain outside the phrase, as is the case in Gibbon & Selting (1983) and Selting (1987a; 1987b). In the prosodic system developed in these works, the focus is also on prominent syllables: their phrase-level units are cohesive sequences of pitch accents ('rhythmisch kohäsive Akzentfolgen'). Boundary loci in this work are unimportant. And the prosodic notation reflects the fact that it is the location of the pitch accents in an utterance which gives definition to the prosodic phrases of that utterance. As in Brazil et al. (1980), the portions of the utterance which are not bracketed by the prominent syllables are given no role in the analysis. In this case, however, the prosodic notation is given in interlinear fashion, allowing a complete separation of prosody and text. In (1) below, an excerpt from the conversation transcribed and analyzed in Gibbon & Selting (1983) illustrates the characteristics of this notation.

In this notation, pitch accents are indicated by plus '+' or minus '-' (corresponding to two accent types, which are of no consequence here) placed above the relevant segment. Cohesive accent sequences are represented by grouping accents with parentheses '()', while properties of the sequences are indicated with symbols placed before or after the parentheses. (In the excerpt, an initial 'f' refers to a globally falling, 'r' to a globally rising accent sequence; a final "' refers to a type of 'continuing' intonation con-
tour.) To highlight the effects of the notation, each accent sequence in (1) is listed on its own line; in Gibbon & Selting (1983) the line breaks result arbitrarily from typographical considerations.


| a | S: | ich glaub' 'at | I believe that |
| b | die Wände noch janz [ jut ] in Ordnung sind | the walls still completely good in order are |
| c | C: | [ ja ] | yes |
| d | S: | man sieht 'et | one sees it |
| e | ja is' ja nix schief un' nix | yes is yes nothing crooked and nothing |

a | S: | I think |
b | the walls are all fine |
c | C: | yeah |
d | S: | you can see it |
e | yeah there's nothing crooked or anything |

Note that in each line (except for the backchannel in line (c), which does not contain a pitch accent), a portion of the text falls outside the scope of the accent sequence: *ich* in (a), *die* and *-nung sind* in (b), and so forth. This ungrouped material has been placed according to syntactic structure here—no meaningful prosodic organization can be made on the basis of the transcribed features; indeed, no significance is attached to the prosody of these segments. From the point of view of the prosodic analysis, then, these unaccented syllables do not exist, and in the alignment of prosody
and text they simply fall through the cracks.

Apart from the two approaches discussed above, almost all analysts assume that prosodic segmentation is to be exhaustive. There is a greater difference of opinion, however, regarding the linearity of prosodic units. Since Crystal & Quirk’s (1964) influential work on intonation, many approaches have allowed some degree of hierarchical organization in that ‘subordinate’ units are recognized, if only marginally.

<table>
<thead>
<tr>
<th>Prosodic unit(s)</th>
<th>Strictly hierarchical?</th>
<th>Exhaustive segmentation?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pike (1945)</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Halliday (1967a)</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Crystal (1969)</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Brazil (1978)</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Pierrehumbert (1980)</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Nespor &amp; Vogel (1983)</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Gibbon &amp; Selting (1983)</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Gibbon (1984)</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Ladd (1986)</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

Table 2.2 Some prosodic units and hierarchical organization

Table 2.2 lists several sets of prosodic units, indicating whether they allow some sort of unit subordination. (Whether or not the prosodic analysis exhaustively segments the text is also indicated.) Three positions concern-
ing the hierarchical organization of prosodic units can be discerned. In the
approaches of Halliday (1967a) and Brazil (1978), as well as Gibbon & Sell-
ing (1983)—all three of which have been applied to extensive amounts of
discourse—the segmentation of prosodic units is linear, whether the
phrase-level unit under consideration here constitutes one domain in a
hierarchy of levels or is the primary prosodic domain. The same is true for
many phonologically oriented models, e.g. Pierrehumbert (1980) and Nes-
por & Vogel (1983), which typically have been applied to much more lim-
ited types of language data.

On the other hand, Pike (1945) allows a limited degree of hierarchy, in
that rhythm units may be simple or complex. Crystal (1969), which is based
on Crystal & Quirk (1964), also relaxes the strictly linear organization of
units by incorporating limited subordination. In this descriptive system,
one tone group may be subordinate to another. From a hierarchical stand-
point, one level of tone group embedding is permitted. What counts as
embedding is quite restricted, however, in that only certain tonal config-
urations may be analyzed as instances of subordination. Notably, the large
London-Lund corpus (Svartvik & Quirk 1980) is transcribed using this sys-
tem.

The third position allows multiple recursion. Both Gibbon (1984) and
Ladd (1986) have argued for abandoning the simpler strictly hierarchical
view of prosodic units. The two proposals, however, differ somewhat. Gib-
bon (1984) identifies two levels of prosodic organization in discourse
(which he calls 'frames'). His 'γ-frame' corresponds roughly to the kind of
prosodic phrases which are in focus here, as evidenced by the boundary criteria associated with this level. The ‘π-frame’ constitutes a higher organizational level, one which is associated primarily with global pitch declination. For Gibbon, these higher-level frames are fully recursive; γ-frames, on the other hand, are not recursive, but are merely iterative at each π-frame level. Similarly, as mentioned earlier, Ladd (1986) operates with two organizational levels: ‘tone group’, the domain of the nucleus; and ‘major phrase’, which is identified by timing cues. For Ladd, however, both units exhibit properties of recursion: both may be grouped into larger hierarchical constituents of the same type—into what he calls ‘superdomains’. The resulting prosodic structures in each case may be quite complex. Fully aware of this, Gibbon (1984: 183) states that ‘[t]he apparent complexity of specific structures is perhaps an artefact of the analysis’. Still, it remains to be seen how successful these approaches can be over the long run. Although Gibbon applies his system to discourse data in his exposition, neither system has been applied extensively to large amounts of connected speech.

2.2.3 The ‘intonation unit’

I turn now to a discussion of the particular phrasal unit that I will be using in the prosodic analysis of my corpus of German conversational discourse: the ‘intonation unit’ (IU), after Chafe (1987). Like many other con-

---

18 Ladd & Campbell (1991) recognize ‘major tone groups’ and ‘minor tone groups’, both of which may be ‘compound’, i.e. recursive.
ceptions of prosodic phrase, an IU may be thought of as consisting of a stretch of speech (by a single speaker) uttered with a 'coherent intonation contour' (Chafe 1987: 22). Even though the exact acoustic or perceptual correlates of coherent intonation contours have never been described in detail—a fact which does call into question the usefulness of this concept in defining any type of prosodic phrase—numerous other prosodic (and non-prosodic) cues are present in the acoustic signal and so can be exploited to determine IU boundaries reliably. The nature of these cues will be discussed in the next section, and the issue of reliability will be addressed in §2.4.2.

Now, one initial question regarding terminology immediately arises which should be clarified. The reader may be puzzled by the use of 'intonation' in conjunction with 'prosodic phrase'. In light of the discussion of prosody vs. intonation in §2.1, the label 'intonation unit' would not appear particularly apt for a unit that I will be characterizing in terms of phonetic features that have less to do with intonation (in the narrow sense) and more to do with prosody in general. There are two main reasons that I have chosen to use an existing label rather than coin or recycle one. First, as I have aligned myself with the well-established analytic tradition of Chafe (1979, 1980b, 1987, 1988a) and his colleagues, I have felt it important to maintain a certain terminological continuity. The IU used in this thesis reflects the evolution that the term has undergone in light of a subsequent focus on prosodic particulars, and is essentially the same as that in recent work by Schuetze-Coburn et al. (1991), Chafe (1992a, 1992b, 1993a, 1993b), Du Bois (1992), Du Bois, Schuetze-Coburn, Paolino & Cumming
(1992),¹⁹ Du Bois, Schuetze-Coburn, Cumming & Paolino (1993), and Du Bois & Schuetze-Coburn (1993), even though it may be defined in slightly different ways in each work. Second, it must be pointed out that the name is not completely without merit, considering the phonetic means of segmenting IUs. Chafe (1980b: 14) states that probably the most important identifying feature of an IU is its 'clause-final rising or falling pitch'. Although there is much more to IUs than this, with this original emphasis on pitch movement, it is understandable that the IU was not christened 'prosodic unit' (which itself may be subject to misinterpretation), or some other similar term.

In the framework of the previous discussion, IU characteristics are similar to the those found for a majority of prosodic phrases: IUs are linearly arranged and segmentation is exhaustive. The basic analytic methodology employed reflects these two principles. For each conversational exchange, the entire excerpt is included in the analysis. Working simultaneously with transcript and recording, the speech stream is segmented first into speaker turns (see §1.6), and subsequently into IUs. No utterance or utterance part for any speaker is eliminated, ignored, or altered. This point is important especially in Chapter 3, where the syntactic structure of the discourse will be discussed, but its consideration is not without merit in the present context. As we have seen above, segmental material may be prosodically 'uninteresting' in some approaches, so that it may arbitrarily

be grouped with more prosodically salient stretches of speech. Or a transcription system may not be geared to handle noncanonical units, as is the case in the practice of some British schools.  

The compression of prosodic structure into a single linear stream has an empirical basis, in addition to serving as an aid in perceiving IUs. While it is likely that prosodic units are hierarchically arranged and that units (of any size) may be organized by the speaker into larger ones, it is also the case that in the more unplanned type of discourse which is contained in the current corpus, units at any one level—here, at the phrase level—may be conveniently represented as essentially linear. Indeed, in discussing the different processing conditions between scripted and unscripted speech which affect cohesiveness, Altenberg (1987: 37) suggests that viewing prosodic phrases as more linear is sometimes warranted:

Although Crystal (1969: 236) warns us against regarding intonation in connected speech as ‘something purely additive’, as if T[one] U[nit]s were strung together independently of each other, it is reasonable to assume that impromptu speech is more ‘additive’ than prepared speech.

This is not to deny, of course, that spontaneous discourse contains prosodic organization larger than the domain of the intonation unit. It implies only that unit subordination or other embedding is less dominant in this type of speech than in more planned varieties. Considering the vast analytical

---

simplification that this view affords, I adopt the practice in my prosodic analysis of German conversational stories.

2.3 Prosodic features and the identification of intonation units

Above I examined a variety of prosodic phrases, including IUs, in terms of two general characteristics which relate to the phonological organization of these phrases in the speech stream, hierarchy and exhaustive segmentation. These characteristics do not go very far, however, in actually identifying the prosodic segmentation of a given utterance. For this, reference to specific prosodic features is necessary. In this section, I examine one by one the major features which have been used to delineate various types of prosodic phrases, both in acoustic and in auditory analyses. Opportunity is thus provided to discuss the relevance of these features in determining IU boundaries. The particular categories of prosodic features systematically employed in the current analysis, and their symbolization in the transcription, are explained at this point. The system of discourse transcription used is a modified version of that detailed in Du Bois et al. (1992), and outlined in Du Bois et al. (1993).

A major conflict to be resolved in defining any linguistic unit is the competition between offering a generalization sufficiently abstract to account for the observed phenomenon as a whole and providing a description that will allow each instance of the phenomenon to be adequately identified. For many units, the former is more highly valued, due to research goals and methodological practice. But such units often exhibit a weak empirical foundation, in the sense that it is difficult to derive them
from natural language use. For auditorily determined prosodic units to be viable, however, emphasis must be placed on the latter, so that potentially ambiguous cases which arise in the course of an analysis do not require frequent reference to ad-hoc procedures or to syntactic or semantic structure to resolve the ambiguity. Crystal (1969: 204-207), Cruttenden (1986: 35-42), and Du Bois et al. (1992: 100-14) devote needed attention to the required specifics of phonetic features and unit demarcation, in their attempts to supply an adequate characterization, if not definition, of the prosodic phrases they use. Their work thus provides the basis for a closer examination of the parameters involved in segmenting speech, which is necessary as there still are many unanswered questions, and a synthesis of positions has not yet been achieved.

While isolating single parameters might seem artificially atomistic—as Brazil (1985b: 12) puts it: ‘The several physical features that we find we have to separate out in order to make a useable description occur, in reality, as a complex contour’—listeners, like analysts, are able to ‘decompose’ the acoustic signal into component parts to extract the linguistically relevant information (cf. Fowler 1990b: 117-20). Furthermore, when I later illustrate the notion of IU types and the prototypical structure of an IU category, it will become clear that a focus on features serves to highlight the complexity of prosodic contours. That is, an evaluation of the relative value of phonetic features in characterizing IUs is relevant not only to IU boundary identification per se, but also to defining elusive theoretical notions. Focusing on individual features is also useful because there is an obvious split in the way features function as boundary markers: some antici-
pate the END of a unit; others draw attention to the START of one.

In order to impose some reasonable limitations on the following discussion, I will focus primarily on features used specifically as boundary markers. There is a vast literature on other aspects of individual prosodic features which cannot be adequately reviewed here.\textsuperscript{21} I have grouped the phonetically based features into three primary types: those concerned with the pitch pattern of the utterance (reset, declination, nuclear and boundary tones), those related to the timing (pause, speech rate, rhythm), and other features, notably those associated with aspects of voice quality (phonation, mimicry). The role of additional, nonphonetic features—which may in some cases be heuristically useful in utterance segmentation—are also discussed.

2.3.1 Pitch features

Attending to the macrovariations in pitch—intonation in the narrow sense—can be said to constitute the primary goal of many analyses of prosodic structure. Indeed, measuring $F_0$ is often held to be sufficient to segment speech into prosodic phrases, among other things. While impressive results have been achieved in some cases (e.g. Huber 1989; Shapley 1989), computing only the moment-by-moment values of $F_0$ conveys but a part of the total prosodic information available to the listener (and analyst) which is necessary for the perception (and identification) of phrase bound-

\textsuperscript{21} A good, though selective, overview is provided in Couper-Kuhlen (1986). Crystal (1969) remains the most complete source book for English prosody; there is no German equivalent.
aries, especially in unrehearsed conversational speech. The limitations of relying solely on pitch or $F_0$ have been recognized by many researchers, especially those focusing on natural discourse, but recently there has also been interest in expanding the transcription of prosody in phonologically oriented models beyond tonal contrasts as in, for example, the TOBI system (Silverman, Beckman, Pitrelli, Ostendorf, Wightman, Price, Pierrehumbert & Hirschberg 1992), which incorporates nontonal prosodic information in terms of a single scalar measure of coherence between lexical items. Nevertheless, there are several important aspects of pitch that may be distinguished and exploited for prosodic phrase identification.

A frequently encountered characterization of prosodic phrase is that it embodies 'a coherent intonation contour' (Du Bois et al. 1992: 17; cf. Chafe 1982: 37; 1987: 22; Altenberg 1987: 47). A similar notion is found in Gibbon (1984), who writes of 'trajectory gestalt impressions' (168). The problem here, of course, is to express the exact acoustic and perceptual correlates of such a contour. While this characterization may be interpreted as a metaphor for the sum of all prosodic (and perhaps other) features that contribute to the perception of unit coherence, it may also be taken to represent the totality of prosodic features dependent just on $F_0$. Such features are not completely understood, but evidence to date suggests that the perception of coherence in a prosodic phrase (and the lack thereof across phrases) is influenced at least by four major aspects of the total pitch pattern: (1) a shift in pitch relative to the speaker's preceding utterance ('pitch reset'), (2) the global pitch trend throughout the phrase ('declination'), (3) degree and direction of pitch movement on at least one prominent syllable ('pitch ac-
cent/nuclear tone’), and (4) a final change in pitch following the nucleus ('boundary tone'). Other aspects, such as the narrowing of pitch width over the course of a phrase, are certain to play a role, but have been less well investigated, especially in connected speech.

Pitch reset. Frequent, abrupt shifts in pitch are readily observable in acoustic records of connected speech. Many of these quasiregular local pitch variations can be ascribed to phonetic manifestations of physiological events reflecting, in part, the cognitive processing of language. The sudden jumps up (or down) in pitch that provide some indication of how speaking is planned and executed are called ‘resets’. The identification of resets, however, is unfortunately clouded by other factors that may influence local variations in pitch. For example, subtle, but perceptible, ‘micro F₀’ (Kohler 1990a) perturbations are caused by segmental effects. For instance, voiced segments appear to have a relative ‘intrinsic pitch’, for example /i/ and /u/ are higher than /æ/ and /a/; and F₀ on vowels following voiceless stops is higher than that on vowels following their voiced counterparts (Lehiste & Peterson 1961). It is usually assumed that listeners’ perception of pitch compensates for these kinds of fluctuations, but there also seem to be contexts where they are actively processed (Silverman 1991: 147). A second source of interfering pitch variation is that used to signal linguistic prominence, or ‘macro F₀’ (Kohler 1990a). Of course, in contrast to segmental perturbations, this variation is perceptually quite salient. The realization of lexically stressed syllables usually, though not always, results in a large F₀ excursion. The relevance of this kind of pitch movement to utterance segmentation is discussed below (see Pitch accent and nuclear tone).
Resets appear to constitute a particularly strong class of cues for an initial phrase boundary. Cruttenden (1986: 41), for instance, states that ‘... a change in pitch of unaccented syllables is a fairly clear boundary marker’. The variable aspect of this feature is emphasized here because not all resets are equivalent: at least in experimental work, $F_0$ reset has been found to be an effective boundary marker which can also convey the importance of the boundary by the degree of reset (Vaissière 1983: 57; cf. also Sorensen & Cooper 1980: 422). Many researchers have included reset as an integral part of unit delimitation. Gibbon (1984: 171), for example, writes of the importance of ‘head/onset height’ as an indicator of an initial boundary. Johns-Lewis (1986b: xxii) mentions the general usefulness of ‘pitch height’ in a similar context. For Crystal (1969), pitch reset (‘step-up’ or ‘step-down’) is one of two primary cues.

It is clear that pitch reset is directly relevant to the segmentation of speech into IUs, too. Reset is one of five prosodic cues to IUs listed in Du Bois et al. (1992). Moreover, Schuetze-Coburn et al. (1991) find that virtually everywhere an $F_0$ reset significant enough to trigger a new pitch declination line is measured, an IU boundary is perceived as well, even when other cues are factored out. Note that while some researchers have used $F_0$ reset alone—making no reference to the lexical material (i.e. using no knowledge of the syntactic or semantic structures)—to segment various

---

22 The location of $F_0$ resets in this case resulted from the identification of ‘declination units’, whose scope was determined by the application of a standard envelope patterned after the model of Willems (1982).
types of speech into phrase or clause-sized units (e.g. Huber 1989, Shapley 1989), these units are, it appears, not equatable to IUs. While a (major) pitch reset always marks an IU boundary, such resets occur in (American English) conversation, on average, only once every two IUs (Schuetze-Coburn et al. 1991: 221). Whether or not ‘minor’ resets occur at other IU boundaries is a question open for further research, but initial indication is that a reset need not be present for an IU boundary to be perceived.

The occurrence of a pitch reset can be identified when there is an abrupt change in pitch which cannot be explained by segmental effects or intended prominence. If the difference between the final $F_0$ of one unit and the initial $F_0$ of another is small (5-15 Hz), then influence from segments cannot be ruled out. On the other hand, if the the difference is large, but is also associated with a prominent syllable, then the contribution of the pitch accent cannot be factored out. (This latter ambiguity is, in fact, one of the major sources of uncertainty in the identification of IU boundaries.)

Figures 1-2 illustrate cases that allow for the observation of pitch reset from the acoustic record. Example (2) below gives the excerpt from the primary corpus of the $F_0$ contour plotted in Figure 2.1 below. There are two IUs in this excerpt, with the pitch of the first IU falling to 178 Hz, i.e. low, transcribed ‘\•’ in (2). The beginning of the second IU starts at a mid-range value (204 Hz), in other words, at a significantly higher pitch, so that there is no question about the location of the boundary.
Figure 2.1. F₀ reset (step up) at an IU boundary.

(2) Corpus excerpt from ‘Cello lessons’ [Afternoon tea 242-3].

242 [E]: .. ~aber nur dann \bullet
   but only then

243 ... ~'s' ja wohl klär @@ \bullet
   is yes certainly clear

242 [E]: but only then.
243 that's for sure.

Figure 2.2 below shows the occurrence of F₀ reset which results in a step down, rather than a step up. Example (3) below gives the corpus excerpt corresponding to Figure 2.2. (Note, however, that only the final portion of IU 61 and the beginning of IU 62 are plotted in the figure.) As is evident from the figure, F₀ on Bett rises to 286 Hz. In contrast, the start of the following IU registers almost 50 Hz lower (227Hz), again at a mid-range value for the speaker.
Figure 2.2. \( F_0 \) reset (step down) at an IU boundary.

(3) Corpus excerpt from ‘Best kid’ [Afternoon tea 61-2].

\[
\begin{align*}
61 \quad [E]: & \quad (H) \text{'3'} \text{'un' sätzt sich so auf 's ... auf 's Bett} \quad + \\
& \quad \text{and sets self thus on the on the bed} \\
62 & \quad \text{~der is' jetzt 'änderthàlb \ nè} \quad + \\
& \quad \text{he is now one and a half ok} \\
61 \quad [E]: & \quad \text{and sits down on the on the bed,} \\
62 & \quad \text{he's one and a half now,}
\end{align*}
\]

Determining resets themselves by ear alone is as challenging a prospect as is working with acoustic records. The problem remains: what degree of pitch change constitutes a reset. While larger resets are quite obvious—they also typically cooccur with other prosodic cues; note the accompanying pause and accelerated speech in (2)—listening for smaller resets can prove to be chimerical. Thus, indicating the occurrence of pitch reset is probably less reliably accomplished by ear than the transcription of other prosodic features. Although Du Bois et al. (1992) provide a convention for systematically indicating auditorily perceived reset, this feature was not indepen-

99
dentilly transcribed for the corpus and thus is not noted in the transcription excerpts. This is not to say that pitch reset is ignored in the identification of other IU boundaries. On the contrary, reset is most definitely included as one of the features which are 'understood' to contribute to the perception of the 'coherent intonation contours' that make up the majority of IUs. Admittedly, the lack of the systematic transcription of such an important phonetic cue is problematic, but given the uncertain perceptual reliability of this feature, it seems prudent to make reference to it only when acoustic records are available.

**Declination.** Related closely to the resetting of pitch at the beginning of a phrase—but distinct from it—is the course of the global pitch pattern, which when gradually falling, is usually called 'declination' (Cohen & 't Hart 1965). Much work on intonation appeals to the notion of declination, even if the controversy surrounding the extent of its occurrence in natural speech and its role in a phonological model has yet to be resolved. Strong evidence for its existence in read speech is reported in Pierrehumbert (1979) and Huber (1989). Shapley (1989) is able to make systematic use of declina-

---

23 An exception has been made for the example text presented in Chapter 5, for most of which I have \( F_0 \) tracings. As reset in this case is determined from the acoustic record (by inspection), not auditorily, the convention recommended in Du Bois et al. (1992) is not used; instead, a raised caret [^] at the beginning of a line marks a reset.

tion for the segmentation of ordinary conversation, which also suggests, contra e.g. Lieberman, Katz, Jongman, Zimmerman & Miller (1985), that it is also a regular property of spontaneous speech. The usefulness of declination as an indicator of a prosodic phrase boundary is connected to the observation that for each speaker there is ‘... remarkable stability of the terminal endpoint F0’ (Boyce & Menn 1979: 374; cf. also Liberman & Pierrehubert 1984: 178-82; and Huber 1991: 192). Given a relatively fixed endpoint value and the rate of declination, the end of a phrase can be projected (cf. Ohala 1990: 42). That is, declination may be used to anticipate an END boundary (hence its separation from reset here). However, the global course of pitch has only limited application in the identification of completed IU's. (It may have some significance for the recognition of uncompleted, or ‘truncated’ IU's, however.) As mentioned above, Schuetze-Coburn et al. (1991) determine that the scope of declination is generally not the IU, but a larger prosodic unit. Nevertheless, its role in creating prosodic coherence is significant, and therefore must be acknowledged (see e.g. Gibbon 1984: 172-3, 183).

Indications of the global trend of the pitch are not independently given in the current transcriptions (as they are, for instance, in the transcription system of Selting 1987a; cf. example (1) above and the representation of pitch trend per Gibbon & Selting 1983). As with pitch reset, declination is difficult to transcribe reliably by ear for every IU, but the presence of declining pitch height is accounted for to some degree in the notion of ‘coherent pitch contour’ that is captured in the basic convention for transcribing IU's. In some cases, it is evident that declination phenomena contribute directly
to the delimitation of IUs.

Pitch accent and nuclear tone. The smallest perceptual unit of intonation is generally taken to be a pitch accent, whether it is considered to be a component of a 'grammar' of intonation ('t Hart & Collier 1975), one of various accent types comprising intonational 'tunes' (Pierrehumbert 1980), or an element in the description of rhythmically cohesive sequences (Gibbon & Selting 1983). Moreover, researchers agree for the most part that pitch accents '... form the foundation ... of the intonation pattern, and the pattern is largely built upon them' (Fox 1984: 7). There is also no doubt that significant pitch movements play an important role in determining prosodic phrase boundaries. In addition to contributing to the shape of an intonation contour, accents serve as perceptual landmarks in an ever-fluctuating acoustic signal. Terken (1991: 291) writes of the 'demarcative function' of major prominences. Bolinger (1987) draws attention to the accents at the beginning, as well as at the end, of phrases, i.e. at their boundaries.

In many approaches, one pitch accent—the 'nuclear' or 'tonic' accent—is viewed as playing the most important role in the characterization of a prosodic phrase.25 Altenberg (1987: 47), for example, states that '... the nucleus is undoubtedly the primary feature'. The nucleus, which is usually considered to be the most prominent accent of the phrase, is seen as the one obligatory prosodic constituent. In other words, there is by definition exactly one nuclear accent per prosodic phrase. Furthermore, its identifica-

tion is crucial to the identification of the phrase boundaries. A couple of problems, however, arise in this regard. The nucleus is usually characterized as the final pitch accent of the phrase. Von Essen (1956: 23), for example, states: ‘Als Schwerpunkt wirkt immer diejenige (Haupt-) Akzentsilbe, der kein gehobener Silbenton mehr folgt’ (‘the (primary) accent which is not followed by a pitch prominence functions as the nucleus’). But when the domain is not given in advance (as it usually is in such descriptions), the notion ‘final accent’ does not help determine the boundaries. In an utterance with a series of accents, no utterance-medial accent is phrase final until after the boundaries are defined. Put another way, there may be difficulty reliably distinguishing the nuclear accent in a prosodic phrase from other prominences. Often enough, more than one syllable in a stretch of speech is a possible candidate for nuclear status. If one pitch accent is clearly most prominent, then it is, of course, regarded as the nucleus. But what if no one accent stands out? Brown, Currie & Kenworthy (1980) report this problem both in their own analysis and in experimental situations; in fact, they complain of ‘... considerable difficulty in identifying tonics ...’ (154). These problems have led some researchers to degrade the importance of the nucleus—Klein (1982: 296) calls the idea of a nucleus (‘Satzakzent’) a ‘hilflose deskriptive Kategorie’ (‘unhelpful descrip-

---

26 Cf. also Winkler (1979: 23), Cruttenden (1986: 49), and Gibbon (1987: 4). Bolinger (1972: 644) frames the question of nucleus placement in a characteristically interesting way: ‘... the speaker will put the main accent as far to the right as he dares ...’.  

27 This may be true even in a delineated utterance. Currie (1981: 19) concludes that an unmarked tone group will contain ‘several peaks of prominence’.

103
tive category)—or at least to relax the usual restriction of one primary accent per phrase, as is the case in Du Bois et al. (1992, 1993) and Chafe (1993b). Nevertheless, there are properties of prosodic phrases which make the idea of a nucleus attractive. Crystal (1969: 207) points out, for instance, that ‘[t]he presence of a nucleus is what accounts for our intuition of “completeness” at the end of a unit ...’; And Cruttenden (1986: 52) counters that abandoning the concept is a position which ‘... does not take into account the very large number of cases where nucleus assignment is straightforward’.

As mentioned above, the question of nucleus is tied up with the status of other accents in the prosodic phrase. It is common to recognize at least three levels of prominence. Fox (1984), for example, recognizes ‘primary accent’, ‘secondary accent’, and no accent (‘unaccented’). But under another view, the perception of different levels of prominence is not phonetically motivated. In his classic study, Lieberman (1965) finds little support for more than one degree of prosodic prominence, although there may be other (e.g. semantic) reasons for insisting on two or more. That is, whether a syllable has a primary or secondary (or tertiary) accent is said to depend on factors other than local variations of pitch, loudness, and duration.\[28\] To clarify the confusion of prosodic relationships, Vanderslice & Ladefoged (1972) posit a set of binary suprasegmental features. Among other things,

\[28\] This is not to deny necessarily that more than one level of prominence may be meaningful. Wells’ (1986) findings, in fact, support the view that listeners are sensitive to varying degrees of prominence. But the mechanisms involved remain unclear.
they separate the phonetic aspects of prominence (labeled [±accent]) from phrase-level effects ([±intonation]). Regarding the nucleus, they state:

Nuclear accents are of course subjectively more prominent than prenuclear ones. But nuclearity is a matter of intonation .... Therefore, it is not only possible but (we believe) obligatory to factor out that extra prominence, attribute it to the feature Intonation, and declare all accents equal. (825)

This position accords well with the view expressed in Zwirner & Richter (1966: 117-8) that secondary accents ('Nebenakzente') should be considered realizations of lexical stress, and that primary accents ('Hauptakzente') must be a property of phrases.

In transcribing IUs, no special indication of a nuclear (or tonic) syllable is generally given. All perceptually prominent syllables are marked with a grave accent [\']. A standard 'default' convention is used to interpret the relative prominence: unless otherwise indicated, the final accented syllable in the unit is understood to be most prominent. In other words, the IU-final pitch accent may be taken to be the nuclear accent, if such an interpretation is desired. This convention follows the simplification outlined in Vanderslice & Ladefoged (1972); operationally, it frees the analyst from making difficult—if not impossible—decisions about the relative weight of prominences: a fully articulated syllable is either accented or unaccented.\(^{29}\)

\(^{29}\) Of course, there is much attenuation in connected speech, so that a syllable may also be weaker in relation to its neighbors. Such reduction is viewed as a separate (nonboundary) phenomenon, and is discussed in Chapter 1 in conjunction with orthographic representation.
Example (4) illustrates this principle:

(4) Corpus excerpt from 'Aldi drunk' [Skat 5-8].

5  [T]: (H) «4» zu dem Lesebrief \+
    to the letter.to.the.editor

6  ... ich will nämlich 'n Lesebrief schreib'n /+
    I want namely a letter.to.the.editor write

7  ...(2.2) an= ... =die Verkäuferin vom Aldi \
    to the cashier of.the A

8  ...(1.5) sie soll den =lèsen \•
    she should it read

5  [T]: about the letter to the editor,
6  I want to write a letter to the editor,
7  to the cashier at Aldi.
8  she should read it.

In IU 5 there is only one prominent syllable, so the question of the degree of prominence is not very important. In IU 6 two accented syllables are perceived. Here, the level of relative prominence corresponds to the 'default' case (the accent on Les- is weaker than that on schreib-), so the accentual pattern is adequately represented with one accent mark. A similar situation is found in IU 8. In IU 7 there are three prominent syllables, and the situation is more complex. The initial prominence on an is due in large part to the perceived loudness of the vowel, while prominence on -käu-
    can be attributed to the pitch excursion there. (There is also a pitch accent on Al-
    the perceptually strongest syllable.) Because various factors have con-
    tributed to the perception of prominence, it is difficult to say whether the first and fourth syllable of the IU differ or have the same degree of prominence. Hence, it is more satisfactory in this case—and certainly more ana-
    lytically reliable—to say simply that both syllables are prominent. The final
accent of the IU is also marked the same way, following the accent convention.

Now, it does not always happen that the most prominent syllable in an IU coincides with the final accent of the unit, as the terms are applied here. Occasionally, a syllable earlier in the IU is prosodically highlighted through loudness, duration, or pitch (or a combination thereof) and so is perceptually the most prominent. This is sometimes the case for so-called 'focused' constituents, as well as for accents which convey or imply 'contrast'. Clearly, the stronger syllable should be specially marked in these cases, otherwise important information about the prosodic structure of the IU would not be represented. The notation that I use reserves the acute accent ['] for nonfinal prominent syllables that are stronger than the final prominence in the IU. Example (5) illustrates:

(5) Corpus excerpt from ‘Getting by’ [Henrika’s visit 7-9].

7  [H]:  sie hat das nicht .. stringént durchgehalten /+
        she has that not strict maintained

8   aber sie hat 's versucht /+
      but she has it tried

9   so güt wie möglich zu mächten \*
      so good as possible to make

7  [H]:  she hasn’t kept to it strictly,
8     but she’s tried,
9     to do it as best as possible.

In IU 7 there are three prominent syllables: sie, -gent, and durch-. The pitch excursion on -gent is quite large, and the accent is perceptually stronger. The relatively stronger prominence on -gent as compared with the unit-fi-
nal prominence on durch- is thus indicated by the acute accent.

**Boundary tone.** Distinct from any pitch accent associated with the realization of lexical stress is the course of pitch that occurs at the edges of prosodic phrases. Pierrehumbert (1980: 26) states: ‘The end of the intonation phrase has distinctive tonal characteristics, apart from those attributable to the pitch accents’. The pitch configuration following the final prominence in a prosodic phrase is generally referred to as a ‘boundary tone’ (‘Grenzton’), or ‘tail’ (‘Nachlauf’ or ‘Abtakt’) of the unit (Crystal 1969: 207-8).\(^\text{30}\) And it is usually assumed that this configuration serves as an important marker of a phrase end. Various studies have confirmed the general importance of boundary tones for the perception of prosodic units. Henderson & Nelms (1980), for example, find that listeners seem to be preoccupied with ‘decoding’ an utterance after a final fall: reaction times to cooccurring events are longer there than in other locations.

Apart from signaling the occurrence of a phrase boundary, the course of pitch on syllables following the final pitch accent in a phrase-sized unit gives important information about the type of phrase. Working from connected, spontaneous speech, Delattre, Poenack & Olsen (1965: 137) establish acoustically what others have noted on an auditory basis: that it is the

---

\(^{30}\) Cf. also ‘Phasenton’ (Bannert 1985), which is associated with some boundaries. In generative models, ‘phrase accents’ occur in this position as well, though they apparently lack a motivated role in accounts of German intonation (Féry 1988: 53; Uhlmann 1988: 75). The earlier influential paradigm developed by the American structuralists generally referred to the prosodic phenomena at phrase boundaries as ‘juncture’ (see e.g. Hockett 1942; Trager & Smith 1951).
shape of pitch on the unaccented syllables at the end of a unit which indicates whether the unit expresses a sense of completion or not. Unaccented syllables following the last prominence in a unit are typically low for finality, high for continuation. The importance of this dichotomy is voiced in many approaches. Brazil (1985a: 65), for example, states: 'The central opposition in the part of the meaning system realized by tone is that associated with end-falling tones on the one hand and end-rising tones on the other'.

The basic functional considerations involved have led some researchers to distinguish just two boundary tones, which is the position I take here. The perceived direction of the final pitch movement at the end of an IU is represented either as 'falling' (iconically marked with a back-slash [\]) or as 'nonfalling' (marked with a slash [/]). It should be pointed out that 'nonfalling' pitch subsumes the 'rise' and 'level' categories in Du Bois et al. (1992, 1993). It should also be mentioned that there is an additional motivation behind this notational system. Various studies have shown that only a two-way distinction for boundary tones can be justified on perceptual grounds. Hadding-Koch (1961: 61), for example, finds that the basic opposition is between 'final' and 'non-final' pitch movements. Lieberman (1965: 52) suggests that the three 'terminal junctures' of Trager & Smith (1951) could be reduced to two: 'falling' and 'nonfalling'. Apart from this, a simple consideration was addressed: the level of effort involved in an auditory transcription of pitch increases the more decisions the analyst has to make. A binary system minimalizes the process while maintaining a meaningful distinction.
There is more to be said about pitch at prosodic phrase boundaries, however. Some indication can be found that the boundary tone alone does not provide the characterization of the phrase, but that it is the combination of a pitch accent plus the pitch movement in the tail which defines the contour and thereby signals a prosodic phrase boundary. In a series of psychoacoustic experiments (Hadding-Koch & Studdert-Kennedy 1964; Studdert-Kennedy & Hadding 1973; Hadding & Studdert-Kennedy 1974), it was found that the perception of final pitch movement depends on aspects of the intonation such as $F_0$ peak and ‘turning point’ (the level at which the last pitch change occurs), in addition to the actual direction of $F_0$ at the end of the contour. Wichmann’s (1991) study is congruous with this result. She finds that the height of the pitch peak before a final fall affects the perception of how ‘final’ a phrase is. In other words, all falls to low pitch do not sound the same functionally.

Because of the importance of a basic functional classification of IU types and the apparent need for the representation of more than the boundary tone to convey them, a feature which summarizes the contributions of the whole pitch contour is included in the current transcriptions. The feature is referred to as the ‘transitional continuity class’ of the IU (Du Bois et al. 1992: 28-31), and represents, specifically, the class of contours which are functionally similar with respect to the degree of continuity at the transition from one IU to the next. Transitional continuity here is viewed again as a binary feature, either ‘complete’ (marked with a bullet [•]) or ‘non-complete’ (marked with a plus sign [+]). (Thus, no use is made of a separate ‘appeal’ class, as in Du Bois et al. 1992, 1993; see below for discussion). The
combination of final pitch movement and transitional continuity features therefore allows for a four-way classification of IUs based on intonation contour, signifying the following:³¹

[\•] Falling pitch in the tail to a level at or near the bottom of the speaker’s range, typically with overall declining pitch, and no large sustained pitch excursion towards the end of the IU. It is often called a ‘final fall’, and will be referred to as ‘low fall’ here.

[/•] Rising pitch either over the expanse of the IU or in the tail, but in each case to a very high level. Sometimes the pitch of the whole IU is heightened. It will be referred to as ‘high rise’.

[/+] Rising or level pitch in the tail, with a rise generally not as high as in a high rise. Overall pitch may be high or sustained. Occasionally, the end rise is quite high (mostly as part of the final pitch accent), but in the tail the pitch levels off, or even falls somewhat. The overall auditory impression, however, is of a rise.

[\+] Falling pitch in the tail, but not to the bottom of the speaker’s range. In addition, overall pitch may be higher than in a final fall, or sustained over the course of the IU.

³¹ Cf. the four types of boundary tones used in Price, Ostendorf, Shattuck-Hufnagel & Fong (1991: 2962).
I have used bullet and plus sign to mark transitional continuity—instead of the more usual period and comma—in order to emphasize (through size) the importance of this category.\textsuperscript{32}

Example (6) illustrates common occurrences of these contour classes:

(6) Corpus excerpt from ‘Project’ [Afternoon tea 7-12].

7 S: ... ~ich kànnd das Projekt ja wâhl'n \+
    I can the project yes choose

8 .. dann <p wèiß [ich= p\+ \bullet
    than know I

9 (E): [mhm ((EATING)) \+

10 * ...(1.1)

11 E: ~un' w's war'n nóch für Projèkte \+
    and what were still for projects

12 .. kànnsnt 'e dich entsinn'n \*
    can you yourself remember

7 S: I can choose my project,
8 then I'll know.
9 (E): mhm,
10 * (pause)
11 E: and what other projects were there,
12 can you remember?

\textsuperscript{32} Du Bois et al. (1992, 1993) and many others use the ordinary punctuation. There is something to be said for this practice. First, the symbols are familiar; they have long been used by discourse researchers. Second, comma and period convey the basic dichotomy outlined here, even to the uninitiated, with little or no explanation. On the other hand, because of their widespread use, there may be some misunderstanding as to their technical meaning here. Furthermore, the two symbols are small and similar in form. My experience using these symbols has made it clear that readers will sometimes ignore or confuse them. Thus, I have opted for new, I hope equally iconic symbols.
IU 7 in (6) exemplifies a noncomplete falling contour, which is very common when speakers are winding down their turns. IU 8 illustrates a low fall, in this instance the final IU of Sabine’s turn. IUs 9 & 11 bear noncomplete nonfalling contours, which are typical when the speaker intends to continue speaking. IU 12 exhibits a high rise, as a direct address to the listener following the question in IU 11.

A few remarks about the four contour types are in order. To the low fall contour class belong realizations of what many linguists believe to be ‘normal, declarative intonation’. The isolated words, phrases, and ‘sentences’ of psychoacoustic experiments often have to do with intonational patterns in this class. Instances of the high rise contour class typically indicate an ‘appeal’ (Du Bois et al. 1992: 30-1). Appeal includes what is referred to as ‘(yes-no) question intonation’ (cf. IU 12 above), but it is not limited to any one syntactic form or pragmatic use. The concept of ‘finality’ or ‘completion’ that these first two contour classes have as part of their characterization is intonational, not interactional (despite the above example): after uttering an IU with a low fall or a high rise, the speaker may well continue talking without break or interruption. On the other hand, these are indeed points that have a high potential for speaker change (see Ford & Thompson 1993). The two noncomplete contour classes are harder to provide with descriptive labels. This is in part due to the wider variety of contours that are understood to be part of this class. Meinhold’s (1967: 478) conclusion from a statistical analysis of a corpus of read speech applies here as well:

... es [handelt] sich beim progredienten Verlauf um keine streng
normgebundene Intonationsform, sondern [es muß]—im Gegensatz zum terminalen Verlauf—eine erhebliche Variationsbreite der Realization eingeräumt werden.

(there is no strict norm in the intonational pattern of continuing contours; instead, wide variation in the realization [of this class] must be allowed for, in contrast to the final contour class.)

Still, it is useful to have some label. With this in mind, I will refer to ‘mid rise’ \([/+]\) and ‘mid fall’ \([\backslash+]\) contours. Finally, one additional notation is needed: truncated (i.e. uncompleted) IUs are indicated with a double hyphen \([-\cdot-]\), regardless of whether they are abandoned or interrupted.

At this point, it makes sense to compare briefly the above system with the treatment of boundary tones in the discourse-oriented studies of spontaneous German listed in Table 2.1 (§2.1), as most offer alternate views. Interestingly, Winkler (1979) also makes use of four contour types (‘Kadenzen’), but it is not immediately obvious how they relate to the four types proposed here. He distinguishes ‘rise’ (‘Steigkadenz’), ‘level’ (‘Schwebekadenz’), ‘half fall’ (‘Halbschluß’), and ‘full fall’ (‘Vollschluß’). The clearest correspondence is probably between my ‘low fall’ and his ‘full fall’, as he states (22): ‘Die sog. “Lösungstiefe” der Stimme wird jedoch erst beim Vollschluß erreicht’ (‘The speaker’s lowest pitch is only achieved in a full fall’). Presumably his ‘level’ would be grouped with some instances of ‘rise’ in my ‘mid rise’, and the remaining instances of ‘rise’ would constitute my ‘high rise’. His ‘half fall’ is likely to correspond to my ‘mid fall’, although he states that this contour also evokes a sense of completion.
In contrast, Delattre, Poenack & Olsen (1965) describe just three contours for German: 'minor continuation', 'major continuation', and 'finality'—although they investigate only, in their words, 'declarative intonation' (135). Continuation is characterized by a rise in pitch (a major continuation rises higher), finality by a fall. They apparently have not considered any instances of 'high rise', so there is no correspondence to be found there. From their 'intonation graphs', it is evident that their 'finality' class corresponds to my 'low fall'. In each case, the speaker's pitch reaches the baseline. Likewise from the pitch tracings, it seems that both of their continuation types would be grouped together in my 'mid rise', although one or two instances of 'minor continuation' would probably correspond to 'mid fall'. They note that continuation in their German samples is usually manifested by rising pitch (unlike in their English material, where it is usually falling).\textsuperscript{33}

Schaeffer (1979) and Selting (1987b) also distinguish three boundary tones: 'falling', 'rising', and 'level', much as is outlined in von Essen (1956). Jin (1990), too, has three: 'falling', 'rising', and 'rise-fall' ('Gipfelakzent'), but in her work (based in part on experimental data) she does not separate the contribution of the boundary tone from the neighboring pitch accent.

On the other hand, Royé (1983) recognizes five tones at boundaries: 'rising', 'level', and three degrees of 'falling', depending on how low the pitch drops. The correspondence between his system and mine is likely to

\textsuperscript{33} Indeed, it is the frequent slight fall in pitch at the end of a noncomplete prosodic phrase that motivates the representation of both continuity class and final pitch direction in the system of transcription of Du Bois et al. (1992; 1993).
be similar to that of Winkler's, discussed above, with the inclusion of two (intermediate) degrees of fall coming under my 'mid fall'.

Rhode & Roßdeutscher (1973) stand apart from other approaches in that they work directly from auditorily transcribed contours, i.e. they do not reduce the pitch movements to a finite number of types. However, they do refer to 'rises' ('Stimmhebung') and 'falls' ('Stimmsenkung') in describing the role of intonation in specific examples. (Girouard 1984 also does not treat the topic of boundary tones directly, but he does seem to transcribe final rises and falls.)

To recap, we see that there are a variety of possible ways to classify boundary tones. All approaches make use of at least two types, rise and fall. A level tone may also be distinguished, and degrees of fall may be specified. The system proposed here combines two aspects of the pitch contour for which there is some empirical support—the direction of the final pitch movement ('boundary tone') and the overall contribution of the contour in conveying completion ('transitional continuity').

One additional aspect of boundary tone representation that is specific to colloquial German must still be mentioned: the pitch on final 'lexical boundary markers' (for a discussion of this class of particles, see below). In particular, the particle ne usually carries a rapidly rising tone (often to a fairly high pitch), which is almost always made even more prominent by the incidence of an immediately preceding fall, often to low. Despite this large pitch movement, when there are no other prosodic indications of a boundary, it appears that the configuration is best analyzed as a single,
complex pitch movement extending from the last pitch accent to the end of
the phrase, i.e. as a fall-rise. Rhythmically, there is usually no break be-
fore ne. Thus, in these cases, ne indicates the presence of a boundary in re-
lation to the next IU, but it itself is taken to be part of the preceding IU. This
complex pitch movement is indicated by the addition of a backslash (for
the falling pitch) before ne, as illustrated at the end of IU 66 in example (7):

(7) Corpus excerpt from 'Weavers' fair' [Doppelkopf 64-70].

64 (D): .. @ [ @
65 [M]: [ der wär schon so in der vördersten Rèihe / +
there was already thus in the frontmost row
66 dass man [z das schon überblicken ] [ konnte ] \ ne / +
that one that already view could ok
67 (D): [z ach sô= ] \ •
oh thus
68 (Z1): [ @N ]
69 (J): @ [z ] @@@@@@@
70 (D): [z @]

64 (D): (laughs)
65 [M]: it was in the front row,
66 so you could see everything,
67 (D): oh.
68 (Z1): (laughs)
69 (J): [(laughs)
70 (D): [(laughs)

Other aspects of pitch. Other pitch metrics, such as maximum and min-
imum F₀ height, maximum F₀ movement, F₀ peak shape, and F₀ peak
alignment with respect to the accented syllable also provide potentially

34 Cf. Esser (1983) on the difficulty of separating movements of this type.
useful information. To the extent that it occurs, pitch width trend—specifically, the tendency for the difference between F₀ peaks and valleys to decrease within a prosodic phrase or over a sequence of phrases—can be considered an independent feature which contributes to the perception of the completion of an IU. (The narrowing of pitch width is explicitly modeled in some accounts of declination, e.g. Thorsen 1985.) An accurate representation, however, is virtually impossible without systematic acoustic analysis. One useful aspect of pitch can be auditorily discerned, though. Speakers occasionally talk with a much wider pitch range than normal (e.g. in uttering exclamations) or with a much narrower one (e.g. in parentheticals), and the stretch of speech over which the wide or narrow pitch extends often corresponds to an IU. When it seems important to the delimitation of IUs, this type of supplementary pitch aspect may be indicated in the transcriptions by means of an open-ended ‘angled-bracket notation’, whereby the stretch of speech that exhibits a particular feature is bracketed and an index added to indicate the feature type. See Appendix 1 for a list of types used here, and Du Bois et al. (1992: 52-8) for a fuller list. In example (8), IU 15 is bracketed for narrow pitch [ < N > ]:

(8) Corpus excerpt from ‘Gas stove’ [Morning visit 14-6].

14 [K]: .. Thèo is’ ja% --
T is yes
15 .. < N ~wèišt ja wer Thèo is' N > \•
know yes who T is
16 T: .. < p nõ= p > \•
no
14 [K]: Theo is--
15 <you> know who Theo is.
16 T: no.
To sum up the use of pitch features, numerous aspects of pitch can be distinguished which may be used to help segment utterances into IUs, including initial reset, global declination, pitch accent sequence, boundary tone, and pitch width trend, among others. The difficulty in reliably transcribing many of these features by ear and the level of effort required for a detailed notation has led to a minimal system along the lines of Du Bois et al. (1992, 1993). Of the intonational cues discussed, two are systematically indicated in the transcriptions: (pitch) prominence and boundary tone. In addition, ‘transitional continuity’ serves as a cover term for aspects of the pitch contour not otherwise represented that convey the degree of intonational completion expressed.

2.3.2 Timing features

Cues to prosodic phrases having to do with the timing, or temporal organization, of an utterance, like pitch cues, consist of a complex of features which both contribute to perceived unit cohesiveness and serve to mark breaks between units. Apart from pause, timing involves the temporal relationship between segments or syllables, as Nooteboom (1991: 228) explains:

Durations of realizations of one and the same vowel phoneme may vary from practically zero to many hundreds of m[illi]s[econds]. Such variation is not random, but rather rigorously controlled by many factors and their interactions. The result is what we call the temporal organization or temporal patterning of speech.
The particular features investigated in this section include pause between units, accelerated rate of speech on unit-initial syllables (including 'anacrusis'), decelerated rate of speech on unit-final syllables (including 'prosodic lengthening'), and changes in the perception of rhythm.

**Pause.** After the contribution of pitch contour, the prosodic feature most often mentioned in conjunction with the identification of phrasal units is 'pause'. Pike (1945: 31-3), Jassem (1952: 43), Crystal (1969: 206), Chafe (1980b: 14), Cruttenden (1986: 36-9), Du Bois et al. (1992: 100) all consider pauses very important in unit segmentation, even if the occurrence of this feature varies from genre to genre and from speaker to speaker. Altenberg (1987: 27) states the importance of pause in more general terms, and in a way that is without a doubt accurate: '... pausing—especially at or near T[one] U[nit] boundaries—plays a crucial role in the production and perception of speech'.\(^{35}\) For Brown et al. (1980), pause constitutes the only boundary feature. It should be noted that not all researchers consider pause to be this useful in segmenting speech, e.g. Henderson & Nelms (1980).

Pause has been claimed to be a valuable cue in signaling various types of linguistic structure besides prosodic phrases: word, subject-predicate distinctions, sentence, and paragraph (which may be governed primarily by

\(^{35}\) Grammarians, of course, have recognized the basic idea here, too; cf. Wängler's (1974: 219) comment: 'Längere Sätze ... werden vom Sprecher des klareren Verständnisses wegen gern in SINNGRUPPEN aufgeteilt. Die Teilung wird durch Zäsuren bewirkt ...' ('Longer sentences are divide into sense groups by the speaker for purposes of clarity. The division is effected by means of pauses'). But their orientation is invariably syntactic, with little attention given to actual speech production.
other strong organizational principles). The role pause is given in evidencing such structures is often tied to variation in its duration (e.g. Crystal & Quirk 1964: 49); the length of the pause is said to correspond to boundary type (Chafe 1979: 162-3; Lehiste 1979: 195-6). These higher-order units do not generally conflict with the identification of phrase-level units, however—a phrase boundary would be indicated at every sentence or paragraph boundary anyway—and so can be ignored when segmenting speech into IUs. Still, it is important to be mindful that speakers pause for many different reasons (cf. Chafe 1980c; Deese 1980; Good & Butterworth 1980), and thus not all pauses serve a segmenting function.

Some attention should be devoted at this point to elucidating what exactly is meant by ‘pause’. O’Connell & Kowal (1990) are rightly critical of apparent inconsistencies in the use of the term ‘pause’. A basic distinction is often made between ‘filled’ and ‘unfilled’ pauses (Maclay & Osgood 1959), the former referring to vocalizations which express hesitation of some kind, and the latter, periods of silence between stretches of speech. But this classification does not suffice to make clear exactly which phenomena may be included or excluded in transcribing spontaneous discourse.

As an illustration, let us consider a few of the complications which may arise in noting unfilled pauses. An undifferentiated analysis of the acoustic signal would identify ‘silence’ (i.e. zero energy) at places in running speech which are not usually conceived of as pauses by discourse researchers. Du

36Cf. Pike’s (1945: 31) ‘tentative pause’, which is usually shorter, and ‘final pause’, usually longer.
Bois et al. (1992: 45) point out one instance of where the lack of acoustic energy is not equitable with pause: 'The moment of silence which necessarily occurs during a lexically or phonologically required voiceless stop should not be classified as a pause ...'. This is a specific case of 'articulatory pause',\footnote{Cf. also the 'intrasegmental pause' of Drommel (1980: 228).} which psycholinguists have sought to distinguish from 'hesitation pauses', largely by means of durational differences. From a phonetician's viewpoint, this is a somewhat curious enterprise, for it is unlikely that any absolute cut-off point will be tenable, whether it be 0.25 second (Goldman-Eisler 1968) or 0.13 second (Hieke, Kowal & O'Connell 1983), if it is applied without reflecting on the context. There is simply too much articulatory variation. Dankovičová (1992), for example, reports stop closures up to 0.44 second in elicited spontaneous Czech monologs. This may seem exceptionally long, but stop closure durations well within the domain of 'slight' pauses in the present corpus (i.e. 0.1-0.3 second) are not uncommon. An instance of an alveolar closure lasting 0.18 second is illustrated in Figure 2.3 below. The transcript excerpt corresponding to the figure is given in (9) below; the voiceless alveolar stop in *toll* (IU 84) is the segment in question.

Another source of 'silence' which may be confused with 'pause' is very weak vocalization. In field recordings, where the signal-to-noise-ratio is hardly optimal, speech of very low intensity may not be evident at all. And even if some vocalization is apparent, the exact timing of crucial articulatory events cannot be ascertained with the accuracy required to measure the pause with any confidence.
Figure 2.3. Example of a stop closure with a long duration.
Closure begins at the left arrow and lasts until the release at the second arrow.

(9) Corpus excerpt from 'Best kid' [Afternoon tea 83-5].
83  (S):  @@@  [ @ (H)
84  [E]:  [ ~und dër is' sô  [2 tô=11] ] + and he is so great
     85  (S):  [2 @sû=per ] + super

83  (S):  (laughs)
84  [E]:  and he's so great,
85  (S):  super,

A more problematic case—because of its high frequency—occurs in conjunction with normal breathing (Du Bois et al. 1992: 49-50). In the transcription of breathing and adjacent pausing from audio recordings, quiet inhalation or exhalation—which should in principle be marked as a separate event—may be obscured, giving the appearance of a longer pause than is actually present in the acoustic signal. (Alternately, a pause may be transcribed in place of breathing.) In both cases, however, no uncertainties as to the segmentation are introduced, as an inbreath or outbreath usually sig-
nals a boundary, regardless of whether a pause cooccurs (see *Inhalation* in §2.3.4).

Complications also arise from the point of view of the pause notation. In some transcription systems, pauses which occur in conjunction with unit boundaries may not be indicated. For example, in the widely used HIAT system (Ehlich & Rehbein 1976), a pause at the end of a ‘sentence’ (indicated as in normal orthography with a period) is not transcribed separately, but is ‘understood’.\(^{38}\) There may be some justification for this practice; in psychoacoustic studies, subjects fail to notice (or underestimate the duration) of these short ‘boundary pauses’ (e.g. Butcher 1981: 205; cf. also Zvirner & Richter 1965: 119). Conversely, it may be justifiable to speak of ‘pause’ in places where no silence actually occurs. An instance of this is embodied in the notion of lengthening as ‘compensation’ for the absence of actual physical silence, here reported in Pike (1945: 31):

The tentative pause has one very important alternate form: instead of a gap in the speech, a complete cessation, there may be a lengthening of the last sound or two of the preceding word. This length takes up the same time as the physical pause would have done.

Pause often enters into the problem of ‘virtual’ boundaries as well. For example, in Jassem’s (1952) monograph, a ‘tonal juncture’—necessary for defining the boundaries of the ‘tone group’—can occur at ‘a place in an ut-

\(^{38}\) ‘... markiert dieses gleichfalls eine kurze Pause’ (24) (‘this [period] marks a short pause as well’).
terance having no phonetic value at which a normal pause may be inserted without causing any change of meaning ...’ (45). The notion of ‘possible pause’ is also mentioned in conjunction with prosodic boundary markers in Bierwisch (1966). A similar conception of boundaries as a phonological attribute of intonational phrases can be found in more recent work: ‘As a rule of thumb, an intonation phrase boundary ... can be taken to occur where there is a nonhesitation pause or where a pause could be felicitously inserted without perturbing the pitch contour’ (Pierrehumbert 1980: 19). In light of the differing treatment of pause, perhaps the most judicious statement that one can make is that pause, in one guise or another, ‘contributes to the perception of intonational units’ (Johns-Lewis 1986b: xxii). It is one of many cues which taken together reinforce the perception of a boundary.

Pauses as they are represented in transcriptions presented here fall into three classes, ranging from shortest to longest: ‘slight’, ‘default’, and ‘timed’. There may be functional significance to this division, but the

39 ‘Grenzsymbole determinieren mögliche Pausen ...’ (186) (‘Boundary symbols determine possible pauses’).
40 Even this statement would be controversial. The degree to which pause is said to determine or correspond to other prosodic structure is highly variable. In many approaches, pause is not a defining (or even characterizing) criterion, but is merely a cooccurring phenomenon (e.g. Brazil 1985b), or a different matter entirely (e.g. Gibbon 1984). Still, under the conceptions of prosody and prosodic phrase considered here, pause clearly plays a strong role as a boundary cue.
41 These classes correspond exactly to the ‘short’, ‘medium’, and ‘long’ pauses of Du Bois et al. (1992: 42-5; 1993: 61-3). I have opted to refrain from using such labels here, because the
primary concerns are with reliability and readability. Timed pauses [(seconds)] can be reliably measured by ear using the stopwatch method to an accuracy of about a tenth of a second (Du Bois et al. 1992: 42-3; Schuetze-Coburn, in progress a). Default pauses [ ... ] contain a clearly perceptible period of silence, about 0.3–0.6 second, but cannot be measured in the same way. Slight pauses [ ... ] are barely noticable, about 0.2 second or less. As durations of this length border on the limits of human perception (Lehiste 1970: 14-5), the perception of slight pauses may actually be due more to sudden changes in speech rate and to an alteration of the rhythmic pattern than to the period of silence itself. When there is no pause between IUs at speaker turns, a special convention is used to highlight this fact [(0)]. Example (10) below illustrates the various pause conventions.

Usually, the pause with occurs between IUs is written with the following IU, as in IUs 280 & 281 in (10). Sometimes, however, a pause cannot be convincingly associated with a particular speaker. In such cases, the pause is written on a line by itself, as is done in IU 278. This situation normally arises in two situations: (a) when the current speaker is finished talking and there is a time lag before the next speaker begins talking, or (b) when the current speaker is apparently at turn’s end, pauses for a while without anyone else taking the floor, and then chooses to continue talking.

duration of silence that is considered to be functionally ‘short’ or ‘long’ varies considerably from system to system (cf. Edwards 1993: 23-4). While my labels may also seem arbitrary, they are motivated for the primary corpus: the mean duration of pauses occurring between IUs corresponds to a default pause.
(10) Corpus excerpt from 'Propane tank' [Morning visit 277-82].

277 [T]: genàu= \*  
exactly
278 *  ...(9)
279 T: stimmt \*  
correct
280 K: ... das-- 
that
281 .. ~die habt 'hr hierher g%-- .. getrà=g'n /•  
it have you to here carried
282 T: (0) jà= \+  
yes

277 [T]: exactly.
278 * (pause)
279 T: right.
280 K: that-
281 you carried it here?
282 T: yeah,

(11) Corpus excerpt from 'Washing machine' [Morning visit 167-71].

167 [T]: solang--  
as long as
168 ... (H) «.7» solang das «A dann noch irg'ndwie A» geh=t \+  
as long as that then still somehow goes
169 (K): .. rh \*  
170 *  ...(9)
171 K: (H) «1.1» née /+  
no

167 [T]: as long as--
168 as long as it's still working,
169 (K): m h
170 * (pause)
171 K: no,

127
Finally, the total duration of complex events which include pauses (i.e. inbreath plus pause, laughter plus pause, etc.) is usually indicated in double angled brackets [« »]—in addition to the estimated pause duration—due to the difficulty mentioned above of measuring the length of such pauses. A separate notation is used to avoid possible confusion with timed pauses. Example (11) above illustrates.

Speech rate. Regularly occurring changes in the rate of speaking constitute another important group of boundary cues. Considerable attention has been devoted to these cues. Crystal (1969) suggests that ‘... tempo has probably the most highly discrete grammatical function of all prosodic parameters other than pitch’ (153), and he finds that it ‘... quite clearly stands out as independently contrastive ...’ (176) in his own data. As boundary markers in spontaneous speech, Winkler (1979: 18) states: ‘Häufig ist ein ebenfalls gliedernd wirkendes Ritardando und Accelerando ...’ (‘Also frequent is the segmenting function of ritardando and accelerando speech’). This property of speech rate would seem to be in conflict with acoustic studies (e.g. Goldman-Eisler 1968) which report that the articulation rate for a given speaker is fairly constant (rather, it is the pause rate that varies). But Miller, Grosjean & Lomanto (1984: 218) clarify the situation by pointing out that: ‘... the data reported in the literature are, for the most part, based on averages across long stretches of speech, and thus do not reveal the extent of local variation in speaking rate within a given utterance ...’. In their acoustic study of 30 radio interviews, they find that the average articulation rate for periods of pause-free speech fluctuates considerably (219).
Measuring speech rate is complicated endeavor. As mentioned above, the choice of units and time interval is crucial. Furthermore, since speech rate is clearly a gradient property, specific (possibly arbitrary) thresholds would have to be established in order to make use of any acoustically determined values. It appears, however, that for the purpose of phrase boundary identification it is possible to simplify the analysis and identify stretches of speech which are (perceptually) faster or slower relative to adjacent stretches. Winkler (1979: 18), Gibbon (1984: 171), Cruttenden (1986: 39), and Du Bois et al. (1992: 100) consider junctural acceleration and lengthening in this way. Thus, there is a three-way tempo contrast in prosodic phrases to be indicated in a discourse transcription: accelerated, modal, and lengthened. As a boundary diagnostic, accelerated speech typically occurs phrase initially, while lengthened speech occurs phrase finally, so that the crucial juncture falls at the shift from lengthened to accelerated rates, modal to accelerated, or lengthened to modal, depending on which rates are realized in adjacent units. The onset of accelerated speech that functions as a boundary cue is indicated in the transcriptions with a prefixed tilde [~]; lengthened speech of a word or more at the end of an IU, by a prefixed wavy equal sign [=].

Lengthened segments within a word are marked with a postfixed equal sign [=]]; modal speech is unmarked. Significant tempo variation also occurs over longer stretches of speech. Occasionally, an entire IU will be set off from adjacent IUs by a change in speech

---

42 All speech marked for tempo is indicated by the angled-bracket notation in Du Bois et al. (1992; 1993). I have used a visually simpler notation here for tempo variation at boundary edges due to the importance of these cues.
rate. These cases are handled as voice quality distinctions (see below), and are indicated with the angled bracket notation. Rapid speech is bracketed ‘<A A>’ (allegro), slow speech ‘<L L>’ (lento).

Figure 2.4 below illustrates tempo variation in a stretch of speech consisting of four IUs; the full corpus excerpt is given in example (12).

(12) Corpus excerpt from ‘Cello lessons’ [Afternoon tea 219-23].

219 [E]: (H) «.5» <L ûnd .. ûhm L> ++
        and uh
220     da wàr er so =zwòlf /+
        there was he thus twelve
221     oder drèi=zehn \ nè /+
        or thirteen ok
222     .. [ (H) «.5» ~un' die war gànz =schrècklich \ nè /+
        and she was completely horrible ok
223   (S):    [ mhm ++

219 [E]: and uh,
220 he was like twelve,
221 or thirteen,
222 [and she was really horrible,
223 (S): [mhm,

In the figure, speech waveforms are presented in IU style (i.e. one IU per line), with the approximate centers of the syllabic segments (in this case, all vowels) marked below the waveform with small filled squares (•). The text (in orthography) is aligned below that.
Figure 2.4. Waveforms of corpus excerpt (12) with syllabic centers marked (·) illustrating tempo variation in connected speech.
Highlighting the syllabic segments in this way enables one to contrast visually local speech rate variation without detailed calculations: the closer together the squares, the faster the tempo. The whole of the first IU in the figure is uttered at a relatively slow tempo; most of the segments are drawn out. Thus, the entire IU is bracketed ‘lento’. In the next unit—IU 220—the entire last word, zwölf [tsvöl], stands out as being uttered at a slower rate than the other words in the IU, and is so marked. In IU 221 syllable drei- is lengthened. And in IU 222 both accelerated speech at the beginning of the unit and lengthening at the end occurs. In this example, an instructive comparision can be made between the tokens of war in IUs 220 & 222. Although both occur in initial stretches of speech that are considerably faster than the later portions, in the first instance no acceleration is indicated, while the second is marked for accelerated speech. The primary reason for this is evident from the waveforms: the instance of war in IU 220 is accented (that in IU 222 is not), and the word is comparatively longer. In fact, it is more than 50% longer (0.23 vs. 0.15 second), despite the similarity in duration of the preceding words. The tempo of the initial words in IU 222 come across as uniformly fast, while in IU 220 the initial words are rhythmically uneven.

Accelerated speech at the beginning of a prosodic phrase (which thus has a demarcative function) is often called ‘anacrusis’ (Jassem 1952: 40; Gibbon 1984: 185; Cruttenden 1986: 24). Specifically, the term refers to one or more unstressed syllables before the first stressed syllable in a prosodic unit, usually spoken rapidly. (The term for this position in the prosodic phrase is ‘prehead’, ‘Vorlauf’, ‘Auftakt’, cf. Sievers 1901: 236-7; Klinghardt
1923: 20.) Strictly speaking, however, anacrusis is a metrical feature—the syllables in question being extrametrical (proclitic) to their rhythm group; but the apparent tendency for an increase in tempo to be associated with the anacrustic syllables has led to a focus on this property instead. In the context of phrase boundary cues, I prefer to use the expression ‘accelerated speech’, which does not carry the same metrical connotation, as it seems to be the change in speech rate which helps to signal a new prosodic phrase, rather than any extrametricality of the initial syllables.\footnote{Cf. Cruttenden (1986: 39): ‘The sudden acceleration ... indicates that [the] syllables [in question] are anacrustic and hence that a new intonation-group is beginning’.} Furthermore, there appear to be cases where the accelerated tempo continues well into the phrase itself, and the marked stretch of speech includes accented syllables. And instances of phrases which consist entirely of accelerated speech may serve to delimit flanking phrases, like prosodic parentheticals.\footnote{Cf. Winkler (1979: 18): ‘Schroffe Tempwechsel gliedern besonders Einschübe aus’ (‘Quick tempo shifts especially segment parentheticals’).} 

Just as there is apparently a tendency to increase the rate of speech at the beginning of a prosodic unit, the tendency to decrease the speech rate at the end of a unit is also a regular feature of prosodic phrases. This lengthening is variously known as ‘prosodic lengthening’ (Du Bois et al. 1992: 37), ‘segmental lengthening’ (Klatt 1975; Johns-Lewis 1986b: xxi), ‘preboundary lengthening’ (Lehiste 1979), ‘final lengthening’ (Thorsen 1987: 84), and other similar expressions.\footnote{As the names imply, preboundary and final lengthening are necessarily associated with some right juncture; prosodic and segmental lengthening cover this position, but are not re-}
place is controversial, as the difference in labels suggests. Lengthening is often regarded as a feature of the final syllable of a prosodic phrase (Cruttenden 1986: 39-40), but careful studies (using primarily read speech) indicate that is is actually just the rime of the of the final syllable that is lengthened (Wightman, Shattuck-Hufnagel, Ostendorf & Price 1992: 1714). In the spontaneous speech of the present corpus, however, it is observed that any prominent syllable near, as well as at, the end of an IU, together with the syllable(s) following this prominence, may exhibit lengthening. (The lengthening on the actual prominent syllable is concomitant with the realization of stress, cf. Lehiste 1970: 36.) Occasionally, as with acceleration, speech may be decelerated over even larger stretches. Undoubtedly, some differences are due to the way in which the prosodic units are defined (e.g. how syntactically oriented they are). But in other cases it seems that different prosodic mechanisms might be involved. Klatt (1976) reviews reports of lengthening at several structural levels, including word, phrase, clause, sentence, and paragraph boundary, in addition to its well-known function as indicator of segmental composition (i.e. long vs. short, voiced vs. voiceless). He concludes that ‘segmental timing carries a high functional load ...’ (1208). The value of duration as an IU boundary cue (for the listener) might thus seem dubious, considering the numerous loci that lengthening has been identified as coinciding with; however, it should be remembered that in most of the surveyed studies, the lengthened segments do not serve to help pinpoint a boundary, but are only correlated with boundaries which

---

stricted to this use. See Beckman & Edwards (1990), Fowler (1990a), and Selkirk (1990) for a variety of recent positions on what this right juncture is.

134
have been determined on other (typically syntactic) grounds. While the use of duration in truly spontaneous speech seems less consistent than experimental studies have indicated, it proves to be a very important feature in the identification of IU boundaries.

Rhythm. Pause and speech rate do not constitute all aspects of timing in an utterance. There is in addition the notion of 'rhythm'. Broadly speaking, any event which is perceived to occur at more or less regular intervals can be said to exhibit rhythm. For speech in particular, the periodic event is generally the perception of a prominent syllable. Fry (1968: 368) defines linguistic rhythm as follows: 'Every syllable in the message will bear some degree of stress and the succession of these stresses makes up a rhythm or a rhythmic pattern'. The perceived beats—which may include stress, accent, and meter expectations (Gibbon 1988: 18)—are clearly perceptible in controlled experiments (Allen 1972), and may be used in interactional speech for such purposes as the coordination of speaker turn transitions (Couper-Kuhlen 1991). Although the value of rhythm for the identification of prosodic phrase boundaries in connected speech has not been systematically investigated, it may be suspected that a sudden change in the rhythmic pattern of an utterance may trigger the perception of a boundary, similar to the way regular shifts in speech rate do, or that the coherence lent to a sequence of speech by such a pattern may serve to set it off from neighboring sequences.

In transcribing the present corpus, no special rhythmic patterns were noted which in and of themselves determined an IU boundary. Occasion-
ally, however, a particularly striking rhythm may be indicated (using
angled-bracket notation [ «RH RH» ]) to help convey a better ‘feel’ for an
excerpt. A finer transcription of IU 9 in (5) above, for example, could have
indicated the rhythmic nature of the accent sequence resulting from the al-
ternation of function and content words, in this instance roughly an
iambic pattern ( — — — ) :

(13) Corpus excerpt from (5).

9  [H]:  «RH so gut wie möglich zu machen RH » •
      so good as possible to make

9  [H]:  to do it as best as possible.

While the rhythm of the IU does not identify any boundaries, it does add a
measure of internal coherence to the unit. See Du Bois et al. (1992: 53-4) for
additional examples and discussion.

In sum, timing features—in particular, the occurrence of pauses and the
modulation of tempo—give an excellent indication of the prosodic struc-
ture of an utterance, as far as necessary for identifying IUs. In conjunction
with the pitch features outlined in the previous section, the vast majority
of IUs can be delimited. Other temporal aspects, such as internal rhythm,
may also at times contribute to the perception of an IU boundary.

2.3.3 Voice quality

In addition to pitch, duration, and loudness (see §2.3.4), there are other
aspects of the acoustic signal that convey information about speakers, their
messages, and their attitudes towards both. These aspects may be collec-
tively called 'voice quality'. To be distinguished for the purposes under discussion are those aspects which constitute 'one of the most important clues to a speaker's identity' (Fry 1968: 369) and those which may vary over the course of an utterance, either intentionally or symptomatically. Especially interesting here are the features which vary from phrase to phrase, as these may provide subtle indications about speech production processes. Voice quality features are probably as a whole relatively underutilized as active prosodic boundary markers; in most cases, qualities accompany other phonetic features. Still, changes in voice quality can function (heuristically, at least) to signal the start of a new intonation unit. Some aspects of voice quality, such as phonation type (in particular laryngealization), provide direct indication of hesitation or locate shifts in utterance planning so that they play a not insignificant role in boundary identification. Other identifiable qualities, such as voice spoken with a smile—which Crystal (1969: 136) describes as 'spread' articulation—or other expressions of emotion, nondistinguishing nasalization and tension, and 'mimicry' (e.g. in 'quotation' quality), all occur sporadically, and may occasionally be of some value. Generally, voice qualities are indicated in the transcriptions only as necessary; see Du Bois et al. (1992: 52-8) for examples and discussion.

*Phonation*. Variations in phonation type provide subtle, yet significant indications as to the possible presence of a prosodic phrase boundary. Of particular interest is the occurrence of low pitched, aperiodic voicing, i.e.

---

46 For Crystal (1969), only the former is 'voice quality'. The second type of aspects he calls 'voice qualifiers' and 'voice qualifications'.

137
laryngealization or creaky voice (cf. 'creak', Johns-Lewis 1986b: xxi). Several researchers report that laryngealization can be used as a very prominent boundary indicator in phrasing connected speech. Lehiste (1979) finds that it is one of three important cues to the end of linguistic units. Huber (1991: 193) observes that laryngealization cues are present in preboundary positions in both read and spontaneous speech (though to different degrees). Du Bois et al. (1992: 48) note that a related phenomenon—paralinguistic glottal stop—can be a useful cue for abandoned utterances. The laryngealization that occurs at the end of a phrasal unit clearly has an articulatory mechanism. If the rate of speech decreases as the utterance progresses, as it often does, there will be a tendency for the pitch to lower. While low pitch cannot be equated with creaky voice in any simple way, in practice when speakers' pitch approaches the bottom of their range, laryngealization is the likely outcome. Crystal (1969: 173), for example, found that '... almost all instances of low [pitch] ... showed co-occurrence, particularly with piano, narrow, creak and diminuendo'. Still, the variation evident from Huber's study suggests that speakers do manipulate the degree to which laryngealization is realized. It should be noted that the usefulness of this feature is diminished for some speakers, however, in that laryngealization is a semi-permanent voice quality (in the sense of a personal attribute).

In the transcriptions, laryngealized speech is indicated by enclosing the

\[ \text{47 See Ladefoged (1971: 14-6) and Catford (1982 [1977]: 98) for a discussion of the articulation, aerodynamics, and linguistic use of laryngealization.} \]
affected portion of the utterance in bracketed percent signs [% %]. Occasionally, IU boundaries are marked not by creaky voice extending over any significant stretch of speech, but by a glottal stop or glottal constriction. As mentioned, this is especially the case when an IU is not completed by the speaker, but is abandoned or restarted. Glottal stops and constrictions, when relevant to IU boundary demarcation, are indicated in the transcriptions by a percent sign prefixed or suffixed to a word, or when independently articulated, by a percent sign in parentheses [% %].

In colloquial German, there is another speech feature which might be covered in this section, namely the use of a pulmonic ingressive airstream mechanism. Virtually all speech using the lungs is ingressive; only in very exceptional situations is it ingressive (Ohala 1990: 23). One place that many (though not all) German speakers regularly use an ingressive airstream is with the backchannel ja ‘yeah’. When ingressive ja is then followed up by a speaker turn, it is always prosodically distinct from this turn. An instance occurs in IU 137 of example (14) below, where the ingressive speech is bracketed ‘<IN IN>’.

Note that the shift from pulmonic ingressive to pulmonic egressive in Henrika’s speech at the juncture of IUs 137-8 provides an additional cue which makes the boundary unequivocal.

---

48 The fact that speech is normally pulmonically egressive may provide a motivation for the use of ingressive backchannels: such an utterance would not be construed as a claim on the floor, and so is interactionally less challenging.
(14) Corpus excerpt from ‘Getting by’ [Henrika’s visit 132-40].

132 [G]: .. das ist sicher nicht das was sie= .. wîll /+
that is certainly not what she wants

133 oder was sie= möchte \+
or what she would like

134 [ode=r ] \+
or

135 (I): [mhm= ] \+

136 .. oder das was sie= so als Perspektive für P sich sieht P \•
or that what she thus as perspectives for self sees

137 (H): (0) <IN ja IN> \•
yes

138 H: .. nur ich denke \+ #
only I think

139 ~sie [brauch'] einfach mal 'ne Zeit /+
she needs simply once a time

140 (I): [ja= ] \•
yes

132 [G]: that’s certainly not what she wants,
133 or what she’d like,
134 [or,
135 (I): [mhm,
136 or what she sees like as prospects for herself.
137 (H): yeah.
138 H: but I think,
139 she simply needs some time,
140 (I): yeah.

Mimicry. A change in voice quality to mimic or impersonate the original speaker in reported speech (‘quotation quality’) often constitutes a complex prosodic signal of a new IU. As such, it functions more or less as a feature in the sense considered here. Even though many aspects of mimicry are directly related to pitch and timing, further phonetic elabora-
tion is usually not necessary for the purposes of segmentation. Example (15) illustrates the prosodic separation of the imitated speech:

(15) Corpus excerpt from 'Aldi drunk' [Skt 20-3].

20 [T]: und hat/+ and has
21 .. Q "hahaha Q"/+ ah ah ah
22 was Sie hier aufmachen/+ what you here open up
23 müssen S'e auch kaufen"/+ must you also buy

20 [T]: and [said],
21 "ah ah ah,
22 whatever you open up here,
23 you have to buy",

In this example (as in the transcriptions), the extent of quotation quality is indicated by enclosing the prosodically qualified speech in angled brackets indexed with a small capital Q, i.e. [Q Q ]. While the reported speech encompasses the three IUs 21-3 (indicated by the double quotation marks), voice quality changes used to imitate the original speaker are maintained only over the stretch of speech in the first of these IUs. The shifts in quality in Tamina's speech at the beginning and end of IU 21 contribute significantly to the perception of this utterance as a separate IU.

2.3.4 Other prosodic features

Loudness. This important feature is 'the perceptual correlate of variation in amplitude of sound-wave vibration, which in turn is the physical
correlate of physiological *intensity* of utterance' (Crystal 1969: 120). Loudness constitutes one of three primary perceptual correlates of intonation, the others being pitch and duration. There is an established cooccurrence of low-volume (quiet) speech with low pitch range (Crystal 1969: 174), suggesting its potential use as a boundary marker: the loudness of an utterance will generally decrease towards the end of a phrase, rather than elsewhere. Apart from the use of loudness to demarcate parenthetical utterances, this feature will not be apparent in every phrase; as with declination, it is instead more likely to coincide with the (planned) end of a larger prosodic unit, or with a speaker's turn. Unfortunately, the practical value of this feature is quite limited. The problems associated with obtaining a valid recording of intensity usually cannot be overcome, except in laboratory situations. Even under ideal recording conditions, measurements of intensity are at best comparative, due to differences in integration time and frequency response of instrumentation (Fry 1968: 386). A further complication is that loudness may not only be related to observable properties in the acoustic signal, but to the listener's perception of the speaker's vocal effort as well (Rosenblum & Fowler 1991).

Despite its theoretical importance, loudness cannot play a significant role in the determination of IUs from audio recordings. Thus, only highly prominent shifts in intensity which are clearly present in the signal are indicated in the transcriptions. The angled-bracket notation is used: loud speech is indexed with a small capital *F* ('forte'), i.e. [ F ]; soft speech, with a small capital *P* ('piano'), i.e. [ P ].

142
Inhalation. As inhalation is part of the rhythm of speaking, like pause, it functions as part of the speaker’s prosody, and so is treated here with other prosodic features. The presence of an inbreath has been noted to virtually guarantee a prosodic phrase boundary. Winkler (1969: 17) comments that ‘Einatmung bewirkt zwar stets Gliederung, sagt jedoch über die Tiefe des Einschnitts nichts Zwingendes aus’ ('inhilation always results in segmentation, but says nothing about the strength of the boundary'). The apparent reason for this seems to be largely physiological. In reporting on their experiment using read speech, Gelfer, Harris & Baer (1987: 58) suggest that during an inhalation, ‘... such variables as subglottal pressure and/or laryngeal muscle activity’ will return to some initial state, resulting in an F0 reset in the next utterance.

As was mentioned in a previous section, reset is a strong cue to a new prosodic phrase. It should be noted, however, that all inhalations in connected spontaneous speech are not the same. Sometimes the inbreath (and resulting break in phonation) is extremely short—on the order of 0.1 second. Other times it is rather long, around 1.0 second. Thus, it is doubtful whether all instances of inhalation carry an identical reset; this matter requires further examination. Still, the basic observation seems to hold.

Inhalation is marked in the transcriptions using a small capital H enclosed in parentheses [(H)]. Exhalation is similarly marked, only with the addition of x, i.e. [(Hx)].
2.3.5 Nonprosodic factors

*Change of speaker.* Implicit in this analysis (and virtually all others) is the notion that a change of speaker, or ‘speaker turn’, constitutes an absolute limit of a prosodic unit. The reason for this may be primarily physiological. As discussed above, after a speaker takes a breath, or after an interlocutor pauses speaking for a sufficient length of time to induce a return of the vocal tract to an initial prespeech state, a system reset is practically unavoidable, and there will thus be prosodic evidence for the start of a new phrase. One may surmise that, in most instances, a turn-initial utterance will follow a period which includes a normal inhalation or a ‘resetting’ of the vocal tract, so that this speech will be prosodically distinct from preceding utterances. Schuetze-Coburn et al. (1991: 222) found that 98% of new turns started with an F<sub>0</sub> reset, thus effectively separating them off.

Depending on the interactional nature of the exchange, new speaker turns may provide unambiguous indication for a significant portion of the IU boundaries in a conversational excerpt.\(^{49}\) While the determination of turns is not always a straightforward process—for instance, overlapping

\(^{49}\) However, in very ‘lively’ speech, or when numerous interlocutors are involved, there is likely to be much overlapping speech. Occasionally, speaker B will start or complete a turn after speaker A has paused, but before the completion of speaker A’s current IU. In such cases, the ‘second part’ of speaker A’s IU does not constitute a new turn, but is simply the continuation of a turn that has not yet been completed. See my discussion of turns and turn representation in §1.5; see also Du Bois et al. (1992: 67-8) for a discussion (with examples) of this kind of structure. Note that in this situation it may be the continuity afforded by the IU that contributes to the perception of speaker turn continuity.
speech by another speaker may be treated as a concurrent turn, but a mere
signal of continued listener involvement ('backchannels') are not consid-
ered to constitute turns per se—the dynamics of verbal interaction is such
that it is generally clear when a speaker change plays a role in establishing a
new prosodic boundary. In general, I considered a new turn to begin when
it was clear that a speaker gained—or attempted to gain—the floor.
Backchannels exhibited no such intention.\textsuperscript{50} Thus, even a short utterance
such as \textit{mhm} was marked as a new turn when it constituted a response re-
quested by the previous speaker requiring that speaker to relinquish the
floor, if temporarily (e.g. as an answer to a yes-no question); but when no
such request was evident and no other indication of claim to the floor was
found, the utterance was marked as a backchannel response.

In the transcription excerpts, speakers are indicated in the second col-
umn by a single letter speaker code (e.g. 'A', 'B', etc.), followed by a colon.
The reason for the use of one character codes is purely out of considera-
tions of space: there is not enough room in the allotted column for full
names. However, as it is generally preferable to use (fictional) names (Du
Bois et al. 1992: 22), names, not codes, will be used in the discussion. This
practice should not cause any confusion, since speaker names are unique
across recordings and transcription segments (i.e. each name signifies one
and only one speaker, even in different transcriptions), and a code is sim-
ply the initial of the speaker. (See Table 1.1 for a full list of speakers.)

\textsuperscript{50} See Sacks, Schegloff \& Jefferson (1974) and Oreström (1983) for a full discussion of the pa-
rameters involved in determining speaker turns.
In order to indicate new speaker turns, codes in the excerpts are not explicit for each IU (i.e. each line of the transcription); rather, the last speaker indicated (ignoring backchannels) remains the 'current speaker' until a new turn is signaled. (Backchannel utterances are indicated distinctly, with the speaker code in parentheses.) Thus, a new speaker turn begins whenever a speaker code without parentheses or brackets is given.

Syntax. Disregarding the use of syntax as the PRIMARY determinant of prosodic phrase boundaries, which was previously discussed, I wish now to draw attention to cases where syntactic structure is explicitly cited as a SECONDARY determinant. Crystal (1969: 206-7) mentions the occasional (rare) use of grammatical criteria, but he does not go into detail about his findings. Chafe (1980b: 14) is more specific, as he includes syntax as one of three cues, after intonation contour and hesitation markers. Cruttenden (1986: 43-4) mentions particular (hypothetical) cases where there may simply be no other usable cues.51 One of these cases involves what Cruttenden calls 'intonational sandhi', the intonational merging of two independent prosodic phrases. But in this and other situations, researchers must have in mind POTENTIAL boundaries, whereby a prosodic boundary is considered present if a pause can be inserted felicitously at the syntactic juncture, or if a boundary is expected, based on what has been observed in other material. On the surface there is no phonetic indication of a boundary. In such cases, I would like to emphasize, a boundary is NOT to be indicated; syntax plays no direct role in the prosodic segmentation of the current study. Although

51 Cf. the discussion of Pheby's (1975) position in §2.1 above.
it must be acknowledged that when an analyst (or listener!) has access to
the grammatical (and semantic) structure of a text, this structure undoubt-
edly affects judgements, particularly in cases where other evidence is weak.
Nevertheless, it is deemed possible to ‘avoid syntactic thinking’ (Du Bois et
al. 1992: 101) to a degree sufficient to make consistent and valid judgments
of IU boundaries.

One point concerning the placement of boundaries and what might be
construed as syntax needs to be made. Words (roughly, orthographic words
in the transcription) are generally to be regarded as wholes in the place-
ment of IU boundaries. Although words (as constituents) are the building
blocks of syntactic structure, they are also the primary locus of processes in
other linguistic dimensions. Thus, while at a micro-phonetic level a pitch
accent may be said to occur in relation to some vocalic (segmental) center,
in the functional domain it is the word that is perceived as prominent
(Ladd 1991: 283). Similarly, tempo changes frequently occur within a word,
but it is the change in relation to other words that is important. The unity
of words in the perception of IU boundaries may be due in part to the in-
fluence of the written language. In a language such as German which has a
long written history and where the literacy rate is very high, this supposi-
tion must be entertained. Of course, in the exceptional case, an IU bound-
dary MAY be located within a word—there is no absolute prohibition at issue
here. In German especially, with it propensity for long lexical compounds,
it might be suspected that breaks in the middle of a compound will occur.
Indeed, there are cases of this in the corpus, as example (16):
(16) Corpus excerpt from ‘Gorges du Verdun’ [Uli’s visit 64-5].

64 [U]: \~das ist das Kl\"etter% /+
that is the climbing

65 .. \~gebiet "{u}berh\~aupt =in Fr\~anreich dort \~, region at all in France there

64 [U]: that's the climbing,
65 region par excellence in France.

In the middle of the the ‘word’ Kl\"ettergebiet ‘climbing region’, Uli hesitates (as evidenced by the glottal stop at the end of IU 64 and the slight pause at the beginning of IU 65) and then resumes at an accelerated rate of speech. The sum of these cues produces a strong prosodic break within the word, so that an IU boundary is perceived at this point.

Semantics. Similar to the role of syntax discussed above, certain aspects of the meaning of utterances are occasionally mentioned as a necessary last resort in determining boundary placement (e.g. Crystal 1969: 207). Perhaps due to the influence of standard examples, there is often an expectation of semantic coherence and completion in prosodic phrases. As noted by Du Bois et al. (1992: 106), the analyst may fail to ‘... recognize as an intonation unit a stretch of discourse that expresses no substantial idea, or which does not seem to contain its own independent referential meaning’. But again, it is possible to disregard the semantic structure of an utterance to a large degree, with a complete focus instead on the prosodic features which shape the speaker’s phrasing.

Lexical boundary markers. Although IU boundaries are directly determined only by prosodic features, a set of lexical expressions called ‘lexical
boundary markers' ('Gliederungssignale', Güllich 1970) constitute an overt nonprosodic means of recognizing that a prosodic boundary has been produced.\textsuperscript{52} Significantly, lexical boundary markers normally exhibit many of the above-mentioned prosodic features which warrant drawing a prosodic boundary; the lexical form serves, in part, to highlight, if not index, the boundary (cf. Stellmacher 1972: 522). Viewed from a cognitive perspective, it may be the case that such markers—being lexically fixed, semantically reduced 'idioms' ('Floskeln')—buy the speaker more processing time, as has been suggested for filled pauses and prosodic lengthening (cf. Quasthoff 1979: 44ff). Clearly, though, they also play a delimitative role interactionally, and so it is legitimate to take them into consideration when establishing intonation unit boundaries.

While lexical boundary markers may generally be characterized by their positional distribution (i.e. utterance initial or final, 'Eingangs/Schlußsignale') and by their reduced semantic content, syntactically they form a rather heterogeneous class, ranging from particle/adverb, e.g. also 'thus', to clause, e.g. \textit{ich würde sagen} 'I would say' (cf. Wackernagel-Jolles 1973b). One crucial question that has not been satisfactorily answered is the degree to which they constitute a relatively stable set of items versus a more open-ended class. If the latter is the case—and this must be suspected, given the diversity of forms—we ultimately need to know how to differentiate the

\textsuperscript{52} Trubetzkoy (1958 [1939]: 242) introduces the related term 'Grenzsignal', which he intended to apply to phonetic indicators of sentence or utterance boundaries, as well as word and morpheme boundaries.
occurrence of a lexical boundary marker from the use of a lexical item (or syntactic construction) in its ‘usual’ sense. This line of inquiry, however, would take us far afield from the present goals; in her approach to defining the related group of ‘discourse markers’, Schiffrin (1987: 31-41) has some suggestions which could be pursued further. Here it will suffice to concentrate on the few lexical boundary markers which are prosodically rather prominent, and which I have used in segmenting the speech stream into IUs.

One such marker in German conversation is *ja* (‘yes’). In situations where *ja* cannot be construed as an answer to a previously posed question, it may serve merely as an ‘introductory marker’ (‘Eingangssignal’) to what the speaker has to say (Wackernagel-Jolles 1973b: 167). This type of marker often occurs turn initially, in a separate intonation unit with marked syllabic lengthening and a nonfalling (level) tone. In this position, it does little to help delimit IUs, which have already been delimited by speaker turn. But turn internally it draws attention to a probable boundary between IUs, whether it itself constitutes a separate IU or not.

A second prosodically prominent marker is *ne*. It (and its variants *wa* and *nich*) is perhaps the most frequently occurring marker in informal, spontaneous conversation—and the most useful in terms of prosodic boundary identification. It has a range of functions, roughly comparable to the tags ‘isn’t it?’, ‘ok?’, and ‘right?’ in English. The effect is often rhetorical, and while a response may be elicited, a verbal reply is usually not expected. (The prosodic description of *ne* was discussed above in conjunction with
boundary tones.)

In sum, a variety of prosodic features can be isolated which serve to demarcate prosodic phrases in connected speech. Besides the various aspects of pitch discussed above, features of pause and speech rate are most significant. Other prosodic cues play an occasional role. The contribution of speaker turn is very important, but the role of syntax and semantics is minimal; however, other, nonprosodic, features such as lexical boundary markers have been recognized as indexing transition points in discourse which may also be prosodically marked as well. Thus, such markers can serve as heuristics to help identify the location of IU boundaries, though they do not demand a boundary in and of themselves.

2.3.6 The representation of prosodic features

Having covered the various cues which are used to identify and demarcate prosodic phrases, I now turn briefly to the representation of these features in a discourse transcription. For the purposes of locating prosodic boundaries, one’s transcription would, ideally, track and notate each prosodic parameter completely independently of all others, off the text line. Off-line notation is in principle more flexible (see, e.g. Gibbon 1984, and his depiction of ‘prosodic frames’). The prosodic notation is not hampered by the segmental (or lexical) material—it can be precisely aligned—and the readability of the text itself is not degraded. Such a notation could be accomplished, for example, by indicating the values for each parameter on a separate vertical coordinate, with the horizontal axis representing time, and the segmental material appropriately aligned.
Figure 2.5 below illustrates this representational scheme using a very short excerpt, the full transcription of which is given in (17):

(17) Corpus excerpt from 'Best kid' [Afternoon tea 103-9].

103 [S]: (H) «.S» irg' ndwann kommt denn der W ünsch @ Kinder zu krieg'n @ \\
          sometime comes then the wish kids to get
104 E: ... (H) «.8» nà ma' ècht \\
          now just real
105 .. [ ~is' das nich' ] schrè~cklich= \\
          is that not horrible
106 S: [ wie es so kòmmt ] \\
          how it so comes
107 E: .. nèe= \\
          no
108 ~also bei mir bestimmmt nich' \\
          so by me certainly not
109 (H) «.4» «F nèin F» /+ \\
          no

103 [S]: sometime or other one wants to have kids,
104 E: well now really,
105 [isn't that just awful,
106 S: [it just happens,
107 E: no,
108 certainly not to me.
109 no!

The analytic benefits of this kind of transcription are evident. The convergence of several features at particular points in time in the speech stream is made clear, indicating prosodic boundaries. The 'strength' of the boundary can easily be calculated by tallying the number of features occur-
Figure 2.5. Coordinate transcription of a corpus excerpt from ‘Best kid’ [Afternoon tea 104-8].
ring, for example, at word boundaries.\textsuperscript{53} In the figure, the top line gives the 'boundary weight' resulting from such a tally. Between \textit{na} and \textit{mal} there are none of the tracked features present; between \textit{mal} and \textit{echt} there is one feature (laryngealization); between \textit{echt} and \textit{ist}, three features (pause, boundary tone, accelerated speech); and so on. In this short excerpt, an IU boundary has been transcribed at places where more than one feature is aligned with a word boundary.

There is, of course, a large investment involved in producing this detailed a transcription. Just as important, having separate lines for each parameter reduces the readability of the transcript as a gestalt significantly, and it violates several other principles in the design of an adequate discourse transcription (Du Bois 1991). Such an arrangement may be more appropriate for a database. One way to make off-line notation more usable is to collapse all prosodic notation into one line, as is done in Gibbon & Selting (1983) and Selting (1987a). But then, whether the prosodic notion is further integrated into the text line or not does not seem to be particularly significant, despite the observations of Gibbon (1984: 180-1). An in-line prosodic notation seems perfectly adequate for the purposes of this study, and so, for simplicity's sake, is used here, which should already be clear from examples (2)-(16). In fact, it may be more appropriate. As mentioned in the previous section, some prosodic features are not individually notated in the transcript, but are taken into account during the auditory seg-

\textsuperscript{53} Cf. the use of 'break indices' in Price, Ostendorf, Shattuck-Hufnagel & Fong (1991) to indicate boundary strength.
mentation. Each marker of final pitch movement and transitional continuity, then, in effect represents the sum of those cues for locating the boundary at that point.

2.4 Intonation units in the corpus

I have stressed in numerous places that I consider the division into IUs (or other similar prosodic phrases) to lay the foundation for any subsequent linguistic analysis of natural discourse. In this study, the prosodic analysis provides a framework for the syntactic and discourse pragmatic analyses in Chapters 3 and 4. At this point, the results of the prosodic analysis of the corpus will be presented and briefly discussed. Numerous aspects of the characterization of IUs need not be brought up at this time, but will surface in subsequent chapters. In addition, the mechanics of the prosodic analysis is not examined further here; a sample excerpt (57 IUs long) is analyzed in detail (from all three perspectives) in Chapter 5, so that the major considerations involved in such an analysis are fully exemplified there. I will, however, address questions about the reliability and validity of auditory analysis at the end of this section, and I will suggest how some legitimate concerns can be overcome through a prototype model of IUs.

2.4.1 Description of corpus IUs

In the corpus database there are 1991 IU entries, of which 1736, or 87%, form the main corpus of utterances upon which most of the analysis in Chapters 3 & 4 will be based. The typology of the whole database which
leads to this subset is depicted in Figure 2.6.

```
IU lines
  Non-Native       Native
    49 (2%)        
  Non-Vocal       Vocal
    84 (4%)        
  Non-Linguistic  Linguistic
    104 (5%)       
                    Uncertain
    18 (1%)        Certain
                    1736 (87%)
```

Figure 2.6. Types of IU entries in the corpus database (N=1991).

In the most general sense, 'IU' refers to one line of the transcript, as each IU is placed on its own line (hence 'IU lines'). The first distinction made is between 'non-native' and 'native' IUs. As there were two non-native speakers present for one speech event (see Table 1.1 for details), the 49 IUs (2% of the total IU lines) by these speakers—half of which are backchannel utterances—are treated separately from all others.

The next two distinctions concern IU lines which contain no linguistic material, and so should not be tallied in the same way as the remaining IUs. These IUs are of two main kinds, 'nonvocal' and 'nonlinguistic'. There are some IU lines that are simply notational artefacts, as when timing measures and comments occupy a line by themselves. Pause indicators may be placed alone on a line (instead of at the start of the following IU) if the pause cannot be associated with the talk of any person. Nonverbal events that are part of the interaction (inserted as bracketed comments by the analyst) may also be placed on separate line, usually with some indication of
the timing. These lines constitute ‘nonvocal’ IUs, of which there are 84 instances (4% of the total IU lines) in the corpus. Transcript lines may also comprise speaker vocalizations and sounds which are ‘nonlinguistic’ in nature, including laughter (all kinds), inhalation and exhalation, coughs, and throat clearing. Also grouped together with the nonlinguistic IUs for the purposes of tabulation are any IUs containing only ultra-short nonsyllabic segmental vocalizations which cannot be interpreted with any certainty (only 2 such cases were found). There are 104 cases of ‘nonlinguistic’ IUs in the corpus (5% of the total IU lines). Although there may be a rational basis for considering some nonlinguistic IUs equivalent to an IU containing linguistic material (e.g. the interturn pauses in an interactional analysis), including these IUs in the calculations would only skew the figures, since they automatically must be placed in ‘other’ categories. They are, of course, always included in the transcriptions, and they do figure in textual interpretation.

Finally, there are 18 IUs (1% of the 1858 ‘vocal’ IUs) which contain uninterpretable words. (Each syllable of such words is transcribed with an X in sharp angled brackets, i.e. [ <X> ].) These few IUs are distinguished from the other linguistic IUs.

Thus, all totaled, the number of ‘IUs’ not included in the following tabulations equals 255, or 13% of the 1991 IU lines in the corpus database. The remaining 1736 IUs constitute the ‘primary corpus subset’.

Let us proceed with a characterization of the IUs in this corpus subset. As characterizations which involve syntactic units will be covered in
Chapter 3, and those which involve considerations of information flow will be covered in Chapter 4, it is appropriate here to examine the nature of IUs containing linguistic material in other terms. One possible measure is the duration of IUs in seconds. Altenberg (1987: 17) states that the size of prosodic phrases in terms of length and speed has been 'little investigated', so this might seem useful. On the other hand, absolute duration may not be useful parameter, since it will not necessarily give an accurate picture of the amount of linguistic material in an IU. An IU of long duration uttered at a slow rate may contain the same number of words, for example, as a short IU uttered at a fast rate. A more informative measure is the number of syllables per IU. While the absolute duration will not be taken into account with this measure—depending on the rate of speech, two IUs having the same number of syllables could differ greatly in their duration—it does gauge the size of the IU in linguistic terms. (Knowing the number of words per IU is, of course, potentially informative, too. This measure will be discussed in the next chapter.)

Furthermore, as discussed in Chapter 1, orthographic conventions were chosen in part to facilitate calculations with syllables, so this is where I will begin. Figure 2.7 below presents the number of IUs in the primary corpus subset in terms of the number of syllables per IU. Several things are worth noting about the distribution depicted in the figure. The most striking characteristic is the obviously high number of IUs which are two syllables in length, i.e. the mode value is 2, while the median and mean values are obviously higher (the median is 5; the mean, 5.2 ± 3.4). I will return to this fact shortly.
First, notice that the frequency of IUs as a function of syllables per IU is not distributed evenly. Two groups can be discerned. Ignoring the large number of two-syllable IUs for the moment, one group includes IUs of 1-8 syllables. The frequency of each of these IUs differs little. A second group includes IUs which contain 9 or more syllables. As is evident from the graph, the drop in frequency in this group is quite rapid as the number of syllables increases.\textsuperscript{54} In other words, it is common to find IU of up to eight syllables (this group constitutes 85\% of the IUs), but IUs with more syllables become increasingly infrequent. While it would be a mistake to say that there is an absolute upper limit to the number of syllables that an IU may contain, it is clear that IUs of more than twelve syllables are fairly rare (3\%)

\textsuperscript{54} It is nearly exponential ($r^2 = 0.972$).
of the IUs). This distribution will be relevant to the discussion of information flow taken up in Chapter 4.

Returning now to the low modal value of the distribution, an explanation for the discrepancy between the mode and median/mean values can perhaps be found if the contribution of speaker turns is taken into consideration. Familiarity with the corpus (indeed with virtually any conversational material) makes it apparent that there are many short IUs in the corpus (primarily 1-2 syllables) due to the presence of backchannel utterances such as ja, ja ja, and mhm. Given this observation, as well as the obvious functional difference between utterances that constitute turns and those that only minimally acknowledge attention, agreement, surprise, and so on, it is legitimate to factor in speaker turn status. Figure 2.8 continues the typology of IUs begun in Figure 2.6, with speaker turn added.

The IUs that are referred to as ‘turn proper’ are those which constitute all or part of a speaker turn; ‘backchannel’ refers to those which constitute utterances by a speaker other than the one holding the floor. There are 182 linguistic backchannels (9% of the total IU lines) in the primary corpus subset. Figure 2.9 below replots Figure 2.7, with the backchannels as a separate category.
From this new figure, we can observe that, indeed, (linguistic) backchannel utterances are usually 1-2 syllables in length. We can also observe that, while the frequency of backchannels 2 syllables long reduces the great predominance of this value for the other IU group, an IU of 2 syllables still represents the mode value, so the contribution of speaker turn status is only part of the explanation for the discrepancy between the mode and the median/mean originally noticed.

Another functional distinction that may have some correlation with the length of IUs is the contour type of the IU. Recall that four contour types are distinguished through a combination of final pitch direction and transitional continuity (with a fifth category of 'uncompleted' (truncated) IUs). Figure 2.10 below graphs the breakdown of the primary corpus subset into these five types.
Figure 2.10. Frequency of IUs in the primary corpus subset (N=1736) as a function of syllables per IU, by contour type.

In the figure, the values for the noncomplete contours (/+ and \+) are plotted with a small cross ('+'), and the complete contours (/* and \*) are indicated with a filled circle ('*'). The final pitch types are likewise grouped, with values for the nonfalling contours (/+ and /*) connected by a dot-dash line, the falling contours (\+ and \*) by a solid line. Truncated contours (') are represented with open squares ('•') connected by a broken line.
A couple of points are worth noting in this figure. Most salient is that the contour types clearly differ in their frequency of distribution in terms of syllables per IU. While the early peak at 2 syllables/IU is still apparent for the noncomplete contours, there is no corresponding peak in either of the complete contours. Moreover, the mode value for the falling noncontinuing type is similar to the median and mean (see Table 2.3). The remaining discrepancy between the modal and median/mean values discussed above would therefore have to be sought in the profile of noncomplete contours. The truncated contours also show a peak at 2 syllables/IU, thus this type contributes to the preponderance of IU of this length as well. The median for this type, however, is also 2 syllables/IU, so this is a typical size for truncated contours.

<table>
<thead>
<tr>
<th></th>
<th>Continuing</th>
<th>Noncontinuing</th>
<th>Truncated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nonfalling</td>
<td>Falling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>/+</td>
<td>+</td>
<td>/\</td>
</tr>
<tr>
<td>Mean</td>
<td>6.0</td>
<td>5.6</td>
<td>4.8</td>
</tr>
<tr>
<td>±</td>
<td>3.3</td>
<td>3.2</td>
<td>4.1</td>
</tr>
<tr>
<td>Median</td>
<td>6</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Mode(s)</td>
<td>4 &amp; 8 &amp; 2</td>
<td>2 &amp; 4 &amp; 8</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2.3. Descriptive statistics for the distribution of contour types (in syllables/IU) corresponding to Figure 2.10.

55 Incidentally, the overall similarity between the two noncomplete contour types (vs. the falling complete contour vs. the nonfalling complete contour) could be taken to constitute evidence for distinguishing exactly three basic types, as has traditionally been the case (e.g. von Essen 1956), and as is done in Du Bois et al. (1992, 1993).
Despite the remaining early peak (which is not so dominating in the nonfalling continuing contour type), the rough plateau from 2 (or 4) syllables/IU to 8 syllables/IU and the steep decline in frequency after that which was pointed out earlier is accentuated for both noncomplete contour types.

At this point, little more can be added without going into the syntax and discourse pragmatics of IUs. Thus, further exploration of the early peak at 2 syllables/IU, as well as the smaller peaks at 4 and 8 syllables/IU in the continuing contours, must await the next two chapters, where the distribution will be discussed in conjunction with the syntactic composition of the IUs and their use in the conversational stories.

In the preceding section, I have outlined the types of 'IU lines' in the corpus database and have discussed a characterization of the IUs containing lexical material in terms of the number of syllables per IU. When the primary corpus subset is broken down further, it is readily apparent that IUs differ in length (expressed as syllables/IU) depending on their speaker turn status and on their contour type.

2.4.2 The use of auditory data and the problems of reliability and validity

One area of unfinished business has yet to be resolved. A systematic description of any of the various instantiations of 'prosodic phrase'—of which the IU is one—is often said to be problematic for three reasons: relevance, reliability, and validity. Now, some of the criticism regarding prosodic phrases hinges more on the semantic and cognitive claims made about these units—discussion of these points will be taken up in Chapter 4,
where I examine the informational issues pertinent to the present analysis. But as far as the relevance of prosody to discourse is concerned, I believe I have provided ample justification of the need for a prosodic unit approximately the size of an IU in interpreting the prosodic structure of discourse. On the other hand, I have not discussed the issues of reliability and validity. It is to these points that I now turn.

There is the general view shared by many linguists—due perhaps in great part to Lieberman (1965, 1967)—that the methodology commonly used to segment speech into prosodic phrases is hopelessly unreliable and lacks validity. Numerous researchers since then have commented on various inadequacies associated with the auditory analysis of prosodic phenomena. Pierrehumbert (1980: 13) is typical: ‘Transcription by ear of intonation is the source of so many gross errors in the literature that we feel it cannot be relied on’. Thus, many linguists have been skeptical about any use of auditory data, insisting that if they are to have any value at all, they must be verified either by experiment or machine. Brown, Currie, and Kenworthy (1980, p. 48) reflect this view in justifying the necessity of acoustic as well as auditory data: ‘The dangers of relying on auditory analysis alone are well known’. Such overgeneralizations are, however, due in part to a misrepresentation of Lieberman’s (1965) comparison between auditory analysis and the acoustic record. The conclusions he draws were not meant to apply to all forms of auditory analysis, but only to the system of intonational transcription he investigated: the Trager & Smith (1951) representation of pitch levels, degrees of stress, and terminal juncture. In fact, he notes that a tonetic notation made for comparative purposes fared much
better in the experiment. It may be suspected that the statements by some researchers (e.g. Crystal 1969; Chafe 1980b) that segmenting spontaneous speech into prosodic phrases is ‘unproblematic’ have not contributed to a feeling of certainty among those who have tried their own hand and ear from published materials. Cruttenden’s (1986: 36) comment concerning opposing views of the ease of auditory analysis, that ‘... the truth lies somewhere in between ...’, is more helpful in building confidence, as it takes into account some of the real problems faced by any analyst without rejecting well-serving methodologies. I will emphasize the point here: identifying IUs requires keen and systematic attention to the full range of prosodic cues, but the analysis is quite possible.

Regarding the question of reliability, there is good evidence that native speakers can segment connected speech consistently. Kreckel (1981), for example, finds that the segmentation of conversational speech by four untrained individuals was remarkably consistent. Working independently from the audio recordings, transcribers agreed with each other to a high degree. Significantly, they also agreed with three trained analysts versed in the intricacies of tone unit identification.

There is also evidence that judgements can be consistently made for a given transcriber, i.e. intratranscriber variation is minimal. For the present corpus, some specific figures regarding the reliability of my own coding are available. I resegmented eight of the stories in the corpus after a period of at least six months of not having worked with the texts (in a few cases the
time frame was closer to a year). Out of a total of 867 IUs, there were 48 (6%) significant adjustments in IU boundaries: 26 (3%) new boundaries were drawn by splitting original IUs, and 22 (3%) boundaries that had originally existed were erased by joining IUs that had been separate. I take these results as support for the view that the auditory assessments of the kind I outlined in §2.3 can be made very consistently.

Regarding the question of validity, there is also strong indication that the system of discourse transcription that is presented in Du Bois et al. (1992, 1993), which I have adopted here, can make a claim to representing actual acoustic events. As mentioned in the previous section, Schuetze-Coburn et al. (1991) demonstrate that transcribers in nearly all cases transcribed an IU boundary when an $F_0$ reset occurred (pitch reset being one of the cues to a new unit). In Schuetze-Coburn (in progress a), I show that the transcription of another important cue—pause—corresponds in the vast majority of cases to silence in the acoustic record. I conclude, then, that given an appropriate analytical framework, both reliable and valid auditory transcription of prosodic features is possible.

Let us return a moment to Brown, Currie & Kenworthy (1980). Evidently, part of the problem that these researchers run up against is their expectation about the nature of prosodic units. They claim to have run into

---

56 'Aldi drunk', 'Getting by', 'Weavers' fair', 'Visa for India', and the four stories in 'Afternoon tea'.
57 There were also 12 'minor adjustments', having to do mainly with the placement of non-lexical material.
a series of difficulties in a 'very substantial number of cases' (40-1):

In prose read aloud most readers will, to a large extent, agree on the tone group organization of the text and will produce tone group units which are comparable to those of other readers, especially where punctuation in the text clearly marks the bounds of such units. In spontaneous speech, however, there are many occasions where no clear indication is given of tone group boundary markers.

They mention the following specific points which presented the most consternation (41-4):

a) pause in the middle of a grammatical unit
b) pitch movement on a grammatical item or filled pause
c) pause dividing what would have been a coherent intonation contour
d) lack of syntactic ambiguity resolution by the intonation contour

The first three problems they tie to planning difficulties on the part of the speaker. Whenever speakers were having a hard time articulating what they wanted to say, nonfluent speech would result, resulting in non-canonical prosodic structure that could not be easily segmented into tone groups. While Brown, Currie, and Kenworthy do pinpoint a source of uncertain boundaries—it is exactly these places that presented the most difficulty in the present corpus—they tend to overemphasize the problem. They themselves state that '... in many cases one is able to assign tone group boundaries with some confidence' (41). More importantly, though, they worry too much about the resulting syntactic or semantic consequences of the segmentation. If it is true, as they state, that they are inter-

168
ested in identifying the prosodic units through phonetic features, then they should have little concern about the syntactic and semantic make-up of these units, at least at the segmentation stage of the analysis.

Brown, Currie & Kenworthy further state: 'If we accept that tone groups are realised by phonetic features which include pitch and pause phenomena, we will necessarily keep encountering what appear to be conflicting criteria' (44). Their skepticism regarding the possibility of segmenting speech is obviously tied to their (mistaken) expectation of finding fixed phonetic criteria for delimiting tone groups in every instance. Given the way tone groups have been defined in the literature and lack of attention that has been paid to the boundary problem, especially in Halliday's influential work, it is perhaps not surprising that they would expect to find a better correspondence then they reported. On the other hand, prosody as a whole constitutes such a complex set of phonetic phenomena that it also seems unreasonable to expect any simple correspondence between features and unit boundaries. A number of years ago, Fry (1968: 371) stated what should be obvious today:

It is one of the basic facts about speech and language that there is never a one-to-one correlation between linguistic units and the characteristics of the articulatory, acoustic or perceptual patterns with which they are associated.

When the requirements are stated in an alternate fashion—in terms of a prototype—the 'problem' of conflicting criteria vanishes (see Schuetze-Coburn 1992; in progress c). Individual features are no longer to be viewed negatively (i.e. in conflict) when they do not coincide, but rather positively
(i.e. as complementing) when they do. This situation is not unlike the problem faced by dialect geographers. Individual isoglosses do not generally define (major) dialect boundaries, but bundles of isoglosses can serve to identify them. Similarly, relying solely on intonation contour (or pause or any other single phonetic feature) paints a distorted picture of the complex realization of prosodic phrases in natural speech and sets up the expectation of the possibility of determinate measures for a perceptual category. Defining a prosodic phrase in terms of only intonation contour, pause, and syntax to unambiguously demarcate boundaries makes use of only a small part of the information that speech act participants have at their disposal to segment and interpret the speech stream.

2.5 Summary

This chapter began with a detailed examination of the terms 'prosody' and 'intonation', and with a consideration of the idea of prosodic units in spontaneous speech. I argued that some conception of prosodic phrase is necessary (and valid) for the analysis of natural discourse, but that it need not match those that have been proposed for the analysis of other types of language data. In the second section I evaluated various types of prosodic phrases that have been put forth in the literature in terms of their basis of segmenting speech and their hierarchical sequencing. I then discussed the characteristics of the prosodic phrase used in my analysis, the 'intonation unit' (IU). Section three provided a detailed examination of various types of prosodic and nonprosodic features that have figured in identifying prosodic phrases. I concluded that pitch and timing features play the most
important role in determining IUs, but that numerous other features contribute to IU boundary perception. I also discussed the problem of representing these features in a discourse transcription. In the final section I outlined some of the characteristics of the IUs in my corpus of conversational German, and I addressed the problem of the reliability of transcription.
3. Discourse Syntax

Eine allgemein anerkannte Definition des so allgemein gebrauchten Wortes giebt es in der That nicht. Nach einer solchen, die allseitige Anerkennung finden könnte, also nach der richtigen, zu suchen, wäre ein Bemühen, das ziemlich müßig erscheinen müsste, wenn es sich dabei nur um einen Wortstreit handelte. Es kommt aber nicht darauf an, einen thatsächlich feststehenden Begriff mit dem (oder den) passendsten Worte(n) zu bezeichnen, sondern darauf, den Begriff selber richtig zu bestimmen.

There is indeed no generally recognized definition of a so commonly used word ['syntax']. To look for one that would find broad acceptance, and thus be the correct one, would be an undertaking that must appear rather superfluous if only a semantic dispute were involved. However, it is not a matter of labeling a well-established notion with the most suitable word or words, but rather of correctly determining the notion itself.

John Ries
Was ist Syntax?, 1894, p.1

In this chapter I examine the production of syntactic structure in spoken discourse by describing the basic patterns which occur in my corpus of German conversation and by contrasting the relationship of these patterns to the discourse prosody, as outlined in Chapter 2. Unlike standard syntactic accounts, the orientation of the investigation is toward the process, rather than the result, of speaking. In other words, discerning the interlocutors' construction of the syntactic form of their utterances, over time, and relating these constructions to the prosodic structure described in the previous chapter, constitutes the primary objective of the present chapter.

First, I discuss the question of syntactic units in general and consider the role that 'word', 'phrase', 'clause', and—above all—'sentence' might play in the grammatical description of spoken language. In the discussion termi-
nological issues specific to German are taken into consideration. It is determined that 'clause' best approximates the syntactic unit which meets the goals of the analysis, though it is also evident that this unit by no means covers all cases. Second, the definition of 'clause' and problems associated with clause delineation are examined. In this context, the notions of 'valency' and 'topological fields' are introduced. Third, I present an innovative system of marking clause structure in natural discourse as a means of describing the speakers' incremental construction of clauses. I then report the distribution of these 'clause constructions' in the corpus. Finally, the basic syntactic structure is contrasted with the basic prosodic organization of the stories by examining where units in each dimension converge and where they diverge.

3.1 Basic syntactic structure

The principal question I address in this section is: Which hierarchical level(s) of syntactic structure should be attended to in an analysis of spoken discourse? The answer is by no means obvious, for several reasons.

First of all, in examining language in actual use, we are treading into an area of language that mainstream syntacticians have relegated to researchers in other fields, such as anthropology, psychology, and sociology. Yet, from the linguist's point of view, such researchers often seem to adopt fairly simplistic approaches to analyzing linguistic structure. I argued in Chapter 1 that many linguists steer clear of discourse and corpus-based linguistics—with all their attendant 'performance' issues—for spurious reasons. As an additional example of the dubious argumentation used to ex-
clude the structure of language use from serious linguistic study, the much-discussed 'errors' that it is fraught with can be mentioned.¹ Coseriu (1988: 216) points out that our perceptions about concrete speech hinge on our assumptions regarding syntactic units:

... weil man ... die syntaktische Konstruktionen unzulänglich rechtfertigte, konnte man zu dem Glauben kommen, daß die Abweichungen die Regelmäßigkeiten übertreffen und daß man in der Performanz so gut wie nur Abweichung feststellen könne. (because syntactic structures have not been adequately justified, one could come to believe that deviations outnumber regularities and that practically only deviation can be found in linguistic performance.)

In other words, when discourse is analyzed by means of an uncritical application of the units of traditional structural linguistics and then judged against these norms, a distorted picture of language use is bound to result.

Also significant is that in systematically dealing with connected, spontaneous speech, we are entering territory that has not yet been thoroughly surveyed. According to Linell (1982: 34):

... it must be said that linguists (and other scholars) have worked much too little with spoken language (much less than we think we have), and have not sufficiently probed the question of what it really means to analyze speech in its own terms, i.e. as a temporally, dynamic behavior.

¹ Labov (1966) calculates that fewer than 2% of utterances are actually 'ungrammatical' (cited in Taylor 1984: 109).
Although our understanding of discourse in the decade since this statement was made has improved, and sensitivity to the issues involved here has generally been heightened, it is only quite recently that linguists have begun to address the deficiency that Linell draws our attention to (e.g. Ono & Thompson 1993). In view of a shift in focus from written to spoken and from object to process, one immediate question that is worth asking is to what degree the customary measures in the surveyor-linguist's analytic kit—forced specifically for the analysis of written language qua object—can find proper use in charting the syntax of actual utterances. Indeed, we need to ask, if we are to be confident that our analysis has any linguistic validity. Past practice amply illustrates the pitfalls of the uncritical overapplication of a linguistic unit. A prime example is the wide use of the term 'sentence'.

The unit 'sentence' is undoubtedly the traditional basic level of syntactic description and theory. Most syntactic analysis is undertaken with some set of sentences as the primary data, and grammars, too, are typically compiled using sentence data. Furthermore, researchers in other fields and in the 'applied' subfields of linguistics have adopted a similar reliance on the sentence. Witness publications in psychology: Sentence Production (Rosenberg 1977); in sociology: On the syntax of sentences-in-progress (Lerner 1991); in computational linguistics: Planning English Sentences (Appelt 1985); in acoustics: Fundamental Frequency in Sentence Production (Cooper & Sorensen 1981); and in many others. Yet how often do we encounter a discussion of the assumptions underlying the use of such a notion? As with many other linguistic items, all too often too little attention is paid to fundamental questions. Because justifying familiar grammatical concepts is
not *de rigueur*, the original motivation for a given concept can be easily overlooked or forgotten (or conveniently ignored). Thus, even though 'sentence' finds its widest application as a syntactic unit for the analysis of normative (i.e. written) language, it is such an integral part of so much linguistic analysis that many, if not most, linguists would not find it objectionable to talk about 'sentence production' in speech; nor would they find reference to 'the intonation pattern of sentences' at all peculiar. But given the many warnings of scholars, including those who have worked extensively with natural language data (e.g. O'Connell 1977; Crystal 1980) as well as those concerned with theoretical pragmatic issues (e.g. Levinson 1983), caution would have us refer instead to 'utterance production' and 'the intonation of utterances' or 'utterance-strings' in those instances where spoken language is meant, at least until a melding of views regarding the relevance of the notion 'sentence' to the analysis of spoken discourse has been achieved. Note that simply taking the polysemy of the term 'sentence' into account, i.e. by distinguishing 'text-sentence' from 'system-sentence' (Lyons 1977: 622), does nothing to alleviate the problem inherent in the use of the term.

Another reason why the issue of hierarchical level needs to be broached is that any decision concerning syntactic units is intricately connected to the framework of one's analysis, as it is dependent on the assumptions and objectives of the researcher. In the present study I wish to avoid, as much as possible, an uncritical acceptance—or rejection—of terms like 'sentence' in analyzing the syntactic structure of spontaneous speech. And so before addressing the ultimate goal of this chapter—explicating the speakers' con-
struction of syntax in discourse and its relation to the prosody of the utterances—I tackle the issue of basic syntactic structure.

3.1.1 Levels for a prosody–syntax comparison

It should be clear from the goals outlined in Chapter 1 that the task at hand is not to put forth a complete structural description of spontaneous utterances, but instead to find an appropriate inroad into the syntax. Put simply, a basic-level account is needed in order to evaluate the grammatical dimension of the stories, and painting a picture of the speakers' frequent syntactic patterns will suffice. Thus, by 'basic' I do not mean 'primary' or 'prime' in the sense of 'not derivative', but rather I mean 'focal'. The question thus narrowed becomes: Which level(s) need we focus on in a comparison of the syntactic with the prosodic (and in the following chapter, with the discourse-pragmatic)? Clearly, the objective is twofold: on the one hand, the speakers' use of syntax to structure their talk must be accounted for; on the other, the scope of the units in the various dimensions being investigated must be matched as closely as possible to facilitate such comparison.

In pursuing these objectives I assume that reference to hierarchical syntactic levels will be useful. Of course, despite the claim that '... hierarchy is the organizing principle of surface structure' (Longacre 1976: 255), the notion of hierarchical structure can itself be questioned. In fact, Hudson (1984) takes a general position on syntactic hierarchy that is similar to the view expressed above regarding 'sentence'. He speculates: '... it is at least tempting to suspect that the attractions of hierarchical analysis for linguists lie in
its relevance for written rather than spoken language’ (249). Considering that written language has a more ‘integrated quality’, while spoken language is more ‘fragmented’ (Chafe 1982), this speculation may be quite well-founded. In any case, we may at least wonder about the relevance of syntactic models which posit (or allow for) the opposite, i.e. extreme hierarchical differentiation. For example, in $\bar{X}$-syntax (Jackendoff 1977) a phrase may have as many levels as it does elements. Since there are additional reasons to question the necessity of a multiplicity of levels—e.g. von Stechow & Sternefeld (1988) doubt the need for any syntactic structure that is based solely on semantic reasoning—I will favor representing a minimum of complexity within syntactic units, following other approaches which are oriented around surface constituency (e.g. Eisenberg 1989; Schlobinski 1992).

As for interunit complexity, there are several hierarchical levels common in syntactic analysis to be considered in answering the question posed earlier. Of course, varying sets of units have been employed by different researchers. Engel (1982), for instance, distinguishes between word, phrase, sentence, and text. The ‘ranks’ of Halliday (1985) include morpheme, word, group, clause, and sentence. Longacre (1976) additionally considers stem (between morpheme and word) and paragraph (between sentence and discourse). Thus, there is no standard set or number of levels.

Among the possibilities offered in the literature, four syntactic units come into consideration for the purposes of a comparison with the basic prosodic structure outlined in Chapter 2: word, group, clause, and sen-
tence. While syntagmatic relationships between elements can be recognized at all levels, units at both poles can be safely ruled out a priori. On the basis of even a casual familiarity with IUs, the largest and smallest units are obviously not appropriate. The lower hierarchical levels—morpheme and stem—cannot be considered, as it is clear that speakers' utterances normally consist of sequences much larger than these units. It is equally clear that the higher levels—paragraph and text—do not correlate with IUs (except perhaps in possible minimal cases not represented in the corpus). We have, however, seen examples of IUs corresponding to words, groups, clauses, and sentences, in a traditional sense. Thus, it is sensible to examine the status of each of these units as a basis for a description of discourse syntax. For reasons that will become clear below, however, I will discuss the occurrence of 'sentences' apart from the other levels; in addition, I introduce the category 'complex', which I mean to be a level between clause and paragraph.  

The correspondence between IUs and the viable hierarchical levels is an empirical question—although one which does have important theoretical implications. Hypotheses concerning the speakers' expression of semantic units, the discourse-pragmatic structuring of stories, and the interactional

---

2 I will follow Halliday and use the term 'group' rather than 'phrase' to refer to a level between word and clause. While 'noun group' and 'noun phrase' are basically synonymous (though see Du Bois & Schuetze-Coburn 1993 for a specialized use of the former), the controversial constituency of 'verb phrase' (especially for German) suggests electing 'verb group' instead.

3 Cf. the term 'multi-clause combination' (Ono & Thompson 1993), following Halliday's 'clause complex'.
organization of conversation bear upon the choice of analytical units. Thus, both empirical and theoretical considerations must be balanced.

At this point, in order to gain a working knowledge of the syntactic scope of IUs, let us examine a characterization of IUs in terms of words, groups, clauses, and complexes, with a view at sentences as well.\footnote{Cf. the similar undertaking in Tao (1993), although his categories differ substantially from mine.} Since the immediate objective here is merely a rough characterization, I will not dwell on the definitions of these units for the moment. As a first approximation, I will take a pretheoretical stand and use intuitive notions based on fairly traditional concepts. However, I adopt one nonstandard constraint in part to simplify the discussion: the syntactic characterization of each IU is assigned the lowest possible rank in the hierarchy, thus excluding mutual category membership. In other words, the syntactic contents of each prosodic unit is counted as belonging to only one of the four categories, minimally a word, maximally a complex.

For the purposes of a tabulation, by ‘word’ I mean a single lexical item as determined by form class and usually represented in the discourse transcription as an orthographic unit. (Thus, reduced or contracted forms and conventionalized backchannel utterances such as \textit{mhmm} are also treated as lexical words). A ‘group’ is a sequence of words containing a (head) word from a major form class (noun, verb, adjective/adverb, preposition) together with its dependents and accompanying minor forms, which usually functions as a constituent part of a larger structure. A ‘clause’ is taken to be
a sequence of groups containing a single verb group and all associated (dependent) words or groups. A 'complex' is a sequence containing two (or more) clauses. I take a 'sentence' to be a clause or complex which has all its grammatical arguments overt and which does not show syntactic dependency on another clause or complex. These definitions are not intended to be generally applicable, but are perfectly adequate for a simple IU characterization of a corpus of German discourse. In other schemes, it is, of course, common for forms to belong to more than one hierarchical level. For example, IU 279 in (1) comprises the form stimmt, which would be viewed simultaneously as an instance of the category verb (i.e. word), as a verb group, and as a clause (but not a sentence under the above definition) in a phrase structure representation.

(1) Corpus excerpt from 'Propane tank' [Morning visit: 274-9].

274 T: [ <als ich gegrillt habe> \•
   when I grilled have
275 K: [ das nich' an der Schildischer da unten /•
   that not on the S there below
276 T: (0) jà= /•
   yes
277 genàu= \•
   exactly
278 * (0,9)
279 T: stimmt \•
   agrees

274 T: [when I grilled.
275 K: [not that down on Schildische <street>?
276 T: yeah,
277 exactly.
278 * (PAUSE)
279 T: <that's> right.
But since forms and sequences are tabulated minimally here, I will view IU 279 as containing simply a word, like IUs 276 and 277.5

<table>
<thead>
<tr>
<th>Word</th>
<th>Group</th>
<th>Clause</th>
<th>Complex</th>
<th>Other</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>355</td>
<td>184</td>
<td>462</td>
<td>17</td>
<td>718</td>
<td>285</td>
</tr>
<tr>
<td>(20%)</td>
<td>(11%)</td>
<td>(27%)</td>
<td>(1%)</td>
<td>(41%)</td>
<td>(16%)</td>
</tr>
</tbody>
</table>

Table 3.1. General syntactic characterization of IUs in the primary corpus subset (N=1736).

Based on the above definition of word, group, clause, complex, and sentence, Table 3.1 gives the rough and ready syntactic characterization for the 1736 IUs in the primary corpus subset.6 The column labeled ‘Other’ tallies IUs which do not comprise any single unit. The column labeled ‘Sentence’ represents a separate tabulation, as sentences constitute a subset of clauses and complexes. Four things are noteworthy about this classification: (a) over two fifths of the (linguistic) IUs in the corpus do not fit into this simple scheme (i.e. a plurality of cases fall into the ‘Other’ category); (b) among the five hierarchical groups (including sentences), clauses are most prominent; (c) sentence is the third most frequent category; and (d) neither clause nor sentence approaches a majority.

The frequent noncorrespondence between IUs and basic syntactic units suggests that a modification of the classification scheme is necessary before an adequate picture of the syntactic contents of IUs can be obtained. While

5 Note that in this particular case, the formulaic nature of stimmt is judged to play some role in this decision, but in general the quantitative difference of such forms from ‘full’ groups or clauses suffices to categorize them.

6 This subset of the corpus database constitutes all IUs containing any lexical material. For the typology of the database, see Figures 2.6 & 2.8 in §2.4.
a substantial portion (41%) of the IUs do not correspond EXACTLY to a single word, group, clause, or complex, upon further examination it is apparent that quite regularly these IUs ALMOST comprise single syntactic units. For example, IU 209 in (2) consists of a noun phrase group which, together with the groups in the previous IU, forms a clause. IU 208, then, comprises a clause minus a constituent group. IU 206 illustrates a case of one too many elements; it contains two units: a clause and a word form ne.

(2) Corpus excerpt from ‘Cello lessons’ [Afternoon tea: 202-9].

202 [E]:  (H) früher /+
          earlier
203       .. (%)...«8»~hat er ja Célllo geleért \ ne /+
          has he yes cello learned
204       [ (H) ~un' hatt' ne gänz tolle Cèlllolehrerin /+
          and had a completely great cello.teacher
205 (S):  [ ja \+
          yes
206       ~die kàm aus Ìndien \ ne /+
          she came out of India ok
207       ~un' die wàr [ also ] /+
          and she was so
208 (S):  [ hm ] /+
209       .. %àuch so 'ne gänz hòchbegàbte Fràu /+
          also such a completely high.talented woman

202 [E]:  earlier,
203       he learned cello, right,
204       [and had a really great cello teacher,
205 (S):  [yeah;
206       she was from India,
207       and she was,
208 (S):  [ hm,
209       also a really talented woman,
Thus, allowing for cases where an IU essentially comprises a unit, the syntactic content of the IUs can be more accurately described. Under this modification, reference is made to a three-value scale for each syntactic unit: an IU corresponds to (a) exactly one unit, (b) a unit plus a form (or forms), or (c) a unit minus a form.

<table>
<thead>
<tr>
<th></th>
<th>Word</th>
<th>Group</th>
<th>Clause</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- = +</td>
<td>- = +</td>
<td>- = +</td>
<td>- = +</td>
</tr>
<tr>
<td>7</td>
<td>355 138</td>
<td>25 184 149</td>
<td>292 462 97</td>
<td>10 17 0</td>
</tr>
<tr>
<td>(0%) (20%) (8%)</td>
<td>(1%) (11%) (9%)</td>
<td>(17%) (27%) (6%)</td>
<td>(1%) (1%) (0%)</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>358</td>
<td>851</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>(29%)</td>
<td>(21%)</td>
<td>(49%)</td>
<td>(2%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2. General syntactic characterization of IUs in the primary corpus subset (N=1736), allowing for partial alignment.

Table 3.2 gives the results of the modified classification. Again, the total number of IUs represented is 1736, the primary corpus subset. In the top half of the table, exact matches between the prosodic and syntactic structures are listed in the middle column (labeled '=' for each hierarchical level. (These values correspond to the values in Table 3.1.) The '-' and '+' columns indicate IUs with syntactic units missing a form or constituent, or containing an additional one. Combined figures for each unit are given in the bottom half of the table.

From Table 3.2 we see that about half of the IUs contain sequences which constitute, or almost constitute, clauses. Between a quarter and a third contain only a word or two, while over a fifth contain a group (or a group and a form). These patterns are all more frequent than the occur-
rence of sentences, which comprise just one sixth of the IUs. On the rare side in the corpus are IUs which contain only a partial word or group, as are those containing a full or partial complex. Thus, it is evident from this general syntactic characterization that IUs do vary greatly in syntactic size, but that the range is basically from full word to sentence, with most IUs containing either a full clause (27%) or a word (20%); IUs containing the major part of a clause (17%) also occur with some frequency, and sentences (16%), too, are not uncommon. The empirical evidence, then, suggests that of the five syntactic units considered, ‘clause’ should be used as the basic syntactic level for the proposed comparisons.

As stated above, the numerical plurality of any one unit in a given corpus does not decide the issue conclusively, but is merely one consideration, albeit one that speaks fairly cogently in this instance for a focus on the clause. Still, we should explore the theoretical feasibility of adopting any of the units in order to gain an understanding of the issues involved. Furthermore, we still need to examine the special role that ‘sentence’ has played in grammatical description, given the particular terminology for this unit in German scholarship, and considering the controversy of this unit in spoken language research. I start with an examination of ‘word’ and ‘group’, and then move to ‘clause’ and ‘complex’, where I discuss ‘sentence’ as well.

---

7 Interestingly, Crystal (1969: 258) found that 28% of the tone units in his (British English) corpus were coextensive with clauses.
3.1.2 Word and group

The level 'word' is in some respects the least likely candidate for basic-level status in the sense developed here. Although it has been one of two focal levels since Alexandrian times (Robins 1979: 33) and is the 'terminal' or base level of phrase structure models, modern preference strongly leans towards notions of syntactic-semantic function with respect to larger structures in syntactic descriptions: syntactic rules are designed to operate not on words, but on higher organizational units which exhibit particular functions as constituents. Nevertheless, 'word' remains a theoretically viable alternative; it is, for example, the only syntactic level in some models of dependency grammar (e.g. Hudson 1984).

The advantage of word as a basic level for a syntactic analysis of discourse lies in its closeness to a 'surface' representation of the utterances. All linguistic analysis of spoken language involves interpreting an acoustic signal (among other things) in terms of meaningful elements. In most transcriptions these elements are represented effectively as words of the language. As Hentschel & Weydt (1990: 13) put it: 'Das sprachliche Zeichen par excellence ist das Wort' ('the linguistic sign par excellence is the word'). Of course, a certain abstraction necessarily takes place in the process of representing words from the speech stream—the acoustic signal itself does not

---

8 I am appealing to intuition, tradition, and contemporary linguistic practice in using the term 'word' (or closely related terms like 'lexical item'). A precise, universal definition is just as unobtainable for 'word' as it is for 'sentence' — or any other basic syntactic unit. See Matthews (1981) on the general issue of word as a syntactic unit and Bergenholtz & Schaedler (1977) on word class in German.
contain words—but this is primarily a phonological and semantic operation.

![Waveform of an intonation unit.](image)

Figure 3.1. Waveform of an intonation unit.

Thus, the acoustic signal, here represented in Figure 3.1 by the waveform of an IU-long stretch of speech, consists of a sequence of acoustic events that can be represented as a series of sounds, phonetically something like:

(3)  \[k^\text{\^a}n\text{\^a}st\text{\^a}t\text{\^a}\text{\^a}z\text{\^a}t\text{\^a}z\text{\^a}n:\]\n
which might be perceived by a speaker of German as the segment sequence:\(^9\)

(4)  \(/k\text{\^a}n\text{\^a}st\text{\^a}t\text{\^a}d\text{\^a}t\text{\^a}\text{\^a}z\text{\^a}t\text{\^a}z\text{\^a}n\text{\^a}n/\)

corresponding to the morpheme sequence:

(5)  \(/k\text{\^a}n\text{\^a}nt\text{\^a}+\text{\^a}d\text{\^a}t\text{\^a}+\text{\^a}r\text{\^a}t\text{\^a}+\text{\^a}z\text{\^a}n+\text{\^a}n/\)
\[\text{can + 2SG.PRES.IND + you [NOM] + you [RFLX] + remember + INF}\]

corresponding to the words in IU 12 of (6) below.

---

\(^9\) Cf. Malmberg (1963: 168): 'The number of segments into which the non-segmented continuum is divided, depends on the pattern applied, not on the substance'.
(6) Corpus excerpt from ‘Project’ [Afternoon tea: 11-13].

11 E:  ~un' w's war'n noch für Projekte /*
       and what were still for projects
12  .. kannst 'u dich entsinn'n /*
       can you yourself remember
13  S:   ...(1.1) m= ...(1.8) die Luft /*
       m the air

11 E: and what other projects were there,
12  can you remember?
13  S:  m the air,

A focus on words—themselves the result of a prior interpretation—in effect eliminates the need to motivate additional grammatical units. Furthermore, the basis for a syntactic analysis is established with little reliance on syntactic abstraction, which on the word level enters the analysis only at a later point, when the relationships between words are considered.

A syntactic analysis, however, involves abstracting connections between words or groups of words. One of two primary principles of syntactic organization is that of dependency (Matthews 1981; Engel 1982). Under this principle, the dependency properties of pairs of words constitute the syntactic description (cf. Eroms 1985: 307). As an illustration of this type of analysis, one possible representation of dependency relations at the word level is depicted in Figure 3.2 below. The glossed conversational sequence is given in (7) below.
(7) Corpus excerpt from 'Vacation travel' [Uli's visit: 444-56].

444 [A]: ...(1.0) ~wir sind ei%e --
                we are onc-
445          .. zwei- --
                    twic-
446          .. zwêimal glaub' ich mit der Bahn gefahr'n \•
                    twice believe I with the train traveled
447          .. war gràuenhaft \•
                    was horrible
448          .. <MAR gràu=enhàft MAR> \•
                    horrible
449          .. (H) [àchtzehn Stunden hinter 'nander mit der]mit der Ba=hn \•
                    eighteen hours after one.another with the with the train
450 U:      [ is 'e Strapàze /+
                    is a strain
451          ja ja= ] \•
                    yes yes
452 A:      .. wie /•
                    how
453 U:      .. das is' ne Strapàze [da] \•
                    that is a strain there
454 (A):    [ <F òh F> ] \•
                    oh
455          ... und =g'rade /+
                    and just
456          ...(1.0) durch [ die Àlpen und so wèiter \ ne ] /+
                    through the Alps and so on ok

444 [A]:    we've (taken) onc--
445          twic--
446          twice I believe the train ().
447          was horrible.
448          horrible.
449          [eighteen hours one after another on the train.
450 U:      [it's rough,
451          yeah.
452 A:      what?
453 U:      that's rough.
454 (A):    oh!
455          especially,
456          through the Alps and all,
Figure 3.2. Dependency (a) and constituency (b) of a conversational excerpt
In part (a) of Figure 3.2, an excerpt from one of the stories is marked in a way to indicate which words are (morpho)syntactically dependent. In customary fashion, arcs connect pairs of words that exhibit some sort of dependency relation. Arcs with arrows are used here to show specifically form-based dependencies which rely on the morphological categories case, number, person, or gender. For example, the word form *sind* in IU 444 is grammatically contingent on the word form *wir* in terms of person and number, thus the directional arc from *sind* to *wir*. Other, nondirectional arcs indicate other relationships. Inverted arcs at the beginning of IUs symbolize a nonspecific contextual linkage.

Although the figure is meant to be illustrative rather than present a definitive analysis, it is readily apparent that this type of dependency-based representation more than adequately describes many syntactic relations that can be found within IUs as well as across them. (Other relationships could be represented as well.) Diagramming dependencies makes clear, for example, that the syntactic relations of the words in IU 453 are neatly self-contained in that IU, while a complex set of relations exists across IUs 444-6. A comparison of discourse prosody and syntax at the word level thus would, in the main, entail identifying the cases where syntactic dependencies are internal to IUs and where they cross IU boundaries. The strength of IU-internal and external relations could then be fairly objectively determined by reference to the type and frequency of dependency arcs.

---

10 In the figure, IUs 445, 446, 448, and 456 are indented to the right to be able to graphically depict the relationships between IUs.
What is not as clear is how reference to dependency relations can help elucidate the speakers' on-line construction of syntactic forms, if it is true that '... a surface structure is, by and large, generated from "right to left" as successive fragments of the message become available ... without much lookahead or backtracking ...' (Levelt 1989: 235). Part of the problem has to do with long-distance (morpho)syntactic dependencies. For instance, in IU 444 there are both local and long-distance dependencies. The local dependency 'chains' (series of arcs) are built up incrementally: the adjacent word (or word fragment) pairs *wir* and *sind*, *ei-* and *zwei-* and *zweimal* are linked, the latter pairs connecting IUs 444, 445, and 446 together. Local chains are built up in IU 446 as well, until the lexical verb *gefährden* is reached. At this point, several long-distance dependencies are realized, and the local chains become interconnected. Yet if the lexical verb regularly comes late in the IU—or worse, in a subsequent IU, as in the above instance—a word-by-word procedure cannot account for the correct morphological marking of arguments. In order for the grammatical case of arguments to be properly assigned, the specific lexical verb must be known.

A second problem concerns the syntactic nexus. Without the lexical verb—the final word of the IU sequence—there would be no connection among the locally constructed dependency chains. The effect of having no lexical verb can be seen in IU 449. This IU contains no lexical verb; nor is there one in an adjacent IU, as in the above case. The two dependency chains in IU 449 have no direct connection with other chains, and any indirect connection between them is not realized. Yet we may question whether in cases such as this the lack of structure resulting from the ab-
sence of a lexical verb in the utterance actually reflects the speaker's use of the syntax or is instead due to the limitations of the model.

The benefit of a dependency-based representation—avoiding unnessary hierarchical abstraction—is unfortunately also the source of a major practical disadvantage. The amount of detail required for this type of syntactic analysis exceeds by far the minimal representation desired for the purpose of a comparison with IUs. As syntactic relations for each word in the corpus must be indicated, there ends up being a plethora of individual relations to keep track of—not all of which are of interest here—and the representation soon becomes difficult to manage efficiently. Thus, it is not at all obvious what the advantage of such a tremendous analytical investment would be for the present study.

One way to reduce syntactic detail is to focus on the 'group' (or 'phrase'), the unit which results from the application of the second basic principle of syntactic organization, namely constituency.\(^\text{11}\) Constituency relations (with group as the basic syntactic level) prove to be more manageable than dependency relations (with word as the basic syntactic level), simply because there are generally fewer constituent groups than dependent word pairs in any given utterance. Although dependency chains can generally be converted into constituency groups (Matthews 1981: 84), a group-based representation permits the hierarchical structure within groups to be differenti-

\(^{11}\) It is interesting to note that while the analysis of utterances in terms of words and the relationships that hold between pairs of words is an ancient enterprise, a constituency-based analysis in terms of groups is a relatively recent innovation, traceable to Wundt and Bloomfield (Percival 1976).
ated—or not, as required—allowing generalizations to made about the top node of each group to the exclusion of lower nodes. The simplification gained is apparent from a comparison of the dependency (a) and constituency (b) diagrams in Figure 3.2. Among the words of IUs 444-6, thirteen dependency relations are indicated. In the same stretch, there are just four groups represented.

Moreover, in contrast to the word-level representation with its orientation on dependency relations, the group representation draws attention to the syntactic composition of IUs, and a focus on constituent elements allows for a characterization of IUs in terms of a small set of unit subtypes. Crystal (1969: 260) comes to a similar conclusion (approached from the opposite direction) in his examination of a large corpus of English: ‘... a more useful hypothesis about the nature of tone-unit structure would be to discuss it in terms of ELEMENTS OF STRUCTURE rather than of structures as wholes’. That is, he determines that constituents of clauses provide a better basis for the grammatical description of tone units than clauses themselves. Given that prosodic breaks within groups are reported to be uncommon, at least for English and German (Crystal 1969: 263; Du Bois & Schuetze-Coburn 1993: 249), this would appear to be true.

A problem, however, surfaces with segmenting utterances into sequences of groups classified according to lexical head without reference to higher level syntactic units. While some IUs consist of single, independent groups, it is evident that most groups are syntactically connected to larger

12 Counting interdependent word pairs (double-arrow arcs) twice.
reproductions. Since there is no indication of relationships among groups in a representation at this level, the only way to characterize an IU syntactically is to list the groups it contains. Again, we may ask whether this is a satisfactory result in terms of modeling the speakers' use of syntax in constructing their utterances. Some reference to higher-level units appears to be necessary.

3.1.3 Clause, complex, and sentence

Production models invariably refer to higher-level units: for instance, Levelt's (1989) comprehensive 'lexically driven', incremental model of grammatical encoding is sentence based. This is due, in part, to the qualitative difference of the levels 'clause', 'complex', and 'sentence' on the one hand, and 'word' and 'group' on the other. The former units are generally taken to be the means of syntactically expressing preverbal, propositional messages, while the latter form parts of the proposition, i.e. the particular action or state, the entities involved in the event, the specification of place, time, manner, and so on. It is above all this integrative aspect of the units 'clause', 'complex', and 'sentence'—expressing a 'Gesamtvorstellung' ('complete idea') Wundt (1911: 611)—which makes reference to these levels of structure attractive to philosophers, psychologists, and linguists. But

13 Of course, dependency, constituency, and functional representations can be combined—as is successfully done in Eisenberg (1989). But here, as before, the degree of detail involved is more like a full-scale syntactic representation than the basic one sought after. This problem does not affect Crystal's study, as he is interested in just this kind of characterization; however, note that his group subtypes ('subject', 'predicator', 'complement', 'adverbial', and 'vocative') are functional and presuppose a clausal analysis.
what linguistic substance lies behind the terms? In this subsection, I address this question.

The following examination of 'clause', 'complex', and 'sentence' must be prefaced by a sketch of the German terminological usage, for the picture is quite different from the distinction implicit in the separate terms in English. In fact, all major Western European languages distinguish terminologically between 'clause' and 'sentence', except German, which has one basic word for both (Satz). ('Complex', which would be subsumed under 'sentence', is, of course, included here, too.) Apart from the work of Glinz (1973, 1985) and others adopting his conventions—in which the French term 'Proposition' is borrowed for 'clause'—there is no simple correspondence between the English and German terms. Sommerfeld, Starke & Nerus (1981: 154) statement about varying usage illustrates the situation:


(A clause/sentence can be independent, but also part of a coordinate or complex sentence (main clause, dependent clause, constituent clause, attribute clause). Many linguists ascribe clause/sentence character only to simple independent clauses/sentences, and view clauses in complex sentences not as clauses/sentences.
Others view the complex sentence not as one clause/sentence, but as a combination of clauses/sentences.) When the compound expressions 'Satzverbindung' and 'Satzgefüge' are used, then reference is clearly to sentences, specifically, to those involving coordination and subordination, respectively. Likewise, the expression 'Teilsatz' always refers to clauses, specifically, to dependent clauses which are constituents of larger constructions. However, when 'Satz' is used alone, then it is often difficult to be certain of the intended referent. In the case of 'simple' (i.e. monoclausal) sentences, where the need to maintain a distinction between 'clause' and 'sentence' is driven more by typological principle than grammatical form, little confusion arises. But in many other cases, use of the term 'Satz' may be ambiguous or genuinely vague. For instance, 'einfacher Satz' can mean either 'clause' (=Teilsatz') or 'sentence' (=Hauptsatz'). And when Herringer (1989: 251) states in his pedagogically oriented grammar that 'Sätze sind die eigentlichen Einheiten menschlicher Rede' ('sentences/clauses are the actual units of human speech'), it is not immediately clear which unit he means. Most of the defining criteria he lists could apply to both equally well.

Having no basic sentence-clause distinction, German descriptive linguistics has developed its own terminology for clause and sentence-level structures. However, as there is no uniformly agreed-upon typology of clauses, the terminology used varies from author to author. Differences in structural analysis as well as in subcategorizing criteria are reflected in terminological choice (cf. Helbig 1982: 202). Thus, not all terms are used by all authors, and some terms may encompass markedly different structures,
depending in part on whether form or syntactic/semantic function is emphasized (cf. Bußmann 1983: 444). For instance, van der Elst (1990) classifies dependent clauses into three main types, based on syntactic function; Hentschel & Weydt (1990) list six. Even the 'fundamental' dichotomy 'Hauptsatz'/Nebensatz' ('main clause'/dependent clause') is not immune to critical examination. Engel (1982: 165) writes: 'Die alte Unterscheidung zwischen Hauptsatz und Nebensatz ist viel problematischer, als gemeinhin angenommen wird' ('the old distinction between main clause and dependent clause is much more problematic than is generally acknowledged'). In generative literature, where a clause-sentence distinction is not made, none of these questions ever even arises.14

The disparate usage and varying analytical criteria make adopting any of the customary German terminology problematic. In this exposition, I will continue to use the English terms 'clause' and 'sentence', as well as the term 'complex', when necessary. Where reference to subtypes of clauses is necessary, I will base the classification on specific morphosyntactic properties (e.g. presence or absence of a finite verb, position of the verb), rather than on functional criteria. I will return to this topic in §3.2 when I discuss clauses in more detail.

The higher level which usually constitutes at least the implicit focus of syntactic analyses is, of course, 'sentence'. Both structural and generative

---

14 In older generative models, there is only one category 'S'. In more recent versions, more categories are posited, but the focus is on a main vs. dependent clause distinction, not sentence vs. clause. See von Stechow & Sternefeld (1988, Chapter 11) for a detailed discussion.
accounts of syntax have customarily embraced the sentence *qua* linguistic
object as the form of primary data. Chomsky's (1957: 13) often quoted
statement that '... a LANGUAGE [is] a set (finite or infinite) of sentences ...'
succinctly expresses the extreme position. However, little time is spent
delving into the reasons of this practice—or its consequences. Generative
grammatical analysis especially plays down the importance of the data. For
example, von Stechow & Sternefeld (1988: 3) state only: 'Mit sprachlichen
Fakten sind hier offenbar wirkliche und mögliche Äußerungen gemeint'
('by linguistic facts we obviously mean actual and possible utterances').
What 'Äußerungen' ('utterances') are is not specified. To a neophyte, it be-
comes apparent only through their examples that they are interested in
sentence structure.

Where motivated as a category, the sentence is said to be, on the one
hand, the largest unit which is grammatically governed (syntactic rules op-
erate within this domain), and on the other hand, the smallest unit which
is structurally independent. 'Independence' in this sense can be said to
manifest itself in several dimensions, listed in (8), which converge—at
least ideally—in the sentence.

(8) Dimensions of grammatical 'independence'
   a) syntactic integrity       c) prosodic coherence
   b) semantic unity           d) pragmatic completeness

The constituents which form a sentence thus exhibit tight internal struc-
tural bonds, constitute a meaningful whole, are grouped prosodically, and
are used to perform a certain unitary communicative effect. In other
words, the sentence constitutes the arena in which various aspects of lin-
guistic structure are claimed to be realized.

In their discussion of the sentence as the '... zentrale Einheit in der Struktur der Äußerungen' ('central unit in the structure of utterances'), Heidolph et al. (1984: 154) make this expectation of coinciding dimensions explicit when they stress that: '... die Grenzen der Einheit “Satz” [fallen] stets mit relevanten Grenzen innerhalb der anderen Strukturschichten zusammen' ('the boundaries of the unit 'sentence' always coincide with boundaries of the other levels'). They go on to list the areas where this relationship holds:

In einer Äußerung ... entspricht die Bedeutung eines jeden Satzes einem relativ in sich abgeschlossenen Teilkomplex der gesamten semantischen Struktur der Äußerung. ... Innerhalb einer Folge von Tongruppen fallen die Grenzen von aufeinanderfolgenden Sätzen stets mit Grenzen zwischen Tongruppen zusammen .... Der Satz ist auch die Einheit, in derer Rahmen sich kommunikativ-pragmatische Bedingungen auswirken. (154-5)

(In an utterance ... the meaning of each sentence corresponds to a relatively independent part of the entire semantic structure of the utterance. ... Within a series of tone units, the boundaries of consecutive sentences always coincide with boundaries between tone units .... The sentence is also the unit within which interactional and pragmatic conditions have effect.)

Despite what seem to be bedrock statements, a brief consideration of what these claims entail suggests that such correspondences circumscribe at best a sentence prototype, rather than forming part of a definition or even a de-
scriptive norm. The difficulties associated with determining what constitutes a semantically—let alone a pragmatically—-independent sequence in an utterance are well known (cf. Lyons 1977: 588ff).\textsuperscript{15} The classical idea that a sentence expresses a 'complete thought' is routinely criticized as untenable. Furthermore, it is widely acknowledged that intonation and syntax in connected speech do not necessarily keep step together (Brazil 1982: 289; cf. also Bolinger 1972 on 'sentence intonation'), despite the 'metalocutionary' function of intonation (Gibbon 1984). In recent years, emphasis has thus been placed on specifying syntactic criteria for sentencehood. Hentschel & Weydt (1990: 304) summarize this tendency as follows: '... das Kriterium der syntaktischen Selbstständigkeit [ist] am angemessensten' ('the criterion of syntactic independence is the most appropriate').

The question of syntactic independence is tied closely to two structural aspects of linguistic units: constituency and subordination. In order for a given sequence to be considered a sentence, it must have the proper internal structure, and it must be hierarchically autonomous from adjacent, same-level sequences. The necessary constituents of a sentence are commonly taken to be subject and predicate, or NP and VP. A bipartite conception was expressed by ancient grammarians (Robins 1979: 26), was part of early functional treatises (e.g. Paul 1880) and American structural linguistics (Bloomfield 1933), and surfaces in Chomsky (1957) and subsequent generative work as well. In German scholarship, it is embodied both in tradi-

\textsuperscript{15} Such questions may not even be meaningful. Apart from the problem of analyzing sentences without reference to context (Brown & Yule 1983: 25), we can only hypothesize what an instance of semantic independence would be.
tion-oriented grammatical description, as the quotation in (9) illustrates, and in the phrase structure rules of generative grammars, whether of a simple *Syntactic Structures* type (10) or elaborated in more recent models (11):

(9) 'Die Grundbeziehung, die einen Satz konstituiert, ist die wechselseitige Beziehung von Subjekt und Praedikat' (Brinkmann 1962: 457).

(The basic relationship which constitutes a sentence is the mutual connection between subject and predicate.)

(10) $S \Rightarrow SbG + PG$, where $SbG$ ('nominal group') = $NP$ and $PG$ ('verbal group') = $VP$ (Heidolph et al. 1984: 122).

(11) $[CP \ SpecC \ [C' \ C \ [IP \ NP \ [I' \ INF \ VP]]]]$, where $IP = S$ (von Stechow & Sternefeld 1988: 350).

What constitutes a proper internal structure of a sentence is thus in effect transferred to the equally vexing questions as to the nature of its basic constituents—the subject and predicate. The high degree of controversy surrounding these concepts can be measured by the attention they continue to draw (see, for example, Reis (1982) on 'subject' in German, and Webelhuth (1985) on the 'configurationality' of the German predicate).

Given that the dichotomy subject-predicate is relevant for the description of German sentences, a conservative view would be that in order to have a full-fledged sentence, the subject NP must be overt, and the VP must contain a finite verb form. Hierarchical requirements, however, dictate that presence of an overt subject and finite verb form is not sufficient for sentence status. The sequence in question cannot be part of a larger structure of sentence forms which have been linked together ('conjoined'),
forming a single compound structure with multiple basic constituents. The sequence in question also must not be syntactically subordinate to ('embedded in') any other sequence. That is, it cannot itself be construed as a sentential constituent (e.g. 'subject clause', 'object clause'), part of a constituent ('relative clause'), or predicate modifier ('adverbial clause'). Sentence (12), for example, contains two problematic sequences which cannot qualify as sentences.


\[
\text{Derwal hat [zu beweisen] versucht, [dass der Ball rund ist].}
\]

Derwal has tried to prove that the ball was round.

The bracketing of (12) indicates two embedded clauses, \textit{zu beweisen} and \textit{dass der Ball rund ist}. The first is an instance of clause combining that involves reduction of the embedded clause. Semantically, the subject of this clause is \textit{Derwal}, which is not overtly expressed within the clause. In the second clause, the subject \textit{der Ball} is overt, but the clause functions as the grammatical object of the first embedded clause, and so is not independent.\textsuperscript{16}

\textsuperscript{16} In a standard current generative representation (this treatment of German originally from Thiersch 1978), the situation is the same. The bracketing would be as follows:

\[
\text{[Derwal [hat [s_1, t_1 [[s_2, PRO [t_k, zu beweisen]] versucht t_j [dass [s_3, der Ball [rund ist]]]]]]]
\]

with three S nodes in the sentence, and thus three subject NPs and three VPs as well. \(s_2\) does not qualify for sentencenesshood since its subject position (\textit{PRO}) is not lexically filled and its verb is not in finite form. \(s_3\) does not qualify either: while both overt subject NP and finite VP are present, the sequence is structurally embedded as the object clause of \(s_2\), as it originates at trace \(t_k\). In contrast, \(s_1\) is hierarchically autonomous, and it has both overt subject (\textit{Derwal}, which originates at \(t_j\)) and finite verb (\textit{hat}, at \(t_j\)).
Of course, much more complex sentences can be constructed. In fact, the recursive property of sentences (as well as other linguistic categories) illustrated above is assumed to be one of the hallmarks of natural language (Chomsky 1957: 21f; cf. Newmeyer 1980: 36). This property, however, also makes 'sentence' suspect as a useful basic unit of syntactic description in a process-oriented investigation of spoken discourse. The arbitrary length (from the point of view of the syntax) that any possible sentence may have creates serious descriptive inadequacies directly related to the processing and expression of spoken language. Viewed as a pre-existing object (as is the case in grammars and linguistic argumentation alike) there is rarely any question about the object per se. In standard grammars, every sentence listed is a sentence of the language. In standard linguistic argumentation, the syntactic sequences presented are either sentences or nonsentences (sometimes to greater or lesser degrees)—both are treated as linguistic facts—yet the existence of the object itself is not at stake. But viewed as a product dependent on a number of factors in a particular context, the sequences exhibiting sufficient syntactic integrity to be called 'sentences' are not given before the analysis; they result from it. This issue will be a central focus in §3.3 when I outline my representation of clause construction

---

17 Crystal (1969: 275), for example, reports that sentences in his corpus varied from 1 to 30 tone units.

18 Some may see a sentence/utterance distinction as relevant here. Yet it is evident that there is no consistent application, and methodological contradictions are the rule, rather than the exception. While distinguishing a set of abstract mental entities from particular instantiations of these entities, is helpful in some cases, this line is conveniently blurred for the written language. Hudson (1984: 23) remarks that '[i]t is interesting to notice that linguists have never seen a need to develop a name for written sentence-tokens'.
in discourse.

Concrete examples of the failure of standard conceptions of the sentence to spoken language can readily be found in the literature. The texts transcribed and analyzed in the series *Texte gesprochener deutscher Standardsprache*,\(^{19}\) for instance, make explicit use of the customary sentential constituents:

Die regelmäßige Wiederkehr von Subjekt und Prädikat (Nominal- und Verbalphrase) ist auch im gesprochenen Deutsch in jedem Redeablauf zu beobachten. Sie kann als ein sicheres formales Kriterium für ein Segmentierungsverfahren an Texten eingesetzt werden. (Bausch 1971: 44)

(The regular occurrence of subject and predicate (NP and VP) is also observable in every utterance of spoken German. This fact can serve as a formal criterion for text segmentation.)

While it is true that constituents such as subject and predicate do occur regularly in much speech, the claim about the usefulness of such constituents as a basis for adequately segmenting a text is dubious. In (13) below, an example of where this application of ‘sentence’ runs into difficulties is presented. (Note that the original transcription has been modified to facilitate discussion of the relevant points; see Bausch 1971 on the original conventions).

\(^{19}\) Institut für deutsche Sprache (1971); Os (1974); Fuchs & Schank (1975); and Jäger (1979).
(13) Conversational story excerpt [Institut für deutsche Sprache 1971: 78].

a  [Y]: und ich mein' «ich hab' ja gesagt» #
                         and I mean I have yes said

b  * ich hatte wirklich nich' viel Zeit und war auch 'n bißchen
     I had really not much time and was also a bit
     wütend auf diese alte Tante hat so verglaste blaue Augen gehabt
     angry at this old lady has so glazed over blue eyes had
     so 'n bißchen im Jenseits schon (ja?) #
     so a bit in the hereafter already yes

c  * und dann fing s'e ganz plötzlich an (ne?) und fragte mich
   and then began she completely suddenly on ok and asked me
   «und wie haben wie sind sie eigentlich interessiert?» #
   and how have how are you actually interested

  and how have how are you actually interested

d  * da hab' ich s'e angeguckt #
   there have I her looked at

e  un ich sag' «(ja) (schönen Dank) aber ich hab' wirklich kein'
    and I say yes pretty thanks but I have really no
    Interesse» #
    interest

a  [Y]: and I mean I said 'yes'

b  I really didn’t have much time and was also a little
   angry at this old lady had like glazed over eyes like
   a little in the next world

c  and then she suddenly got going okay and asked
   me ‘and how have how are you interested
      actually?’

d  then I looked at her

e  and I said 'yes thanks but I'm really not interested'

The above transcription includes a sentence-based syntactic analysis
which identifies three major features: sentence boundaries, dependent se-
quences, and nonnormative structures. Here, sentences ('Hauptsätze'/
'Satzgefüge') are consecutively numbered (a-e), and final boundaries are
marked with a number sign '#'. Dependent (complement) sequences are
indicated using double-angled brackets «»; ‘parenthetical’ expressions are enclosed in parentheses ‘( )’; and sentences which contain nonnormative structures are prefixed with asterisk ‘*’.

The syntactic analysis of the spoken language passage given in (13) relies heavily on the norms of standard written German. Sentences (a) and (e) are well formed according to the established criteria (i.e. sequence is governed by a finite verb and not dependent, grammatical is subject present). Sentences (c) and (d) contains sequences which cannot be reconciled with Standard German grammar and are thus marked with an asterisk. Specifically, the third ‘sentence’ (c) contains a syntactic reformulation (the ‘false start’ wie haben corrected by wie sind) in the dependent sequence. The second ‘sentence’ (b) appears to contain a shared argument, diese alte Tante, which is simultaneously the object of the preposition auf and the subject of the predicate hat so verglaste blaue Augen gehabt. (Other interpretations of this sequence are possible; familiarity with the recording would be necessary to evaluate alternatives.) In addition, the final (nonverbal) sequence of this ‘sentence’ so n bisschen im Jenseits schon is syntactically neither part of the preceding predicate nor itself independent, as it lacks both an overt subject and finite verb. It is not immediately clear why the fourth ‘sentence’ (d) has been marked as nonnormative. The sentence contains a sequence which is possibly considered not to be standard in the written language (though it is quite common in spontaneous speech). Alternately, the sentence is so marked because of the lexical item angeguckt, which is colloquial for angesehen ‘looked at’.

207
While there is something to be said for comparing and contrasting the syntactic structure of connected speech with a normative grammar, attempting to force structures which are uncommon or absent in the written language into a rigid, standardized mold does not do justice to the regularly occurring patterns of spoken language. In applying the rules of Standard German grammar to the passage in (13), sequences which are semantically (and probably prosodically) distinct but contain no overt subject or finite verb are interpreted as a type of 'VP-conjunction'. Sequences which contain no predicate must necessarily be lumped together with some other sequence, creating the appearance of structure which, according to the conventions, must then be marked as 'nonnormative' ('normwidrig').

In light of these and other difficulties, many researchers working on spoken discourse have questioned whether it makes sense to look for complete sentences in spoken language. Svartvik (1982: 131) writes that ‘...the traditional unit “sentence” is often impossible or, at best, awkward to work with’ when analyzing impromptu speech. Crystal (1980: 159), after identifying three problem areas that using a concept ‘sentence’ poses, states: ‘It is arguable that all of [these] problems arise solely because of the attempt to impose a descriptive model on the data which uses SENTENCE as a primitive term’. He goes on to say that informal, domestic conversation is not analyzable in a nonarbitrary fashion using the unit ‘sentence’: ‘Rather, the CLAUSE is the unit in terms of which the material is most conveniently organized’. Svartvik, Eeg-Olofsson, Forshedén, Oreström & Thavenius (1982: 70) sum up this view succinctly: ‘... the analysis of spoken language in terms of sentences is difficult; some would even say irrelevant.’

208
Yet other researchers have found use for a redefined unit 'sentence' in analyzing spoken discourse. The question thus becomes whether a unit 'sentence' can be defined and operationalized which takes the norms of spoken, as well as written, language into account. Cumming (1984), for example, answers this question in the affirmative. Using three syntactic criteria to successfully identify 'cohesion clusters' in Chinese and English, she is able to identify a discourse unit 'sentence' consistently in her spoken data. What is especially relevant for the current discussion, however, is the fact that Cumming achieves her goal through the analysis of a smaller syntactic unit: the clause. For Cumming, segmenting a text into a series of sentences does not constitute a basic-level syntactic analysis.

Another approach to defining 'sentence' is strictly nonsyntactic. According to Glinz (1973: 58), 'sentence' is:

... NICHT als eine Einheit [zu fassen], die durch irgendwelche Strukturen der Morho-Schicht bedingt ist ... sondern als eine Einheit, die sich unmittelbar auf die PRÄSENTATION UND SEGMENTIERUNG DER INFORMATION ÜBERHAUPT bezieht, indem sie das Kontinuum des Textes in einzelne kürzere oder längere Schritte gliedert.

(Not to be viewed as a unit which is dependent on structures at a morphological level, but rather as a unit which is actually connected to the PRESENTATION AND SEGMENTATION OF INFORMATION by dividing the textual continuum into individual pieces, shorter or longer in length.)

In written language, such units are, of course, signaled orthographically; in spoken language, intonation—specifically 'sentence-final fall' in pitch—
indicates boundaries of comparable structure (cf. Glinz 1985: 113). Chafe (1980: 20) makes the same observation, noting that a ‘... sentence-final intonation contour often, though certainly not always, coincides with syntactic closure. That is, the sequence of information units which ends with this intonation often constitutes what grammarians would identify as a complete sentence’. (See also Chafe 1993: 132-7, where both intonational and syntactic criteria are used.) Appealing to intonation, however, brings another set of difficulties to the fore. One is now called on to define exactly which set of intonation contours are ‘sentence-final’. But how is this to be determined? Schuetze-Coburn, Shapley & Weber (1991: 228) report difficulty establishing the relationship in conversational discourse that one would predict, i.e. between IUs (segmented auditorily) with a ‘final fall’ in pitch and the end of acoustically determined ‘declination units’ (DUs):

... IUs exhibiting final fall ... do strongly tend to occur at the end of DUs (114/131 or 87%)—however, not all do. Some final falls are DU internal. More importantly, though, only a quarter of all DUs end in a final fall (114/455).

There appears to be no simple prosodic criterion—just as a simple syntactic criterion is lacking—to unequivocally define sentences in discourse.

All of the above-mentioned studies support the view that ‘sentence’ is not primarily syntactic, but textual, in the sense conveyed by the Glinz quotation above. As Chafe (1987: 46) states, sentences ‘... seem to belong to the category of phenomena which are under more rhetorical control, and are more independent of cognitive constraints’. Lyons (1977: 623), too, has observed that the sentence may be better viewed as a rhetorical or aesthetic
construct, rather than as a syntactic unit, i.e. as a sequence whose existence is governed more by writer/speaker purpose and aspects of discourse ‘rhythm’ than structural necessity.\footnote{Chafe (1980: 23-5) gives a particularly telling example of a narrative 40 IUs long that was produced as two intonational sentences.} Plainly, this type of unit is of interest mainly when such issues are in focus.

The negative verdict concerning the use of the term ‘sentence’ for syntactic purposes in spoken language analysis leaves us to consider the suitability of the term ‘clause’. The first thing to be noticed in this regard is that ‘clause’, in contrast to ‘sentence’, can apparently be considered primarily a (morpho)syntactic unit, as will be discussed in the following section. Properties such as prosodic finality, semantic unity, pragmatic completeness, or textual coherence which are mentioned in the context of defining ‘sentence’ are not given the significant weight in definitions of ‘clause’ that they are for ‘sentence’. It is a reliance on syntactic criteria which enables discourse researchers to hold the view that ‘[c]lauses and phrases are easier to identify in speech’ (Svartvik 1982: 131). This is not to say that these other properties play no role in interpreting clause boundaries; as we will see below, reference must be made to the nonsyntactic features in segmenting a text, but these features are always auxiliary to the syntactic criteria.

Secondly, there is a significant difference in scope between ‘sentence’ and ‘clause’. As Longacre (1976: 274) explains, sentence ‘... is preeminently the level of clause combination’. While a distinction may often appear unnecessary—the two units coincide in the case of simple sentences, which
contain a single subject and predicate—clauses are in general syntactically more constrained. By virtue of its limited scope, the clause would seem to be better suited than the sentence for providing the syntactic framework that is necessary for describing the basic structures in spoken language. Other researchers concur. In contrasting the applicability of sentence and clause to text analysis, Glinz (1985: 115) stresses the importance of paying close attention to the clause structure: 'Jeder Text, gleichgültig ob mündlich (gesprochen/gehört) oder schriftlich (geschrieben/gelesen) ist aufgebaut aus PROPOSITIONEN' ('every text whether oral (spoken/heard) or written (written/read) is constructed from CLAUSES'). Crystal (1980: 159-60) also emphasizes the compatibility of clause and the prosody of the discourse.

Finally, 'clause' is the syntactic unit that is most relevant to hypotheses regarding cognitive activity and language. Halliday (1967: 200f), for example, postulates that in the unmarked case the clause is realized as a single 'information unit'. Givón (1975: 204) observes a discourse strategy of 'one bit of new information per proposition', clauses being, in effect, the linguistic equivalent of propositions. Pawley & Syder's (1977) 'one-clause-at-a-time' hypothesis and Du Bois' (1987) 'Preferred Argument Structure' reflect the role of clause as an information processing unit. The 'idea units' of Chafe (1980: 13f) are defined by 'clause-final' intonation, and Bock (1982: 23), after reviewing the psycholinguistic literature, comes to the conclusion that '[c]urrent evidence ... suggests that the planning unit is roughly clausal in scope'.

In the preceding discussion, we have seen the need to keep the concepts
of clause and sentence distinct in dealing with spoken language, for theoretical as well as heuristic purposes. At the same time, the strengths and weaknesses of the units ‘word’, ‘group’, ‘clause’, and sentence’ were considered, and the conclusion was drawn that ‘clause’—and not ‘sentence’, as is often the case—should form the basis of a syntactic analysis of natural conversational speech.

3.2 Defining clause and issues in clause identification.

If ‘sentence’ has received so much attention in linguistic theory, especially for languages with a well-established written tradition, ‘clause’ per se has received comparatively less, at least in terms of attempting to provide an explicit definition. This is probably due in part to the relatively recent distinction of sentence and clause (Matthews 1981: 190), but surely also to its secondary status as a unit between what was perceived as the maximal grammatical domain (sentence) and the minimal (word). Perhaps, too, there has been less concern with ‘clause’ because a general agreement about its characterization is assumed, along the lines of Longacre’s (1976: 273) statement that ‘[c]lauses are, broadly speaking, predicational strings ...’. Under this view, the task of defining ‘clause’ is reduced to relating a semantic notion of ‘predication’ to some grammatical category or operation, as can be done with the concept of ‘valency’.

---

21 It is nearly obligatory in this connection to cite Ries (1894; 1931) and Seidel (1935), with their compilation of nearly 225 different definitions of ‘Satz’. Cf. also the various definitions discussed more recently in Müller (1985).

22 Cf. Glinz (1985). Lyons (1968: 171) points out that a clear distinction between clause and phrase was lacking in traditional grammar.
Because of the usefulness of the concept of 'valency' in defining 'clause', I will discuss it at this juncture. Grammatical valency refers to the ability of a form class to bear a specific number and type of arguments (Helbig 1982b: 12; cf. also Engel & Schumacher 1978: 15-7). For instance, the verb frühstücken 'to breakfast' is said to have a valence of 2 and take an obligatory nominative NP and an (optional) accusative NP argument, as in Er frühstückt ein Ei 'He's eating an egg for breakfast' (Helbig & Schenkel 1982: 125); or the adjective befindlich 'situated' is said to have a valence of 2 and relate an NP and an adverbial of place, as in das sich dort befindliche Lager 'the storehouse situated there' (Sommerfeldt & Schreiber 1983: 75). Valency has a direct connection to predicate calculus (Lyons 1977: 486), and so neatly provides the sought-after bridge between semantics and syntax. While valency is more than a property of verbs, its relevance for a definition of clause in German lies in an application to this form class. In this vein, Brinker (1985: 22) spells out his definition as follows:23

Auf der Grundlage des Valenzmodells kann SATZ als eine sprachliche Einheit definiert werden, die sich aus einem Verb (Prädikat) als dem strukturellen Zentrum und einer Reihe von Satzgliedpositionen (Subjekt, Objekte, Adverbialebestimmungen, usw.) konstituiert, die jeweils in bestimmten Abhängigkeitsrelationen zum 'tragenden' Verb stehen.

(On the basis of a valency model, CLAUSE can be defined as a lin-

---

23 It is clear in this case that Brinker is referring to clauses: 'Sätze können danach als einfache Sätze oder als Teilsätze (Haupt- und Gliedsätze in sog. Satzgefügen) realisiert sein' (22) ('Clauses can thus be realized as simple sentences or as main or constituent clauses in so-called complex sentences').
guistic unit that consists of a verb (predicate) as the structural center and a series of constituent positions (subject, object, adverbials, etc.), which in each case stand in certain dependency relationships with the verb that bears them.)

The extent of a clause is thus limited primarily by the ability of the (main) verb to bear nominal and prepositional arguments (which in German are typically marked by morphological case) and other, largely adverbial, elements. Although the model has a semantic origin, it is applied in terms of lexicosyntactic properties. The ideas of prosodic finality, semantic unity, and pragmatic completeness discussed above do not, in theory, play any role, as they do in definitions of sentence. Determining the boundaries of clauses, then, depends only on ascertaining the grammatical relationship between a lexical verb and adjacent phrases. Naturally, there are complications in application, as well as different possible interpretations of detail (cf. Matthews 1981: 168-89). I will discuss specifics in examples below, and in §3.3 when I exemplify my model of clause representation.

Along with valency, a concept of 'topological field' has been developed in German linguistics to help characterize and subclassify 'Satz'. Since the outlines of this concept are typically described in relation to simple sentences (which are coextensive with clauses), certain aspects can be used quite successfully in clause identification. The concept 'topological field' refers to constituent positions in the clause with respect to its verbal elements. Drach (1937) called these positions 'Felder' ('fields'), distinguishing a

24 Glinz (1985: 114) does not refer to valency, but states simply that a clause is '... was auf einer einzigen Verbalstruktur beruht' ('what is based on a single verbal structure').

215
'Vorfeld' ('prefield') before the finite verb form and a 'Nachfeld' ('postfield') after it, with the explanation: 'Das Verbum finitum ist der standfeste Angelpunkt, um den herum der Satz sich aufbaut und gliedert' ('The finite verb is the most stable anchor point around which the sentence/clause is built and segmented') (16). Since Erben (1954) it has been usual to distinguish three fields in a German clause. Drach's Nachfeld is split into a 'Mittelfeld' ('midfield'), which is the position between the finite verb form and the rest of the verbal complex, and a 'Nachfeld', which follows the last element of the verbal complex. According to this model, then, there are three argument positions and two (verbal) predicate positions in the German clause, as schematized in (14).^{25}

(14) Prefield Verbfield\textsubscript{a} Midfield Verbfield\textsubscript{b} Postfield

When both verbfields are realized, the two predicate parts are said to form a 'verbal brace' ('Satzklammer' or 'Satzrahmen') which in effect brackets the preficld off from the rest of the clausal constituents.^{26} Given this framework, clauses may be subcategorized as to the position of the finite verb, which, as the above quotation indicates, is considered to be its most stable constituent. Three subclasses are generally recognized: (a) 'Kernsätze', or clauses in which the finite verb occupies Verbfield\textsubscript{a} (often called 'V-2', as the finite verb is the second constituent); (b) 'Spannsätze', or clauses in which the finite verb occupies Verbfield\textsubscript{b} ('V-L', for 'verb-last');

---

^{25} See Askedal (1986) and especially Höhle (1986) for various conceptions of this model.

^{26} Despite the controversies surrounding the elevation of this pattern to the status of a 'grammatical rule', it beginnings '...reichen mindestens in germanische Zeit zurück' ('reach at least back to the Germanic period (Ebert 1978: 39).
and (c) 'Stirnsätze', or clauses in which the finite verb occupies Verbfield$_a$, but the prefief is unrealized. (Occasionally it will be useful to group subclasses (a) & (c), as opposed to (b). In these instances I will describe the clause as 'verb-early' (V-E).)

There are several important details of this model for clause identification. First, the prefief normally contains at most one constituent. Thus, by locating the finite verb in a V-2 clausal sequence, the beginning of the clause can be ascertained when the immediately preceding constituent is identified. Second, the postfield is often unrealized. Thus, the end of the clause can be ascertained by identifying the final element in Verbfield$_b$. In V-2 clauses, this position consists of all elements of the 'verbal complex' except for the finite verb, which is in Verbfield$_a$. The exact composition of the verbal complex is controversial (see Van der Elst 1990: 83 for a brief statement of various positions), but generally included are both grammatical and lexical adjuncts ('Zusätze') which together form the verbal predicate. In the main, I will follow those authors who have a broad conception (e.g. Götze 1979; Engelen 1984), by treating any of the elements in (15) as part of the 'verbal complex', to the extent that they are present.

(15) a) auxiliary and modal verbs (e.g. sein 'be', müssen 'must')
b) modality verbs (e.g. brauchen zu 'need to')
c) verbal prefixes (e.g. aufmachen 'open')
d) adjectival adjuncts (e.g. fernsehen 'watch TV')
e) adverbial adjuncts (e.g. hinhalten 'delay')
f) bare nominal adjuncts (e.g. Gefahr laufen 'run the risk')
g) prepositional adjuncts (e.g. zur Entscheidung bringen 'reach a decision')
h) reflexive pronoun with reflexive verbs (e.g. sich lohnen 'be worth')
In V-L clauses, where Verbfield_a is not realized, the finite verb is in most cases the final element of Verbfield_b. For V-2 clauses with no Verbfield_b, other criteria for ascertainning the end boundary must be used. In these cases, boundaries are determined more through expected valency than the position of constituents.

In sum, the concepts of valency and topological field provide a rich framework for the identification of clauses in discourse. Their application is frequently complementary: field structure delineates clausal sequences through the anchorpoints of the split predicate, and valency establishes the specific relationships between the verbal complex and the interlaced arguments. In other words, an utterance can be segmented into clauses by ascertainning the argument structure for each lexical verb. It now remains to show how I will represent clauses in the corpus. But before turning to these details, it will be worthwhile to examine first some specific cases of clause identification and the issues that arise therefrom. I begin with a brief consideration of written German.

As with the previous concern over basic syntactic units, written and spoken language differ somewhat in how clauses may be identified. In written German, it is a fairly straightforward process to delineate clauses, regardless of the actual constituent structure, because of prescribed use of capitalization and punctuation. Of course, for simple sentences—which contain only one verbal complex and so consist of a single clause—there is no question. The first word is capitalized and a sentence-final punctuation mark (period, question mark, or exclamation mark) follows the last word.
When the sentence is composed of several clauses, as in (15), standard orthography (which includes the use of punctuation) often provides direct delineation of sentence-internal clause structure. In this instance, there are four clauses, as the four lexical verbs or verb complexes (boldface) indicate. Clauses (a) & (c) are V-L, with their arguments initially bounded in the case of (a) by the sentence beginning, and in the case of (c) by the verb complex of the previous clause. Occurring between (a) & (c) is a nonfinite clause (b). Clause (d) is V-E, with the end of the clause coinciding with the end of the sentence. In addition to the grammatical criteria, each clause is separated by commas.

(15) Written literary German [Feuchtwanger 1940: 9].

\begin{itemize}
\item[(a)] Als er aber vorsichtig Papier und Bleistift aus der Schublade kramte, 
\hspace{1cm} as he but carefully paper and pencil out of the drawer rummaged \\
\item[(b)] um sich das Motiv aufzuschreiben, 
\hspace{1cm} in order himself the motif to write out \\
\item[(c)] das ihm eingefallen war, 
\hspace{1cm} which to him occurred was \\
\item[(d)] fing er ein Buch von dem gebrechlichen, überladenen Schreibtisch. 
\hspace{1cm} brushed he a book from the fragile overloaded desk \\
\end{itemize}

\begin{itemize}
\item[(a)] As he carefully took paper and pencil out of the drawer \\
\item[(b)] to write down the idea \\
\item[(c)] that had occurred to him, \\
\item[(d)] he knocked a book off the fragile, overloaded desk. \\
\end{itemize}

Most ordinary written language as well can be segmented into a sequence of sentences and clauses in this way without too much difficulty. Of course, not all written language is like this: exceptionally idiosyncratic or creative language will prove intractable to such simplistic segmentation.
Schlobinski (1992: 117) gives an extreme example from Arno Schmidt's *Zettels Traum*, which he describes as 'ein Chaos von Sätzen' ('a clause/sentence chaos'). Moreover, standard orthography does not function only to indicate clause and sentence-level structure. As the above example also illustrates, punctuation marks like comma serves other purposes as well. The comma within clause (d) separates coordinate adjectives modifying *Schreibtisch*, not clauses. The converse, too, can be found. Not all clauses are delimited by punctuation marks, e.g. short infinitive clauses are not always separated by a comma. Nevertheless, the identification of clauses can be achieved for the most part relatively successfully, given some additional systematization, such as a coherent typology of clause types to properly deal with questions which arise due to clause combining. Glinz (1985: 118) reports, for example: 'In 80 bis 90% aller Fälle wird es über die Abgrenzung der Propositionen keine Zweifel geben' ('in 80-90% of all cases there will be no doubt about the segmentation of clauses').

In spoken language, it is doubtful that clause identification is so clear cut. In spontaneous speech, the lack of opportunity for careful editing is the source of much nonnormative structure vis-à-vis written forms, but in all speech the convenient and usually unambiguous orthographic notation of written language does not exist. While spoken language possesses a prosodic dimension—lacking overt form in written language (though see Chafe 1988)—which in some respects mirrors standard punctuation, as I discussed in the previous chapter, prosody is multifunctional, and indexes a variety of processes in addition to the indication of grammatical hierar-
chy. Thus, the prosodic structure of an utterance cannot be relied on to provide a consistent, unambiguous segmentation into grammatical units like clauses. Even so, prosody does give some indication of syntactic structure, in that, minimally, it segments the speech stream into smaller, manageable pieces, which do often coincide with syntactic units, as was shown in the previous section.

The structure of some spontaneous spoken language does give the appearance that clause identification is as easy as I outlined in the previous example. The excerpt in (16) below, which constitutes the opening IUs of one of the experimentally elicited narratives in the secondary corpus (see §1.4), illustrates.

In the excerpt, the clauses are all easily delineated. From the concept of valency, each instance of a finite verb or verb complex (boldface) can be taken to establish one clause. Thus, there are seven clauses to be marked in (16).

---

27 Of course, it is somewhat backwards to view matters in this way, as punctuation was originally devised to capture the aspects of prosody which are most relevant to the grammatical structure of utterances, such as pausing and pitch contour (cf. Nerius 1987: 178-9). The prosodic structures of connected speech are much richer than any existing orthography represents.

28 Some studies have shown a higher correspondence than I am suggesting. In a series of studies, Lehide has examined the relation of prosodic features to sentence and paragraph structure (for summaries, see Lehide 1982; 1984). Price, Ostendorf, Shattuck-Hufnagel & Fong (1991: 2963) find that "... prosodic boundary cues are associated with almost all reliably identified sentences". Most of these (and other) works, however, are psychoacoustic studies involving read speech.
(16) 'Pear Story' narrative [Corpus B, Pb8: 1-9].

1 Y₁:  { also als èrstes /+
        so as first
2  wird eine ... Ländschaftsszene gezèi=gt /+
    becomes a landscape shown
3  { und ~da is' ein riesiger Bäu=m /+
    and there is a giant tree
4  { an dem Baum steht eine Lèiter /+
    on the tree stands a ladder
5  { (H) «.6» und èben auf der Lèiter steht ein MÀnn /+
    and up on the ladder stands a man
6  { und pflßckt ... uh .. Birnen \• }
    and picks uh pears
7  { ... und unten auf der Èrde stehen /+
    and below on the ground stand
8  ... (H) «.7» drei Kòrbe /+
    three baskets
9  { ... die schon hàlb mit Birnen gefüllt sind \+
    which already half with pears filled are

From the concept of topological fields, the boundaries of the clauses are located by determining the position of the finite verb in the clausal sequence. The first clause, for example, is V-2. IU 1, which precedes the finite verb wird, comprises the first constituent of the clause, and the participle gezeigt ends the clause, as it occurs in the second predicate part position. The second clause is also V-2, but in this case there is no split predicate.

222
While the beginning of clause 2 is already bounded by the extent of clause 1, the end of the clause is not indicated by a field position. Instead, appeal is made to the expected argument structure and, looking ahead, to the boundary indicated by the next lexical verb. In this way, clause boundaries can be established between the constituents *ein riesiger Baum* and *an dem Baum*. These procedures can be applied to the remainder of the excerpt to arrive at the clause bracketing indicated in (16). Under the notation used here, the beginning of each clause is marked with an open bracket ‘[’, the end with its closing counterpart ‘]’. (Thus, the clauses beginning at IU 1 and IU 7 each comprise two IUs, the remaining IUs are coextensive with a clause.) In six out of the seven clauses in this initial part of the narrative—the exception being IU 6—the speaker produced clauses with fully overt arguments; each of these clauses would compare favorably with simple sentences in the written language. Even in this one exceptional case, however, the clausal nature of the sequence is clear from the expected valency associated with the verb *pflücken*.

Much spoken language is similar to the above excerpt in that it apparently conforms syntactically to standard descriptions based on the written language—and thus to the ideal outlined by Heidolph et al. (1984) cited in §3.1.3. This view is held to be true especially for more formal types of speech: lectures, orations, sermons, and so forth. There is much variability in more informal discourse, not only in the face of situational differences, but also because of ‘idiosyncratic’ speaker variation. In comparable situations, two speakers will produce clauses with different degrees of fluency. Speaker Y₁ in (16) was highly fluent; the speaker who produced the narra-
tive excerpt in (17)—the beginning of another Pear film 'Nacherzählung'
from the secondary corpus—a bit less so.


1  \( Y_2: \) \{(H) da is' 'n Männ /+ \}
   there is a man
2             \{( ... der pflückt .. Birn'n /+ \}
   he picks pears
3             \{( ... (N) «1.2» und hat zwei Körbe vor sich zu stêhen /+ \}
   and has two baskets before himself to stand
4             \{( ...(1.2) und ... pflückt die èbend /+ \}
   and picks them just then
5             \{ .. und da kommt -- \}
   and there comes
6             \{ .. èrst kommt 'n Männ vorbei mit 'm Èsel /+ \}
   first comes a man by with a donkey
7             \~uh nee mit 'm Kàlb glaub' ich \+
   uh no with a calf believe I
8             \(\text{(H)} \) ... «.9» ùn=d /+
   and
9             \{ ... dann geh- -- \}
   then go-
10            \{( ... er geht dann wieder hùch den Bàum /+ \}
   he goes then again high the tree
11            \{(N) ... «1.0» und pflückt wêiter /+ \}
   and picks further

---

1  \( Y_2: \) there is a man,
2  he's picking pears,
3  and has two baskets in front of him,
4  and <he's> picking them just then,
5  and there comes--
6  first a man comes by with a donkey,
7  no with a calf I believe;
8  and,
9  then <he> go--
10  he then goes up the tree again,
11  and continues picking,
As before, the finite verbs and verb complexes are indicated in boldface. Save IU 8, which consists solely of the conjunction *und*, each IU in (17) contains a lexical verb; thus, we would expect to find ten clauses. I have bracketed only seven clauses, however. All seven clauses are V-2; in three instances—in IUs 3, 6, & 10—the clause-final bracket is determined by the presence of verb complex elements in the second predicate part position (for the clauses in IUs 6 & 10, the postfield is occupied); in four instances—in IUs 1, 2, 4, & 11—argument structure must be used. There are, then, four finite verbs (and associated arguments)—in IUs 5, 7, 8, & 9—whose clausal membership has not been determined. The ‘nascent clauses’ in IUs 5 and 9 have been left unbracketed for two reasons: they are not complete units, and their subject NPs are indeterminate. The clausal status of the lexical verb in IU 7 is unclear, and its relationship to the prepositional phrase and negative element at the beginning of the IU is open to interpretation. Finally, the connection to adjacent structures of the isolated conjunction *und* in IU 8, which falls between two unbracketed IUs, is also unclear. These unbracketed sequences are difficult to evaluate under the valency–topological field model that I have been assuming.

Of course, in natural discourse—as opposed to elicited speech of the type in (16) and (17), where there may be some pressure to monitor one’s speech particularly carefully—the variation in fluency of clause formation that will be encountered is greater. Example (18) below, which is from the primary corpus, provides a glimpse of this variation.
(18) Corpus excerpt from ‘Getting by’ [Henrika’s visit: 251-66].

251 G:  
(H) «9» na jà \+ 
well yes

252 aber\% -- 
but

253 .. ich mèin' /+ 
I mean

254 { ~so wie das âu=ssieht jetzt /+ } 
like how that out.looks now

255 { .. «P ~also wa=s=s P» /+ 
so what

256 .. m%= \+

257 .. (TSK) die ângeht /+ } 
her concerns

258 { ~also kànn man doch wirklich nu=r \+ 
so can one certainly really only

259 .. ~kùck'n } { daß% -- 
look that

260 .. (m=} «8» die Christel sèlbständig [ werd \+ 
the C independent becomes

261 (H): 
[ mhm= ] \+

262 { ~daß ] sie% /+
that she

263 (H) «1.0» «F lò=skò=mmt F» /+
away.comes

264 uh= «.7» \+

265 auf [ jèden ] Fall von den Typen /+
on every case from the characters

266 (I): 
[ (H) «P <ja> P» ] \+ 
yes

226
In this stretch of sixteen IUs, six finite verbs and verb complexes are found, five of which have been bracketed as clauses. Yet, there is hardly a single clause which is fluently produced. Four of the five bracketed clauses contain vocal markers of hesitation, pauses, or breaks for inhalation, which have all simply been overlooked in the bracketing. Here, too, there are unbracketed sequences that cannot be syntactically placed with any certainty, due to the fragmentary nature of the speech.

Some samples of spoken language are evidently more amenable to standard structural analysis than others. Example (16) represents perhaps a 'best case', i.e. where spoken utterances can be unambiguously matched to clauses (and sentences) in the written language. But requirements of descriptive adequacy call for more than dealing with only the best cases. Examples (17) and—especially—(18) illustrate through the speaker's comparative lack of fluency some of the well-known sources of problems in representing syntactic structure in spoken language. 'False starts', syntactic breaks, 'repairs', repetitions, and hesitations all make the descriptive tasks of the syntactician much more difficult. In the above excerpts, it is not immediately clear how the unbracketed IUs are to be integrated into the syntactic representation. Do these utterances themselves constitute clauses? Are they parts of other clauses? To what degree are they extraclausal? Furthermore, the bracketing of the IUs which have been delineated syntactically may require some revision in light of decisions made about the unbracketed IUs. In the following section I will address these issues in conjunction with the introduction of the concept of 'clause construction' and the outline of a system for its representation in natural discourse.
3.3 Clause construction and the ‘free-bracket’ convention

In the preceding sections, I considered whether it was sensible to make use of the linguistic concepts ‘word’, ‘group’, ‘clause’, ‘complex’, or ‘sentence’ in describing basic syntactic structure in spoken discourse. While an assessment of the options concluded that a unit ‘clause’ is best suited for unscripted conversational speech, especially the type included in the present corpus, it is equally apparent that there is substantial variation in speakers’ formulation of clauses, as examples (17) and (18) have demonstrated. As a consequence, a rigid conception of this syntactic unit would prove less than successful. There are therefore two goals to be aimed for. Along with providing a syntactic representation of spontaneous speech with a minimum of structural detail (yet enough to clarify the relationship between prosody and syntax), a certain flexibility in the representation of clauses must be introduced to cover the range of structures that are encountered.

The approach I will embrace here accepts the valency–topological field framework outlined in the previous section, but views clauses as active constructions by the speaker (and sometimes by the listener as well) with a temporal component supplied by the incorporation of the prosodic phrasing, rather than as a purely static object which must fit some normative schema. The system of clause representation I will propose respects this dynamic feature.29 The principal observation underlying my approach is this: ‘clause’ often seems to be a syntactic unit that speakers put together

29 The discussion in this section is based on Schuetze-Coburn (1993c).
on-line, IU by IU, rather than utter as a prepackaged structure which may be analyzed without regard to the way it was formed. Other discourse analysts have come to the similar conclusions. Brazil (1982: 277), for example, explains his understanding as follows:

... utterances are typically produced in real time and decoded in real time. It quickly proves impossible ... to ignore the way the substantive events that constitute speech are organized in linear sequence. In this respect intonation encourages a quite different theoretical stance from that adopted by those who investigate the apparently-hierarchical structuring of syntax. Although linguistic orthodoxy has given little recognition to the fact, there are many questions of both a theoretical and a practical nature that can scarcely be formulated satisfactorily without reference to some kind of step-by-step model.

Furthermore, as I will show below, a focus on the stepwise construction of clauses—taking into account the basic prosodic structure of the utterance—permits the accommodation of most all portions of the utterance into the syntactic analysis, as opposed to the usual editing or shunting aside the 'ungrammatical' sequences, 'repairs', and peripheral elements. As a result, viewing syntax in discourse as clause construction contributes to the methodological aims of whole-text analysis, as stated in Chapter 1.

The notation that I will be detailing below—which I refer to as the 'free-bracket convention'—is designed specifically to take the incremental production of clauses into account. The focus of the notation is on clause boundaries, rather than on constituency. As I have explained, I share the
view of many researchers that a clause generally revolves around a syntactic schema linked to a lexical verb. Yet, when a temporal component is incorporated into the description, the main issue becomes the establishment of boundaries, not the schema, which may itself change as the clause further constructed (cf. Ono & Thompson 1993). With this in mind, the key to representing clauses is the abandonment of the standard convention of requiring matched pairs of brackets. Under the free-bracket convention, each bracket marks a boundary which is independently established. I will illustrate this principle shortly.

First, there are a number of general properties of clause-level syntax in spoken discourse to be mentioned, which, if properly taken into account in representing clause construction, will go far to satisfy the stated goals. In (19) I list twelve observations which constitute the basis for a discourse representation of clauses.

(19) a) Speakers often construct clauses having the extent of one IU.
    b) Clauses may be constructed shorter than one IU.
    c) Constructions may begin and end with interactional elements.
    d) Speakers add to clauses after syntactic closure has been reached.
    e) Syntactic closure may occur more than once per clause.
    f) Clauses may be constructed longer than one IU.
    g) Speakers may restart clauses.
    h) Subordination is less well defined than in the written language.
    i) The beginning of an IU frequently coincides with a new clause.
    j) Closure is frequently reached at the end of an IU.
    k) Speakers may abandon clauses and construct new ones.
    l) Speakers can collaborate on the construction of a clause.
Most of these observations are, of course, not new. Many similar ones can be found in the Conversational Analysis literature, and in linguistics literature in the context of descriptions of the 'aberrant' nature of spoken language, or under headings such as 'anacoluthon', 'ellipsis', or 'repair'. Only occasionally, however, do prosodic features play a role in the discussions, and rarely—if ever—is prosodic structure systematically taken into account.

In order to discuss the above properties, it will be useful to have a more extended excerpt from one of the corpus stories to draw upon. Example (20) below presents an excerpt which I will use to illustrate observations (19a-j). The remaining observations will be discussed with reference to other examples.

The excerpt in (20) consists of 24 IUs by Diana, the main speaker in the conversation, who has the floor throughout this passage. (The three backchannel utterances are ignored here, as backchannels in general will not be integrated into the representation of clause construction.) In 16 of these 24 IUs, there is a finite verb present, which in each case constitutes the basis for a clause. In a number of instances, the scope of the clause coincides exactly with the extent of the IU. This basic observation (listed as property 19a) is true for IUs 31, 36, 44, and 45.

30 Recall that I have taken a fairly restricted view of what utterances are to be treated as 'backchannel' (§1.5). In the present corpus, most backchannels are tokens of forms of mhm and ja. In principle, backchannels such as those in IUs 26, 27, and 37 could be represented as nonverbal 'clause equivalents' (discussed below); however, I have found no compelling reason to do so, and for simplicity's sake, have excluded them.
(20) Corpus excerpt from 'Visa for India' [Doppelkopf b: 20-46].

20  [D]:  ... nà ja \+ 
        well yes
21        und där=nn\% /+
        and then
22        ..(7)(hX)\p und dann \p wir wollten dònerstags fliegen \nè /+
        and then we wanted on Thursday fly ok
23        (H) «.7» nà ja /+
        well yes
24        und dienstags war 's noch nicht dà \nè /+
        and on Tuesday was it still not there ok
25        ... [ "dòh" dachten ] wi' /+
        oh thought we
26        (Z1):  [ ò jè= ] /+
        oh my
27        "jètz' rufen wi' die mal àn" \[2 nè \]/+
        now call we them once on ok
28        (Z2):  [2 \p mhm \p ] \*
        [2 a mhm \p ] \*
29        (H) «3» - und dann wàres aber schon grade irgendwie mittags /+
        and then was it but already even somehow midday
30        oder sò /+
        or so
31        ~auf jeden Fall war da kàum noch einer dà= /+
        on every case was there hardly still one there
32        (H)...(n) «14» - na un' dà haben die uns die Nùmmer gegeben \nè /+
        well and there have they us the number given ok

20  [D]:  well;
21        and then,
22        and then we wanted to fly on Thursday,
23        well,
24        and on Tuesday it still wasn't there,
25        "oh" we thought;
26        (Z1):  oh no,
27        let's call them,
28        (Z2):  mhm
29        and somehow it was already about noon,
30        or so,
31        anyway there was hardly anyone still there,
32        well they gave us the number,
von= --
from

also von dieser Stelle \\+
so from this place

\-wo w'r das Visum krie=gen \\+
where we the visa get

\-und [ haben wi' ] dann nachmittags noch ma' ângerufen \\+
and have we then afternoon still once called

(Z2):
\[ <p mhm p> \]

(H) «.6» \-un' dann éiner /+
and then one

dér hat sofort âufgelegt /+
he has immediately hung up

... bei éiner Nummer hatten wi' n Anrufbeantworter \ nè /+
by one number had we a answering machine ok

(H) «.7» und frégendwann hatten wi' dann= /+
and sometime had we then

so ... frâu d'ran= \+
so one [FEM] there on

die meint dann "o Gott o Gott" \ nè /+
she means then oh God oh God ok

.. \-"das is' ja nur noch éineinhalb Tä=ge /+
that is yes only still one and a half days

und dann fliegen S'e schon" /+
and then fly you already

und dann [ müßten wi' ] ganz schnell [2 gucken \ nè ] /+
and then had to we completely quick look ok

of --
of this place;
where we get the visa;
and then we called in the afternoon;

(Z2):
mhm
and then one <guy>,
he hung up right away,
we got an answering machine at another number,
and then we got,
this woman;
she said then "oh God oh God,
there's only a day and a half left,
and then you're leaving",
and then we had to check real quick,
Using brackets in much the same way as was done in examples (16), (17), and (18), the exemplification of this one-to-one correspondence can be indicated as follows:

(20a) 44 #{ .. ~”das is' ja nur noch eineinhalb Tà=ge /+ } that is yes only still one and a half days

The clause in (20a) is V-2, with the constituent das occupying the prefield. The beginning of the clause is thus marked with an open bracket before das ‘[’. (In addition, a number sign ‘#’ is used to mark ‘new clause construction’. This second convention seems redundant in this case, but it will prove to be necessary in a moment, so I include it at the outset.) The second predicate part position Verbfielbd is not realized in this case, but the expected valency of the copula puts the end of the clause at the end of the IU, which is marked with a close bracket ‘]’. The extent of the clause in IU 44 is thus indicated by the two brackets, applied independently.

The placement of the final bracket in the above example was fairly obvious, though in general, there arise cases where final bracket placement may be more difficult, especially if the predicate is not split. An important working principle, however, clarifies these cases, namely: close brackets are indicated at ‘actual points of syntactic closure’. The clause in (20b) illustrates the application of this principle:

(20b) 31 #{ ~auf jeden Fall war da kàum noch einer dà= /+ } on every case was there hardly still one there

---

31 For ease of recognition, clause-initial bracketing is arbitrarily placed before all prosodic feature markers.
This clause is also V-2, and the preposition auf in the group auf jeden Fall is to be construed as the first word of the clause. Again, Verbfield is not realized in this clause, and an acceptable valency of war would permit a close bracket after einen. But the fact that another constituent occurs at the end of the same IU bearing the primary accent which can be syntactically accommodated signifies that a point of syntactic closure is realized only at the end of the IU, where the close bracket is placed. In other words, 'syntactic closure' does not refer to potential points of syntactic completion, only to points that are actually realized by the speaker. Under this view, the prosodic coherence of an IU guides the listener PAST potential clause boundaries. This clause is thus also coextensive with an IU.

Now, most of the clauses in this passage do not correspond with the prosodic structure perfectly, as do (20a-b). In some instances, the correspondence is close, but the IU contains some internal bracketing. Clauses with this structure are found in IUs 27, 40, and 46. Example (20c) illustrates this, the second observation (19b), that a clause may be less than the extent of an IU:

(20c) 46 #\{ und dann [ mußten wi' ] gänz schnell [2 gucken \ } nè ] / + } and then had.to we completely quick look ok

The scope of the V-2 clause is grammatically indicated at its beginning by the conjunction und and temporal adverb dann in the prefied.\textsuperscript{32} Near the

\textsuperscript{32}This is a simplification. There is a case to be made for considering conjunctions like und to lie outside the clause, or rather BETWEEN clauses. On the other hand, there are also cogent reasons for these conjunctions to be clausal elements (see e.g. Höhle 1986:332). Authors who favor the latter approach generally assign the conjunction to a separate field. Here, I am
end of the IU, the infinitive *gucken* signals syntactic closure. As in (20a-b),
brackets mark these boundaries and delimit the extent of the clause. This
clause, then, does not coincide exactly with the IU: there is a particle *ne*
which occurs after it in the same IU.

In the example, a second close bracket has been placed after the particle
*ne*. Under the free-bracket convention, more than one close bracket may be
associated with a given open bracket. The bracketing in (20c) takes advan-
tage of this convention to represent the particle as an element which has
been added to the clause CONSTRUCTION after a point of syntactic closure in
the clause had been reached. That is, *ne*—a ‘lexical boundary marker’
(Gliederungs–signal, see §2.3.5) serving interactional purposes—is viewed
as peripheral to the clause, but part of the larger construction which may
extend to encompass either prior or subsequent items. Such an additional
element may occur at the end of an IU, as above, or in an IU by itself, as in
(20d):

(20d) 29  #: (H) «3» ~und dann wàres abers schon grade irgendwie mittags /+
       and then was it but already even somehow midday
       30       oder sò / + }
or so

This is the essence of the concept of ‘clause construction’: a central ‘clause
core’ may be formed with a variety of peripheral elements. These example
partially illustrate observation (19c), that interactional elements may begin
or end a clause construction. Example (20e) illustrates the occurrence of ini-
treating *und* much like a ‘particle’, which does not occupy a field by itself, but is nevertheless part of the clause.

236
tial interactively oriented elements that are bracketed to the periphery (there is a final *ne* as well):

(20e) 23 \#\{ (H) «.7» nà ja /+
well yes
24 \{ und dienstags war 's noch nicht dà \ } nè /+
and on. Tuesday was it still not there ok

Incidentally, it is due to examples such as this where an open bracket is used to separate initial elements from the rest of the clause that the number sign ‘#’ is also necessary at the beginning of the clause construction in order to indicate when a new construction is begun. In other words, brackets mark clausal boundaries, and the number sign marks clause construction boundaries. In regard to the placement of construction boundaries, it should be mentioned that the number sign is placed at the beginning of a new construction, rather than at the end of a completed one, because (as we will see in the next example) syntactic closure is never absolute. It is not until the beginning of a new construction that it is apparent a new construction has actually been started, that is, that the previous one has been ended.

As there is no predication connected with forms such as *ne, oder so,* and *na ja,* an independent clause bracketing is not justifiable. These forms fall syntactically outside the scope of the clause proper: the conjunction + adverb sequence that we saw in (20c-d) also constitutes the syntactic beginning of the clause in (20e) as it occupies the prefield, and there is often very good indication of syntactic closure at the end of a clause (e.g. a participle or infinitive in Verbfieldb) before a final peripheral element. Yet, it is unde-
sirable to let such forms fall through the clausal cracks. Strong analogical reasons exist for including these interactional (and other) elements in the clause periphery. Treating elements such as *ne* and *na ja* in this way results from a generalization of other cases. Two of the properties associated with these cases are (19d-e), namely, that speakers may add to clauses after syntactic closure has been reached, and (thus) syntactic closure may occur more than once per clause. The motivating cases for this observation are not interactional, but grammatical in nature. The clause construction in (20f) provides an example:

(20f) 32  #{ (f) ... (m) ~na { un' dâ haben die uns die Nûmmer gege=ben \ nè / } } wll and there have they us the number given ok
33    .. von= --
      from
34    ... ~also von dieser Stèle \+ }
      so from this place
35    {{ ~wo w' r das Visum krie=gen \+ }}
      where we the visa get

In this construction, the constituents of the clause proper are again similar to those in (20c). The clause is V-2 and the predicate is split. The conjunction + adverb sequence *und da* thus indicates the beginning of the clause, and the infinitive *gegeben* indicates its end. The forms *na* and *ne* are construed as above, i.e. as peripheral elements, and thus receive the appropriate bracketing. The constituents that are of special interest are contained in IUs 33-5. The prepositional phrase in IUs 33-4 (with the relative clause in IU 35) constitutes a postnominal modifier to *Nummer* in IU 32, i.e. it is syntactically part of the accusative NP argument of the verb *geben* 'give'. For approaches that view the clause as a static whole, the syntactic relation-
ship that is realized here is a type of ‘Ausklammerung’ (‘exbraciation’). I will describe the PP that follows in the subsequent IUs simply as a ‘clause addition’. It receives a close bracket (at the end of IU 34) to mark the third syntactic closure point in the construction. The sequence in IU 35 constitutes a second addition to the clause construction. In this case, the addition itself is clausal, so the IU is bracketed with both open and close brackets. Cooccurring features of the NP headed by Stelle (specifically, the demonstrative article) suggest that the relative clause in IU 35 is syntactically embedded. In order to set such an embedded clause apart from nonembedded portions of the construction, a secondary convention is introduced in this example, whereby an embedded clause receives double brackets, i.e. ‘{[‘ and ‘]}’. (Embedding is further discussed below.) Patterned after this grammatical addition is the addition of peripheral elements like oder so in (20d) or ne in (20c), (20e), and at the end of IU 32 in the above example.

Example (20f) also provides the proper material to address the notion of ‘repair’. I have informally referred to ‘repair’ in a number of places in this chapter, and wish now briefly to clarify my views. Broadly conceived, the term apparently encompasses a heterogeneous assortment of structures, including any correction, repetition, substitution, or supplementation in the lexical, prosodic, or syntactic domains. The truncation of IU 33 in (20f)

33 See Rath (1979: 159-84) for a discussion of Ausklammerung in spoken German, and Altman (1981) for a general typology.
34 See Schegloff, Jefferson & Sacks (1977) and Schegloff (1979). Neither work defines ‘repair’ beyond the amelioration of ‘... troubles in speaking, hearing, or understanding ...’ (Schegloff 1979: 261), but both provide many conversational examples. For a more linguistically oriented approach, see Fox & Jasperson (1993), where a repair is said to occur when ‘...
following *von* and the subsequent continuation of the phrase and repetition of the preposition *von* in IU 34 thus would constitute an instance of ‘repair’. As the focus of this chapter is on conversational syntax, however, I would prefer to think of ‘repair’ in a narrower sense, i.e. including at most actual corrections or substitutions in the syntactic domain, but is not clear to me that sharp distinctions like this can be consistently made. Due in part to this problem, in part to the potential for confusion, it seems prudent to avoid the term ‘repair’ altogether. Moreover, as syntactic reformulations are, in general, viewed within the framework of the construction of clauses, there is little need to introduce alternate terminology. In (20f) the addition made after syntactic closure at the end of IU 32 is to be considered a contribution to the clause construction whether the ‘repair’ is present or not. The one significant implication of my decision is that there is a substantive distinction between ‘completed’ and ‘abandoned’ clauses. An utterance which the speaker breaks off, but subsequently attends to grammatically, forms a completed clause construction. One which is broken off and is left syntactically incomplete, however, results in an abandoned construction, so that any following utterance begins an entirely new clause.

Returning now to motivating the treatment of peripheral elements as clausal, the bracketing of initial interactional elements is patterned after a set of cases different from the clause additions in (20f). German has a wide range of structures which involve the periphery of the clause (Altmann 1981: 33), especially at the left edge. An instance of an intermediate initial

*an emerging utterance is stopped in some way, and is then aborted, recast, continued, or re-done*. 

240
boundary results from so-called 'Linksversetzung' ('left dislocation'), which occurs in the corpus excerpt at IU 38:

(20g) 38 #| (H) «.6» ~un' dann eíner /+
       and then one
       39 { .. dèr hat sofort àufgele-gt /+ }
       he has immediately hung up

The clause in (20g) is again V-2 and has a split predicate. It begins with the conjunction + adverb combination, followed by the nominative proform eíner. This is the extent of the IU 38, which thus ends without syntactic closure. IU 39 then begins with the correferential nominative pronoun dèr in the position before the finite verb hat, indicating a (intermediate) clause boundary. An intervening open bracket is therefore placed at the beginning of IU 39. The clause core occurs entirely within IU 39, and the constituents in IU 38 thus fall in the clause periphery. Initial 'particles' are taken to pattern in the same way.

Examples (20f) and (20g) also illustrate observation (19f), that clauses may be longer than one IU. In the above instance, two IUs combine to form one clause construction. The additions of IUs 33-5 in (20f), which are constituent parts of the clause in IU 32, form a clause four IUs in length, even though there are intervening points of syntactic closure. An example of clause construction extending over more than one IU where no intermediate closure was reached is given in (20h):

(20h) 41 #| (H) «.7» und ürgendwann hatten wi' ìànn= /+
        and sometime had we then
        42 so .. fràu d'ran= \+ }
        so one [FÉM] there.on
Once more, the clause is V-2 with a split predicate. The familiar conjunction + adverb in the prefieId opens the construction. At the end of IU 41, however, no syntactic closure has occurred. The expected grammatical object or predicate complement (or both) is yet to be uttered. Only at the end of IU 42 is closure reached, so that while distributed over two IUs, the clause construction contains no internal clause boundaries. It is in such cases that the occasional divergence between prosody and syntax is especially striking.

Continuing now with the exemplification of the observations in (19), example (20i) illustrates observation (19g), that speakers sometimes restart clauses:

(20i) 20 #\{ ... nà ja \+  
         well yes
  21 { und dà=nn% /+  
         and then
  22 { ...(7)(Hx) } und dann (wir wollten dönerstags fliegen } nè /+ )  
         and then we wanted on Thursday fly ok

Here, the speaker begins the construction with an interactionally oriented expression na ja, followed by a conjunction + adverb combination in IU 21. In the next IU, however, this combination is repeated—the clause has in effect been restarted, and so an intermediate boundary is indicated. Note also that there is a second intermediate boundary in IU 22 after the repetition of und dann. The clause is V-2, and the pronoun wir clearly occupies the prefieId, so an internal boundary marking the beginning of the clause core is necessary. Normally, a substantial prosodic break at the juncture of clause periphery and prefieId would be expected when the periphery con-
tains more than a short interactional token (cf. 20f); here the sequence is perceived as one IU. However, notice that there is a volume modulation at this point, so the juncture is indeed prosodically marked, if but minimally. Again, the interaction between prosody and syntax is highlighted by the bracketing in the example.

A hint of what lies behind observation (19h), that syntactic subordination is often not well defined in spoken language, can be found in example (20j).

(20j) 43 #{ ~die meint dann [{ "o Gòtt o Gòtt" \ } nè /+ } she means then oh God oh God ok
44 #{ .. ~"das is' ja nur noch èineinhalb Tà=ge /+ } that is yes only still one and a half days
45 #{ und dann flie=gen S'e schon" /+ } and then fly you already

In IU 43 the reported speech is bracketed as embedded clause. (The fact that there is no verb in this 'clause' will be of no concern here; a discussion follows in the next section.) Grammatically, it can be thought of as the object argument of the verb *meint* (see Munro 1982 for a discussion of this issue), though also relevant here is the extension of the clause construction beyond the reported speech with the particle *ne*. That is, the reported speech occurs in the same IU as *meint* and is linearly embedded in the clause in IU 43.

On the other hand, the reported speech in IU 44-5 is not so obviously subordinate. Both clauses are themselves V-2, they occur in separate IUs, and there is no sequential embedding. Without overt signs of subordina-
tion, I will treat such clauses as independent constructions.\textsuperscript{35} Thus, both are bracketed with ‘#’ at the start of the clause, indicating a new clause construction.

(20k) Corpus excerpt from (20) with clause bracketing.

\begin{verbatim}
20 [D]:  #\{ ... nà ja \+
21   { und dà=rn% /+
22   { -(7)(tx) und dann P \{ wir wollten dönerstags fliegen \} nè /+
23   #\{ (H) «7» nà ja /+
24   { und dienstags war 's noch nicht dà \} nè /+
25   #\{[ ... [“öh” ]] dachten ] wi' \+
26   (Z1):  \{ ò jë= ] /+
27   #\{ “jétz' rufen wi' die mal àn” \[2 nè ] /+
28   (Z2):  \[2 «p mhm p» ] \*
29   #\{ (H) «3» ~und dann wàres aber schon g'rade irgendwie mittags /+
30   oder so /+
31   #\{ ~auf jeden Fall war da kàum noch einer dà= /+
32   #\{ (m)-na {un' dà haben die uns die Nùmmer gët=ben \} nè /+
33   .. von= --
34   ... ~also von dieser Stèlle \+
35   {{ ~wo w'r das Visum krie=gen \+
36   #\{ ~und [haben wi'] dann nachmittags noch ma' ängerufen /+
37   (Z2):  \[ «p mhm p» ] \+
38   #\{ (H) «6» ~un' dann ëiner /+
39   { .. dër hat sofort ëufge=gt /+
40   #\{ .. bi ëiner Nummer hatten wi' n Änrußantworter \} nè /+
41   #\{ (H) «7» und irgendwann hatten wi' dann= /+
42   so .. fràu d'ran= \+
43   #\{ ~die meint dann {"o Gòtt o Gòtt" \} nè /+
44   #\{ .. ~“das is' ja nur noch ëineinhalb Tà=ge /+
45   #\{ und dann flie=gen S'e schon” /+
46   #\{ und dann [ mußten wi' ] gànz schnèll [2 gùcken \} nè ] /+
\end{verbatim}

\textsuperscript{35} See König & van der Auwera (1989: 102-6) for a discussion of V-2 as a indicator of syntactic embedding.

244
At this point, it will be helpful to view the full excerpt in (20) with its clause bracketing so that all clause constructions of passage are clearly shown. Seen as a whole, the bracketed excerpt in (20k) above clearly exemplifies observations (19i-j) regarding the occurrence of clause boundaries with respect to IUs. It seems that quite often, a clause opening coincides with the start of an IU, and syntactic closure is reached at the end of an IU. In this sample passage, clause boundaries are overwhelmingly present in both locations: 20/24 (83%) of the IUs are bracketed at the beginning, and 18/24 (75%) have a close bracket at the end. In §3.5 these and other relationships will be presented and discussed for the entire corpus.

The remaining two observations in (19) cannot be illustrated with the conversational excerpt in (20). For this I must turn to other excerpts. The passage in (21) below provides two examples of observation (19k), that speakers sometimes abandon clause constructions.

The first example of an abandoned clause occurs in IU 47, where Uli begins a new turn. Anna, however, jumps in and Uli acquiesces, effectively truncating the construction he had just begun. His clause remains uncompleted, and so there is no close bracket at the end of IU 47. A frequent source of abandoned clauses involves such interruptions.\textsuperscript{36} The second example occurs in IU 52. Anna still has the floor and she has begun a new clause construction. She is apparently having some trouble formulating her utterance (notice the repetition of \textit{da}), and IU 52 ends up prosodically

\textsuperscript{36} Du Bois (p.c.) reports it is usually the case that speakers eventually finish their clauses. With an interruption, though, there may not be an appropriate opportunity.
truncated. In IU 53 she then starts a completely new clause, as evidenced by
the change in grammatical subject and verb. In other words, she has aban-
donied the clause construction begun in IU 52.

(21) Corpus excerpt from ‘Gorges du Verdon’ [Uli’s visit: 47-54].

47 U: 
#( ... (1.0) » (TSK) (H) «.8» und dalla --
and there

48 A: 
#( 0) ja { sie schreibt g’ra=d’ /+ }
yes she writes just now

49 
#( ... “wir fähren jetzt noch in den --
we travel now still in the
... uh? «.6» in-.. an den Rând der Pyrenâen” /+ }
uh in on the edge of the Pyrenees

50 (U): .. (H) «.’”}

52 
#( .. und dalla = .. da <gibt es> nun --
and there there gives it now

53 
#( .. ‘n .. Verwandter von ihn’n [ wohnt da ] \ } ne /+ }
a relative of them lives there ok

54 (U): 
[ unhunh ] /+

47 U: and there--

48 A: yeah she writes,

49 “we’re traveling now in the--

50 uh in at the edge of the Pyrenees”,

51 (U): (INHALES)

52 and there there’s--

53 a relative of theirs lives there,

54 (U): unhunh,

Finally, observation (19k), that speakers collaborate in creating clause
constructions needs to be illustrated. Collaboration between interlocutors
can take many forms in conversational discourse. Clark & Wilkes-Gibbs
(1986) show how establishing reference is often a joint process. Edelsky
Goodwin (1979) and Lerner (1991) describe the joint production of ‘sentences’. Like Goodwin and Lerner, I focus on the collaboration of syntactic units. The excerpt in (22) illustrates:

(22) Corpus excerpt from ‘Ambassador’ [Uli’s visit: 239-246].

239 [A]:  #\{  .. ö\'\'ner  [ f\'\'r 's ] K\'ochen  /+ \}
          one for the cooking

240 U:   #\{  [ w- w\'ô= denn ] \* \}
          w- where then

241     #\{  .. im% - *  
          in the

242 A:   *  .. in Brazzaville  /+ \}
          in B

243 U:   #\{  ~in Brazzaville  \+ \}
          in B

244 A:   #\{  [ ~ is' ] die [ =\'\' H\'auptstadt von ] K\'ongo  \* \}
          is the capital of Congo

245 (U):  [ ja ] /+
          yes

246 (U):  [\ 2 ja ja ] \+
          yes yes

239 [A]:  one for the cooking,
240 U:     where.
241     in --
242 A:     in Brazzaville,
243 U:     in Brazzaville;
244 A:     [ that's the capital of the Congo.
245 (U):   [ yeah,
246 (U):   yeah yeah;

The syntactic bracketing conventions that I have outlined in this section take advantage of the fact that each speaker’s utterances are usually produced as contributions to their own clause constructions, at least in con-
versational discourse involving a small number of participants. In order to represent the occasional case when two speakers collaborate syntactically, an additional convention must be introduced. I have used an asterisk "*" in the text to indicate a syntactic link across speakers, as happens in IUs 241-2. In IU 240 Uli asks (or requests confirmation) about the location of the events that Anna is describing. Instead of waiting for a response, he begins to answer his own question in IU 241. (There is some indication that he has not been paying full attention here, and he perhaps feels he is supposed to know the answer.) But he stops short, and Anna promptly completes the construction he has begun. The asterisk alerts the reader that despite the speaker change, one clause construction has been produced.

This asterisk convention may also be used within the speech of one speaker to indicate the continuation of a clause construction after some intervening exchange. The excerpt in (23) below illustrates.

Anna has the floor and is describing a house her friends have in town. After giving its location in IU 156, she puts the business of the description on hold in order to ask Uli whether he knows the place the house is located. After he answers at IU 158, she picks up where she left off, in effect adding to the clause construction that was started in IU 156. That the prepositional phrase in IU 158 constitutes a syntactic addition to the previous construction, i.e. that it is to be construed as a clause-peripheral phrase, is

---


indicated by the asterisk pair.

(23) Corpus excerpt from ‘Ambassador’ [Uli’s visit: 155-9].

155 [A]:  #{ ...(.7) und hier wohnen s’e {{ wenn s’e hier sind }} /+
and here live they when they here are

156    #{ ... hab’n s’e eine Eigentumswohnung am =Römerlager \+ }*
    have they a property dwelling on the R

157    #{ ... ‘A is’ Ihnen das ‘n Begriff A’ /+ }
    is you that a concept

158 U:  #{ .. ja ja /+ }
    yes yes

159 A:  * ...(1.0) mit sieben Zimmer d’rin \+ }
    with seven rooms there, in

155 [A]:  and they live here when they’re in town,
156    they’ve got a rowhouse on ‘Römerlager’;
157    do you know that,
158 U:  yeah,
159 A:  with seven rooms;

In this section, I have illustrated several properties of spontaneous speech which need to be taken into account in order to represent the clausal structure of the utterances. I introduced the notion of ‘clause construction’ which incorporates these properties, reflecting the speaker’s ‘online’ creation of syntactic structure. In particular, clause constructions (which are indicated in the transcription with the number sign ‘#’) encompass all constituents and elements which are syntactically peripheral to a clause, as well as the ‘clause core’—the innermost bracketing containing the lexical verb—itself. I discussed the ‘free-bracket’ notation, whereby open and close brackets are placed independently of one another. Points of syntactic closure internal to the clause construction are also indicated with brackets. Two supplemental conventions were also introduced: double
brackets for embedded clauses, and asterisk for linking constructions across
speakers or across intervening exchanges.

3.4 Clause constructions in the corpus

In the previous section, I have outlined a general framework for de-
scribing clauses in spoken discourse. In this section, I detail the nature of
the resulting clause constructions from a syntactic point of view, focusing
on a variety of measures, including words per clause, verb forms per
clause, simple clause type (based on interclausal relationships), and occurrence of intermediate clause boundaries. A prosodic point of view, which
also highlights the temporal component of the constructions, will be con-sidered in the next section, when I compare the coincidence of clause and
IU boundaries.

In the corpus database there are 1099 clause entries, of which 973, or
89%, comprise the basic clausal subset that will be described in this section
and the next. The typology of the whole database which results in this sub-
set is depicted in Figure 3.3.

```
All Clauses
   Non-Native 36 (3%)
       Uncertain 16 (1%)
              Abandoned/Uncompleted 74 (7%)
   Native
       Certain
              Completed 973 (89%)
```

Figure 3.3. Types of clause entries in the corpus database (N=1099).
As was the case with IUs in Chapter 2, the first distinction to be made is between 'non-native' and 'native' clauses. The 36 clauses (3% of all clauses) by the two non-native speakers present during one speech event (see Table 1.1 for details) are treated separately from all other clauses.

Two additional groups of clauses have been separated from the remaining 1063 clauses. First, the 16 clauses which contain uninterpretable words (1% of all clauses) are excluded from the basic subset. (As before, an isolated uncertain 'syllable' did not necessarily disqualify a clause from being included in the main subset.) Second, clause constructions that were completely abandoned—either as a result of interruption or otherwise—or ones that remained uncompleted are also treated separately. There are 55 abandoned constructions (5% of all clauses) and 19 uncompleted ones (2%) in the corpus database. The syntactic characteristics of the remaining 973 'native', 'certain', 'completed' clause constructions are examined below.

A simple, but instructive, measure for establishing an initial characterization of clause constructions is the number of words per clause. Figure 3.4 below presents the distribution graphically. The mean number of words per clause is 6.4 (±3.4), with a range of 1 to 21 words. The median value is 6, and from the figure it is apparent that the modal value is also 6 (or 5, as this value is almost as great). Typically, this near congruence of mean, median, and mode would indicate a relatively unskewed distribution. In this case, however, it is evident that the relatively large number of clauses distributed between 1-3 words/clause balances the number distributed between 8-21 words/clause, i.e. over a much larger range. It is also evident
that the values to the left of the mean are somewhat higher than might be expected, particularly at 1 and 3 words/clause.

The reason for this heightened frequency can be found when we consider what exactly these short clauses are. Recall that there are 'free' adjectival, nominal, particle, and prepositional groups which are used with the force of a clause, that is, as an independent predication. These constitute what I called 'clause-equivalent' constructions, of which there are 106 in the corpus. As they comprise a not insignificant portion (10%) of all clauses, clause equivalents have been included in a representation of discourse syntax. Indeed, it is with their frequency as well as their role in discourse in mind that they have been bracketed just as are clauses containing verb forms (cf. example 20j). Yet, this is not to admit that such construc-
tions are clauses. Without a doubt, there are differences, and the mean number of words per clause is one way in which they differ.

Figure 3.5 re-presents Figure 3.4 with clause equivalent constructions treated as a separate category. The new figure makes clear the contribution of these constructions in elevating the over-all frequency of short clauses. Thus, while some kind of functional equivalence for clauses and clause equivalents can be maintained, asserting a categorial equivalence is not warranted. Due to their markedly divergent structure, clause equivalents will not continue to be included in the calculations of this section. They will, however, be further considered in §3.5 when we revisit the prosodic structures of the corpus.

![Bar graph](image)

**Figure 3.5** Frequency of clauses in the basic clausal subset (N=973) as a function of words per clause, with clause equivalent constructions as a separate category.

253
In the syntactic models that have played a role in the present conception of 'clause' and clause construction—i.e. in the models of valency and topological fields—the verb or verb complex of the clause is central. For this reason, a further description of the clause constructions in the corpus in terms of their verbal structure seems appropriate. The main features to consider are: (a) the presence or absence of verb forms; (b) the presence or absence of a finite verb form; and (c) the presence or absence of a nonfinite verb form. (The position of the verb forms in the clause will be taken into account later.) Figure 3.6, in effect a continuation of Figure 3.3, depicts the classification of the 973 clauses in the basic clausal subset.

![Completed Clauses Diagram](image)

Figure 3.6. Verbal characteristics of clauses in the primary clause subset (N=973).

Two type of verbless clauses are distinguished here. I have already discussed the distinction of clause equivalents and 'proper' clauses. This constitutes the first subclass in the figure. Second, a smaller class of 'true verbless clauses' can be delimited. This class consists of a heterogeneous set of constructions that are more syntactically developed than the simple inde-
dependent groups of clause equivalents. One example would be a clausal sequence in which only the copular verb is lacking. Also included are clausal sequences which are more generally dependent on other constructions for their predicative expressions. There are 34 instances of verbless clauses in the corpus (3% of all clauses).

There are, then, 833 verb-bearing clauses in the corpus. The first distinction to be made at this point is whether a finite (tensed) verb form is present in the construction. The vast majority (799 or 96%) do have a finite verb, but there are 34 clauses that contain only nonfinite forms. This small subclass includes infinitive clauses and constructions which contain only a participle. A more important distinction is to be made among the 799 tensed-verb clauses, namely whether the finite verb stands alone in the clause or whether it part of a larger verb complex. This differentiation will be useful when evaluating the role of the verbal bracket in clause construction, as the presence of a nonfinite verb form implies a verbal bracket in the clause. A total of 225 of the tensed-verb clauses have a verbal complex; the remaining 574 clauses do not.

To summarize, a characterization of the majority class of clauses in corpus database is as follows: completed, 'proper', and verb-bearing, with a finite verb form present, but no verb complex. These 574 clauses constitute 52% of all clauses, 59% of the basic clausal subset, 66% of 'proper' clauses, and 69% of verb-bearing clauses.

The above survey considered features of the clause when viewed as a self-contained entity. Also to be considered is the relationship of clauses to
each other. In the previous section, I distinguished three types of clauses on the basis of interclausal relationships: 'embedded', 'complex', and 'simplex'. I defined an embedded clause as a construction that functions as a verbal argument of another clause, or as one contained within the scope of another clause. A clause is complex if it contains an embedded clause of either type. Other clauses are 'simplex'. (I also recognized a separate parameter [±dependent], which applies to clauses exhibiting other types of syntactic relationships. Dependent clauses which are not, strictly speaking, embedded may be either simplex or complex.) The distribution of these basic clause types will now be considered.

First, let us reexamine the measure of the number of words per clause, this time with an eye on the three basic clause types. Figure 3.7 below depicts the contribution of each type superimposed on the distribution for all proper clauses. From the figure it can be seen that the proportion of each clause type differs for the various words/clause values. Most notably, the sharp peak at 5 words/clause for simplex clauses is absent for complex and embedded clauses. In fact, the distribution of these latter types is bimodal, at 4 & 8 words/clause for complex clauses, and at 3 & 6 words/clause for embedded ones. The substantially lower overall frequencies for these types, however, minimizes the effect of these distributions. In other words, the large numerical superiority of simplex clauses more or less determines characteristic values for the entire set.

Second, let us consider the frequency of the clause types in terms of the verb configuration classes outlined in Figure 3.6.
Figure 3.7. Frequency of proper clauses (N=867) as a function of words per clause, with contributions of basic clause types indicated.

<table>
<thead>
<tr>
<th>Verb configuration</th>
<th>Simplex</th>
<th>Clause type</th>
<th>Embedded</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$-V_{f}/-V_{n}$</td>
<td>33</td>
<td>1</td>
<td>0</td>
<td>34 (4%)</td>
</tr>
<tr>
<td>$-V_{f}/+V_{n}$</td>
<td>20</td>
<td>0</td>
<td>14</td>
<td>34 (4%)</td>
</tr>
<tr>
<td>$+V_{f}/-V_{n}$</td>
<td>448</td>
<td>55</td>
<td>71</td>
<td>574 (66%)</td>
</tr>
<tr>
<td>$+V_{f}/+V_{n}$</td>
<td>186</td>
<td>23</td>
<td>16</td>
<td>225 (26%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>687</td>
<td>79</td>
<td>101</td>
<td>867</td>
</tr>
</tbody>
</table>

Table 3.3. Frequency of three clause types (N=867) with respect to the presence of finite and nonfinite verb forms.

257
Table 3.3 above gives the distribution for the 867 proper clauses. (In the
table, ‘±Vf’ indicates the presence or absence of a finite verb form in the
clause; ‘±Vn’, the presence or absence of a nonfinite verb form.) From the
table it is evident that a good majority of proper clauses (687 or 79%) are
‘simplex’ (and that, as should be expected, a majority of these (448 or 65%)
contain a finite verb form, but not a verbal complex). The balance of
clauses is more or less split between the ‘complex’ and ‘embedded’ (9% vs.
12%), as these two types are by definition mutually dependent.

A few noteworthy points can be made about the distribution of clause
types as given in the table. First, verbless clauses (-Vf/-Vn) are almost al-
ways simplex. In order to understand this fact, it must be pointed out that
embedded clauses without verb forms do occur—two instances are found
in (20) at IUs 25 & 43—but all such cases involve clause equivalent con-
structions, which are not included in the table. The interesting distribu-
tion, then, is the lack of complex clauses with no finite verb form (the two
rows with -Vf). In the corpus there is only one such instance. This fact can
be explained when the nature of clauses without a finite verb is taken into
consideration. As noted earlier, these clauses are typically dependent con-
structions, yet the incidence of dependent complex clauses is very low
(about 5%)—complex clause constructions engender dependency through
embedding, but are themselves rarely the object of similar dependent rela-
tionships. Thus, few complex clauses without finite verb forms would be
expected. Also of interest is the distribution of clauses containing a verbal
complex (+Vf/+Vn). The proportion of such clauses is comparable for sim-
plex and complex clause types (27% vs. 29%), but the proportion of embed-
ded clauses with a verbal complex is somewhat lower (16%). This suggests that embedded clauses in the corpus may be syntactically more simple to some degree, a supposition that is in accord with the observation made above regarding embedded verbless clauses.

The position of the finite verb in the clause is of especial interest, as this is the parameter that is traditionally used to subcategorize clauses (see §3.2). Table 3.4 gives the distribution of V-1, V-2, and V-L clauses among the three clause types previously discussed.

<table>
<thead>
<tr>
<th>Finite verb position</th>
<th>Simplex</th>
<th>Complex</th>
<th>Embedded</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ø</td>
<td>53</td>
<td>1</td>
<td>14</td>
<td>68 (8%)</td>
</tr>
<tr>
<td>V-1</td>
<td>22</td>
<td>4</td>
<td>1</td>
<td>27 (3%)</td>
</tr>
<tr>
<td>V-2</td>
<td>562</td>
<td>72</td>
<td>12</td>
<td>646 (74%)</td>
</tr>
<tr>
<td>V-L</td>
<td>50</td>
<td>2</td>
<td>74</td>
<td>126 (15%)</td>
</tr>
<tr>
<td>Total</td>
<td>687</td>
<td>79</td>
<td>101</td>
<td>867</td>
</tr>
</tbody>
</table>

Table 3.4. Frequency of three clause types with respect to position of the finite verb (N=867).

It is evident from the table that clauses with the finite verb in second position predominate: almost three quarters of the clauses in the corpus are V-2. Given that this order is usually associated with 'declarative' clauses, which will form the linguistic fabric of virtually any conversational story, their frequency should not be surprising. This fact does have consequences, though, for the interpretation of clause constructions, and

---

39 This trend, however, is not significant for p < .01 (calculated for the last two rows of Table 3.3, i.e. absence vs. presence of a verbal complex, $\chi^2 = 4.6067$, d.f. = 2).
will be addressed in the next section. The low occurrence of V-1 clauses is also not surprising. This is a marked order associated with yes-no questions, imperatives, conditionals, and certain 'expressive' clauses, all of which occur in highly specific discourse contexts. The distribution of V-L clauses with respect to clause type is more interesting. Verb-late is usually associated with syntactic dependency (though more specifically it should be linked with the presence of a clause-initial 'subordinator'). In the previous discussion, it was mentioned that complex clauses are not often syntactically dependent. The low frequency of V-L order among complex clauses bears out this observation. On the other hand, a substantial number of both simplex and embedded clauses exhibit V-L order. Because dependency is a separate parameter and embedding is viewed narrowly, both clause types will have verb-late clauses. There is a fundamental difference, however: most embedded clauses (73%) are V-L, while for simplex clauses this order constitutes but a small fraction (7%).

We have seen in examining the clauses in excerpt (20) that clause constructions cannot always be viewed as homogenous wholes, but often have some internal structure. This intermediate clausal structure—represented with the free-bracket notation—is a significant aspect of the character of clause constructions. It is to a description of this internal structure that I now turn.

Clause constructions differ, of course, as to the number of intermediate

\[40\] 100% of the 126 clauses with V-L order in the corpus begin with a 'subordinating conjunction' or relative pronoun.
clause boundaries they contain. Table 3.5 gives the figures for the 'proper' clause subset. From the table it is evident that a majority of clause constructions do not, in fact, have any internal clausal structure as marked by clause bracketing. In other words, they are produced—from a syntactic point of view—as a single unit. A substantial minority (36%), however, does have some internal bracketing—usually one additional boundary, and a maximum of four. The excerpt in (20) illustrates the full range: the clause construction consisting of IU 20-2 has four intermediate brackets, while that consisting of IU 31 has zero.

<table>
<thead>
<tr>
<th>Number of intermediate boundaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td>554</td>
</tr>
<tr>
<td>(64%)</td>
</tr>
</tbody>
</table>

Table 3.5. Frequency of intermediate clause boundaries in proper clauses (N=867).

Intermediate brackets are not distributed proportionately across clause type, but vary significantly with it. Table 3.6 below presents the figures. While 260/687 (38%) of the simplex clauses and 31/79 (39%) of the complex clauses have one or more intermediate brackets, only 22/101 (22%) of the embedded clauses do.\(^{41}\) Again, embedded clauses stand apart, suggesting that they are qualitatively different from the other types.

\(^{41}\) This trend is significant for \(p < .01\) (calculated for column 1 vs. columns 2-5 of Table 3.6, i.e. absence vs. presence of an intermediate boundary, \(\chi^2 = 10.2209, \text{d.f.} = 2\)).
<table>
<thead>
<tr>
<th>Clause type</th>
<th>Number of intermediate boundaries</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Simplex</td>
<td>427</td>
<td>192</td>
</tr>
<tr>
<td>Complex</td>
<td>48</td>
<td>26</td>
</tr>
<tr>
<td>Embedded</td>
<td>79</td>
<td>20</td>
</tr>
<tr>
<td>Total</td>
<td>554</td>
<td>238</td>
</tr>
</tbody>
</table>

Table 3.6. Frequency of intermediate clause boundaries with respect to clause type (N=867).

Finally, of interest is the syntactic composition of the clause periphery, that is, of the syntactic fields that are created by the intermediate bracketing. Typically, clause-internal bracketing results in just one constituent (or element) falling in the periphery. Out of the 238 clause constructions with one intermediate bracket, 208 (87%) meet this description. The type of constituent in the periphery does vary, however, as does its occurrence relative to the clause core. Table 3.7 below lists the most frequent peripheral elements for initial and final peripheral fields. We know from the previous discussion that most clause constructions do not have a periphery (see Table 3.4). More specifically, from Table 3.7 we see that 85% of proper clauses do not have an initial field, and 75% have no final field. Still, clause constructions that do have an initial peripheral constituent or element, as well as those that have a final peripheral one, occur often enough to justify a consideration of their basic structure. Furthermore, these positions will be important in the following section when we examine the interaction of prosody and syntax, and in Chapter 4 when information flow questions are discussed. It is apparent that particles are

262
the most frequent peripheral elements, both in the initial and final fields. This fact is to be attributed to the regular occurrence of common 'Gliederungssignale' in the conversational stories. There are, for example, 52 instances of *ja* in the initial field (39% of all initial elements), and 84 of *ne* in the final field (38% of all final elements). Again, excerpt (20) provides typical examples; see (20k) for the full bracketing. The first two clauses (IUs 20-2 & IUs 23-4) each contain an initial peripheral field with *na ja*, and a final peripheral field with *ne*.

<table>
<thead>
<tr>
<th>Element in periphery</th>
<th>Ø</th>
<th>Particle</th>
<th>NP</th>
<th>PP</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial peripheral field</td>
<td>738</td>
<td>82</td>
<td>24</td>
<td>5</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>(85%)</td>
<td>(9%)</td>
<td>(3%)</td>
<td>(1%)</td>
<td>(3%)</td>
</tr>
<tr>
<td>Final peripheral field</td>
<td>646</td>
<td>149</td>
<td>43</td>
<td>42</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>(75%)</td>
<td>(17%)</td>
<td>(5%)</td>
<td>(5%)</td>
<td>(3%)</td>
</tr>
</tbody>
</table>

Table 3.7. Frequency of peripheral elements in clause constructions (N=867).

In sum, a typology of the clause entries in the database produced 973 completed clause constructions whose lexical contents are known. A consideration of the number of words per clause led to the separate treatment of the 106 clause-equivalent groups and 867 'proper' clauses. This latter subclass was further classified on the basis of the presence of finite and nonfinite verb forms. It was found that a majority of clause constructions possess a finite verb, but not a nonfinite one. That is, there is no verbal complex in most clauses. Clause constructions were also be subclassified into simplex, complex, and embedded types. It was found that the simplex
type overwhelmingly dominates. The traditional classification of clauses based on the location of the finite verb is also separately considered. Finally, the nature of the clause periphery was explored. While most clause constructions were found not to have either an initial or a final peripheral field, it is useful to have sketched the basics, as these positions will play a role in Chapter 4 when we address issues of information flow in the clause.

3.5 Clause constructions and intonation units

In the previous section, a characterization of clause constructions was given in terms of verbal structure, clause type, verb position, and intermediate clause boundaries. With this background, we are now prepared to examine clause constructions from the point of view of the prosodic structure. In conjunction with the explanation of clause constructions and the free-bracket notation in §3.3, we have already seen some examples of how clauses may be constructed over more than one IU. Here, a systematic comparison will be undertaken in order to gain an accurate picture of how speakers put together clauses over the course of an extended amount of connected speech. As usual, the focus will be on majority, rather than on specific, cases.

The first point to be made when comparing clauses and IUs is that, on the whole, IUs are somewhat shorter in scope than clauses. One way to show this is to compare the two units is in terms of words per unit. In Figures 3.4-3.5 & 3.7 above I examined the distribution of clauses in terms of words per clause. Adjusting for clause equivalents, the mean number of words per clause is 6.8 ± 3.3 (median = 6). For IUs, after adjusting for
backchannels and clause equivalents, the mean number of words per IU is 4.3 (median = 4).\textsuperscript{42}

![Graph showing frequency of IUs in terms of words per IU with backchannels and clause equivalents as separate categories.]

Figure 3.8. Frequency of IUs in the primary corpus subset (N=1736) as a function of words per IU, with backchannels and clause equivalents as a separate categories.

Figure 3.8 gives the distribution of IUs in terms of words per IUs, with backchannels and clause equivalents treated as separate categories. Evident from the figure is the much shorter range that IUs exhibit: 1-12 words, as opposed to 1-21 words for clauses. Both Altenberg (1987: 25) and Chafe (1993) remark on the relatively narrow range of IU size.

In Figure 3.8, there is a clear modal value of 3 words per IU for IUs which are parts of proper turns and clauses.

\textsuperscript{42} This figure compares favorably with the mean of 4.5 words/tone unit reported by Altenberg (1987) for a much larger corpus of British English, and with the 4.8 words/IU reported by Chafe (1993) for a comparable corpus of American English discourse.
Recall that when a measure of syllables per IU was considered in §2.4 (see Figures 2.7 & 2.9), there were three peaks in the distribution of IUs: at 2, 4, and 8 syllables per IU. In that section, the contribution of backchannels to the over-all distribution was discussed. Here, we may consider the effect of clause equivalents. Interestingly, the three peaks still remain when the contribution of clause equivalent constructions is subtracted, as is illustrated in Figure 3.9. Although the highest peak does shift to 4 syllables per IU (from 2 syllables per IU), which is more in keeping with the mean number of words/IU, an explanation for this trimodal distribution will obviously require more than an appeal to various subclasses of short IUs.
Fortunately, a more general, syntactically based source presents itself. As we saw earlier in the chapter, IUs having the approximate scope of a word, a group, or a clause are all fairly common (see Table 3.2). Given the hierarchical relationship involved, with the consequential increase in unit size—words form groups, which in turn form clauses—it might be suspected that the three peaks in Figure 3.9 correspond to the maximums of each of these units. Indeed, the modal value of IUs with the scope of a word is 2 syllables/IU; of a group, 4 syllables/IU; and of a clause, 8 syllables/IU.

Figure 3.10. Frequency of IUs in the primary corpus subset (N=1736) in syllables per IU, with the individual contributions of IUs having the scope of a word, group, or clause.
Figure 3.10 above replots the bar graph of Figure 2.7 (representing all IUs in the primary corpus subset) with each contribution of word, group, and clause-level IUs separately indicated. From the figure, it is clear that the correspondence of syntactic unit and peak in frequency is very good for word and clause. A good majority (76%) of the IUs that are 2 syllables in length are word units; a greater proportion (84%) of the 8-syllable IUs are clauses. The situation is less straightforward for 4-syllable IUs, as both group and clause-level IUs contribute substantially to their frequency. However, group-level IUs do make their maximum contribution at 4 syllables/IU. In other words, each syntactic unit has its own characteristic effect on the average size of IUs.

<table>
<thead>
<tr>
<th></th>
<th>&lt;1</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>&gt;6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aligned</td>
<td>28</td>
<td>451</td>
<td>213</td>
<td>79</td>
<td>40</td>
<td>8</td>
<td>7</td>
<td>3</td>
<td>829</td>
</tr>
<tr>
<td>Unaligned</td>
<td>20</td>
<td>11</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>38</td>
</tr>
</tbody>
</table>

Table 3.8. Frequency of clause constructions (N=867) in terms of IUs per clause for aligned and unaligned constructions.

Based on the approximate difference in size of IUs and clauses, an appropriate metric to be examined is the number of IUs per clause. A calculation of the range, mean, and mode values with respect to IUs will provide a good basic indication of the distribution of clauses across these prosodic units, and thus of their temporal structure as well. Table 3.8 gives the frequency of clause constructions differentiated by IUs per clause. In tabulating this measure, it becomes apparent that there are two different patterns.

---

43 In this tabulation, for complex clauses the full number of IUs in the construction (i.e. including embedded clauses) is counted. Embedded clauses are also counted separately.
that are best kept separate; they are labeled 'aligned' and 'unaligned' in the
table. The initial and final boundaries of an 'aligned' clause construction
coincide with IU boundaries, whereas those of a 'unaligned' clause con-
struction do not.44

The number of IUs per clause (construction) ranges from less than one
(the first column of Table 3.8) up to 13. There are, however, very few con-
structions over 4 IUs long (18 or 2%). The mean value is 1.7 (± 1.1)
IUs/clause; the median is 1, and the mode is also 1 IU/clause.45 Clearly, the
vast majority of clause constructions (829/867 or 96%) have boundaries
that align fully with IU boundaries. Moreover, of the clause constructions
which are prosodically aligned, 451 (54%) are coextensive with the scope of
a single IU. There is, then, something to be said for Halliday's (1967b: 201)
distinction between marked vs. unmarked 'tonality', the unmarked case
being when a clause is realized as a single tone group (See also Altenberg
1987; Chafe 1987). Furthermore, it is perhaps the prevalence of this pattern

44 Included in the 'aligned' category are 28 embedded clauses that are contained within a
construction that is 1 IU long, e.g. [Uli's visit: 259]:

[U]: #{ ~ich weiß nich' [{ ob das= .. dört dasselbe is' \+ }]
  I know not whether that there the same is
  'I don't know if it's the same there'

Because the construction as a whole aligns (and full alignment of the embedded clause is not
possible in this case), it was felt that such embedded clauses are qualitatively more similar
to aligned clauses of 1 IU or more than those where alignment is really lacking; see excerpt
(18) for an example.

45 The mean value represents only the 'aligned' constructions, as it is difficult to quantify
the unaligned ones. However, using an arbitrary value or 0.5 for the fractional parts, the
mean does remain the same.
which leads many linguists not familiar with discourse data to believe that such clauses constitute the 'normal' pattern. Because single-IU constructions generally correspond to simple sentences in German, their overall frequency may reinforce the dominant focus on written language (cf. Linell 1982), which results in a reduction in the attention paid to the dynamic aspects of speech. There is little need to devote the extra necessary effort required in the study spoken discourse if the linguist's (or native speaker's) intuition will supply the same kind of data.

This view oversimplifies matters for a number of reasons. First, there is the flipside to distribution: nearly half of the constructions (416 or 48%) are not coextensive with IUs. In other words, there is simply more to be said. Second, although single-IU clause constructions are indeed common, some contain internal clausal structure: 73 (16%) have an intermediate clause bracket indicating one or more elements in the prefield or postfield, or the presence of an embedded clause, as is the case for the constructions in IUs 25, 27, 40, 43, and 46 in example (20). Third, even among the 378 single-IU constructions without internal bracketing, there seem to be certain striking limitations to their occurrence in discourse. They tend to be produced singly or in pairs readily, but only seldom are stretches of spontaneous conversational discourse found—even when perceptually very fluent—which have been uttered completely in single-IU clauses. There are only a handful of instances in the primary corpus where a sequence of three or more of these clause constructions is produced. Excerpt (24) illustrates one occurrence:
(24) Corpus excerpt from ‘Ambassador’ [Uli’s visit: 142-6].

142 [A]:  #{ ...(1.2) denn wär’n s’e in Wàshington /+ }
then they were in W

143    #{ ... die sprèchen beide gut .. ënglisch /+ }
they speak both good English

144 (U):  .. mhm \+

145    #{ ... und s- dånn sprechen sie bèide .. fließend franzòsisch /+ }
and then speak they both fluent French

146    #{ ~dann war’ s’e ...(7) in Wàshington /+ }
then were they in W

142 [A]:  then they were in Washington,  
143    they both speak English well,      
144 (U):  mhm;                         
145    and they both speak French fluently,  
146    then they were in Washington,  

The distribution of single-IU clause constructions lends support to the view that data from spoken discourse should supplement intuition not only when issues of language use are involved, but also when language as a dynamic system is in focus. The fact that such clauses do not usually occur in straight succession points to a richness of language that usually remains unexplored in linguistics. It is not only at points of nonfluency or problematic communication that there is more going on in the production of clauses. Syntactic-prosodic variation is interwoven in the fabric of discourse itself.

Of great interest are the specifics of the prosody–syntax nonalignment. There are two basic classes of nonalignment to consider. Obviously, the 38 cases of ‘unaligned’ clause constructions feature prominently here. But, as these cases constitute but a fraction of all constructions, they do not merit

271
detailed consideration here. Two brief points can be mentioned, though. It might be suspected that this kind of nonalignment is due to constructional adjustments. However, in only 6 of the 38 cases (16%) contain a lexical repetition or syntactic reformulation. One thing is apparent that could be pursued if there were more instances to consider. Although most of the speakers produced at least one unaligned construction, two speakers produced one third of the total. The halting speech of one participant in particular (Günther, see excerpt (18) for an typical example) is striking for the number of mismatches between prosody and syntax. It may be suspected that this class of nonalignment is partly idiosyncratic in nature, and partly due to language-external factors. While all speakers will occasionally construct a clause which does not align with the basic prosody of the utterance (for whatever reason), some speakers will produce characteristically unaligned clauses.

More interesting due to their higher frequency of occurrence are the 350 instances of aligned clause constructions that are produced over the course of two or more IUs. There are several different ways in which clause constructions are distributed across IU boundaries. We have already seen how peripheral elements may be produced in their own IUs. In the corpus there are 96 (27%) instances of prefld constituents occurring in separate IUs, and 80 (23%) instances for postfield constituents. As above, only a small number of splits (9 or 3%) can be associated with lexical repetition or syntactic reformulation. Perhaps surprisingly, the majority of IU breaks involve the clause core; there are 197 (56%) cases of the core of a clause construction being produced in more than one IU. It can be noted that these
breaks are far from haphazard. There are only 12 (3%) breaks within a group boundary. In only 16 (5%) instances is the clause core produced in roughly equal portions in each IU. In most cases (169 or 86% of a split clause core), one or two constituents or elements are produced separately, but the remainder of the clause is intact.

As was the case in the previous section with the peripheral syntactic fields, there seems to be a substantive difference in construction parts, depending on whether they occur initially or finally. When initial in the clause construction, prosodically separated elements are most likely to include conjunctions (63 cases) or NPs (35 cases). Adjectives and adverbs (21 cases) and PPs (17 cases) occur less commonly. When final, the prosodically separate elements are likely to include NPs (41 cases) or PPs (38 cases). There are also a number of verbal elements (33 cases). Adjectives and adverbials are not nearly as frequent (10 cases). The import of these distributions is not apparent from the syntax—or from the prosody—alone, as there are a great many individual patterns. A partial explanation will be given in Chapter 4, when the discourse pragmatics of the IUs and clause constructions is examined.

3.6 Summary

In examining spoken language, it is apparent that much of the syntactic structure a speaker produces has an inherent temporal component. To the extent that they are present, sentences do not exist in a pre-defined way, but are constructed, on-line, by the speaker, intonation unit by intonation unit. A substantial proportion of clauses, too, are not produced within the scope
of a single IU, but are instead formed over the course of two or more IUs. On the other hand, smaller syntactic units like group and word are rarely split over IUs. Yet it the case that these smaller units are often prosodically separate—produced in the periphery of larger constructions. I suggested that what happens at the periphery of clauses—syntactically, prosodically, and interactionally—constitutes a legitimate part of their description, and thus introduced the notion of 'clause construction' to be able to incorporate all linguistic elements that have clausal relevance. I also introduced the 'free-bracket convention' to represent these clause constructions. The syntactic looseness that the on-line construction of clauses exhibits is not to be viewed as a problem which causes difficulty for the (abstract) representation of grammatical structure in spoken language, but rather as a fact of the language, to be described in its own right.

After exemplifying the mechanics of bracketing clause constructions, I examined a number of distributions in the primary corpus. Considering the distribution of syntactic (clausal) boundaries with respect to the 'clause core' enabled a focus on the prefield and postfield positions in the clause. I then discussed the occurrence of syntactic-prosodic boundary alignment. The instances of complete nonalignment of prosodic and syntactic boundaries were found to be minimal. It was found instead that most IU junc-
tures occur at points of syntactic closure, but a substantial number of clause cores were split across two IUs. These splits were not random; consistently, only one or two clausal elements were produced prosodically separate. The significance of these features of clause constructions will be explored in the next chapter.
4. Discourse Pragmatics

Erst im Dialog tritt der Zweck und die Absicht auf, in einer bestimmten Weise auf eine angerede Person einzuwirken. ... Der Zweck unseres Sprechens ist stets der, den Willen oder die Erkenntnis einer Person so zu beeinflussen, wie es dem Sprechenden als wertvoll erscheint.

It is only in dialog that the purpose and intention of affecting an addressee in a certain way is evident. ... The purpose of our speaking is always to influence the will or cognition of a person in a way that appears important to the speaker.

Philipp Wegener
Untersuchungen ueber die Grundfragen des Sprachlebens, 1885, p.66-7.

In this chapter I examine aspects of the discourse-pragmatic constitution of the utterances in the primary corpus. Issues in discourse pragmatics are discussed first, with Chafe’s cognitive model of information flow and the properties ‘activation state’ and ‘identifiability’ taking center stage. Next, I turn to a discussion of the model in light of my German data. A distinction between ‘concepts’ and ‘referents’ is shown to be helpful in accounting for certain activation patterns. Transitional and steady-state configurations of activation are discussed next, with problem areas in the determination of activation state addressed. After outlining the revised model, my attention then turns to a systematic analysis of the corpus, with a focus on the noun phrase mentions that present ‘new’ concepts in the discourse. The focus then shifts to the subset of new mentions that also introduce referents into the conversational stories, the ‘tracking mentions’. The distribution of these referents is contrasted with the basic prosodic and syntactic structures from the previous chapters. Finally,
their occurrence is evaluated in light of several hypotheses that have been put forth regarding information flow constraints.

4.1 Pragmatics in linguistics

Ostensibly, pragmatics is the study of language use. But behind this 'definition' lies a variety of enterprises. On the one hand, this is no different from other areas of linguistics, as we have seen in Chapter 2 for prosody and in Chapter 3 for syntax. Depending on one's fundamental assumptions about the relationship between langue and parole, appropriate methodology, immediate goals, and so forth, the approach one takes will to a greater or lesser degree resemble the approaches of others. Yet pragmatics as a field seems to be more varied, because, as Levinson (1983: 9) points out, the term '... covers both context-dependent aspects of language structure and principles of language usage and understanding that have little or nothing to do with linguistic structure'. The basic substance to be described is itself quite vast, if not heterogeneous. Moreover, demarcating pragmatics from other areas of linguistics, especially from semantics, is troublesome. Questions of context—in whatever guise—seem to pop up at every step through the so-called core dimensions of the field.

In attempting to instill some order into the pragmatic morass, it has become common to describe as pragmatic '... those aspects of the meaning of utterances which cannot be accounted for by straightforward reference to the truth conditions of the sentences uttered' (Gazdar 1979: 2). This way of thinking has led to a broad research program investigating the linguistic consequences of speech acts, illocutionary force, presupposition, and
conversational implicature—topics which are now often viewed as synonymous with pragmatics itself. In order to keep the term ‘pragmatics’ for general use, however, we might label the speech act enterprise ‘propositional pragmatics’ for its affinity to (and historical roots in) logic and the philosophy of language.

A wholly different enterprise, which may be called ‘discourse pragmatics’, is concerned not with the implied meaning of utterances per se, but with the linguistic devices manipulated by speakers (and listeners) in ordinary language use. Discourse pragmatics can be viewed as an essential component of a larger effort seeking to explicate the ways in which language use shapes the very systems underlying this use. Familiar semantic/pragmatic notions such as deixis and reference play an important role in discourse pragmatics, but they are considered from a thoroughly contextual perspective, i.e. as part of natural discourse in situ. In other words, they are not strictly defined in terms of logical relations, e.g. between linguistic expressions and entities in some possible world; rather, they emerge from the fabric of an evolving text, and—most significantly—they are taken to involve crucial cognitive and interactional components which are relevant to the explication of particular form-function connections.

4.2 Information flow

Central to discourse pragmatics is the notion of ‘information flow’. The ideas underlying this notion are most clearly developed in Chafe (1979) in the context of presenting a model for describing the ‘verbaliza-
tion' of recalled experience. Metaphorically, this 'flow model' (as opposed to the 'hierarchical model') seeks to represent the way thoughts flow in and out of consciousness and how language flows with them. In the model, the role of the speaker predominates, since it is the speaker '...' who is engaged in a real-time process of focusing on a sequence of ideas and converting these ideas, one after another, into language' (166). It is exactly this real-time cognitive process inherent to the flow model that has been a major concern in my examination of discourse prosody and syntax. It is just this aspect of the model which lends itself to a contrast with the linguistic structures that speakers form in the course of expressing their thoughts. Information flow, then, has to do with the 'packaging' of thoughts, specifically with '... the speaker's assessment of how the addressee is able to process what [the speaker] is saying against the background of a particular context' (Chafe 1976: 27). An important part of the continual assessment speakers make in formulating their utterances involves the distinctions between 'given' and 'new' information, 'identifiable' and 'nonidentifiable' referents, and related notions. I will be concerned with the elaboration of these distinctions for the remainder of this section.

4.2.1 Given and new information

There are many conceptions of 'given' and 'new' (and of the related notions of 'presupposition' and 'focus' as well) in the linguistic and psycholinguistic literature, and many ways in which these terms have been

Two important comparative points can be appropriately made here, however. First, there are—at the most fundamental level—just two primary orientations in approaches incorporating a given/new distinction. Putting aside the question of how the speaker and listener fit into the equation, information status in the above-mentioned works can be viewed as a configurational property in either a textual or a cognitive process. In a textual process, information status values depend largely on the external, product-oriented facts of the speech act, without any special regard for the cognitive states of the human participants. Given informa-

\textsuperscript{1} This is not to mention terms such as 'topic', which often exhibit a significant overlap with the value 'given'; for a variety of positions, see the discussions in Reinhart (1981), Fries (1984), van Oosten (1986), and Schlobinski & Schuetze-Coburn (1991).
tion is that which is '... recoverable anaphorically or situationally' (Halliday 1967b: 211); it is '... already in the discourse-model ...' (Prince 1981: 236), and so is '... present in, and hence retrievable from, the immediately relevant preceding context ...' (Firbas 1992: 37). New information lacks this connection to the context. Information status values under this group of views can be mechanically derived (although different researchers bring different sets of assumptions to bear on their derivation), and in limited cases may be expressed algorithmically (see e.g. Pignataro 1987). In a cognitive process, information status values depend instead on the ever-changing internal mental states of the speech act participants. Each instance represents a hypothesis about a momentary cognitive gestalt. Given information is that which is '... assume[d] to be in ... consciousness ... at the time of the utterance' (Chafe 1976: 30). New information is assumed not to be present. The operationalization of this conception of the given/new distinction is more difficult, but there are cogent reasons to include consciousness in an account of information flow, and so models which incorporate the cognition of the speaker and listener in some way are to be preferred.

A second difference among approaches—one that I have repeatedly appealed to—has to do with the applicability to spoken language in general, and to conversational discourse in particular. Again, a number of models have been developed from and only applied to isolated sentences or sentence pairs, otherwise experimentally controlled language, or written texts. Very few have been fully annealed in the fiery arena of spontaneous, spoken discourse. Those accounts which contain a written lan-
guage bias are of dubious—at the least, unproven—value for the discourse-pragmatic analysis of the present corpus. To the extent that the production and comprehension of spoken and written language differ, models which are based on spoken discourse are to be preferred.

The speaker-oriented model proposed by Chafe (1974, 1976, 1979, 1980b, 1987, 1993) that I mentioned above weighs in positively on both counts. It is thoroughly cognitive in orientation and has developed from a close analysis of spontaneous speech. Importantly, recent articulations of the model have been applied to discourse from a number of languages, including Chamorro (Scancarelli 1985), Chinese (Tao 1993), English (Chafe 1987, 1993; Du Bois 1987), French (Lambrecht 1986; Ashby & Bentivoglio 1993), German (Schuetze-Coburn 1987b), Papago (Payne 1987), Seneca (Chafe 1993), and Spanish (Ashby & Bentivoglio 1993). Chafe’s model is thus especially promising as part of an account of information flow in conversational discourse, and it is to his system that I now turn. In the next section, I sketch the cognitive approach to the given/new distinction (in this context, called ‘activation state’), and in the following section, the equally important notion of identifiability. Other aspects of the model are subsequently mentioned as well. Afterwards, I critically examine Chafe’s proposals and suggest emendations to aid in their consistent application to conversational discourse.

4.2.2 Activation state

The cognitive model of information flow advanced by Chafe seeks to account for what happens in the minds of speakers during the production
of spoken discourse and for the linguistic effect on the ensuing discourse. The determination of information status values is naturally speaker oriented because the values are '... based precisely on a speaker's assumption as to what is in [the] addressee's consciousness at the time of speech' (1974: 111). In the model, values apply to 'ideas of objects, events, and properties' (1987: 25)—in other words, to the mental representations of such entities, not to the real-world referents or linguistic expressions thereof. (Note that Chafe uses 'idea' as a technical term; see §4.3.1 below, where I examine this and related notions in detail.) The specific value an idea is said to have is thus tied directly to the speaker's estimation of the idea's status in the mind of the listener, a process which in essence reflects the 'activation' of ideas in the speaker's own mind. Consequently, the given/new distinction is called 'activation state'.

2 In particular, ideas with the value 'given' are those that the speaker believes are already activated for the listener, i.e. those that are already in the listener's 'focus of consciousness' at the time an utterance is made. Ideas with the value 'new' are those that the speaker believes are as unactivated in the listener's consciousness prior to the point of the utterance. The basic linguistic conse-

---

2 The given/new distinction is subsumed under the notion of 'activation state' in Chafe (1987); however, it is discussed in terms of 'activation cost' in Chafe (1993). It seems to me that the values 'given' and 'new' (or 'active' and 'inactive') do not fit the cost metaphor, for which values like 'cheap' and 'expensive' would be necessary. Thus, I continue to use the label 'activation state' in the following discussion without denying the suitability of describing the activation process as the amount of mental effort ('the cost') required to bring an idea into consciousness. I might add here, however, that 'cost' seems to imply a continuum and 'activation' quantum states, so that there may be theoretical ramifications for using one term or the other.
quences of this difference can be stated as follows: expression of given ideas is generally made using a form which is ‘attenuated’ both prosodically (i.e. weakly accented or unaccented) and phonologically (i.e. pronominalized or elided), while new ideas are generally expressed with unattenuated forms, usually accented full noun phrases.

In order to illustrate the application of Chafe’s model, it will be helpful to examine some discourse from the primary corpus. Example (1) presents three IUs near the beginning of a conversational story by Tamina in which given and new information in the sense just described can be distinguished.

(1) Corpus excerpt from ‘Aaldi drunk’ [Skt: 10-12].

10  T:  #{ ... ~da wår’n wir nämlich gëstern \+ }
     there were we namely yesterday

11  #{ ... da war so ‘n Besöfener im AÌldi= /+ }
    there was like a drunk in the A

12  #{ ...(1.0) und der hât .. so ‘ne Rôle Kà=se \+ }
    and he has like a roll cheese

10  T:  we were there yesterday;
11  there was a drunk in the Aldi <supermarket>,
12  and he <PAST> a roll of cheese;

(As an ad hoc notation in this example and the next, I have underlined the expressions referring to new ideas.) For each IU, the determination of given vs. new information begins by identifying the expressions which convey ideas of objects, events, or states. For the moment, we can

---

3 I am using ‘object’ as a cover term for persons, other animate beings, things, and abstractions. Likewise, ‘event’ includes any dynamic situation: actions, processes, as well as
assume that these ideas are expressed predominantly by noun, verb, and adjective phrases, or through the proforms of such expressions.

In IU 10 there are three expressions of this sort: the locative proform *da* 'there' (referring to the discount supermarket 'Aldi'), the pronoun *wir* 'we', and *gestern* 'yesterday'. Thus, three ideas can be distinguished: the idea of Aldi, the idea of Tamina (together with whoever had accompanied her), and the idea of (being somewhere) yesterday. In the course of initiating the story during the preceding nine IUs, Tamina had made reference to these ideas. Because of these recent prior mentions, the three ideas are assumed to have a status 'given'. For Tamina herself, they probably have not yet faded from consciousness. More importantly, she could hypothesize that the ideas were most likely still active in the minds of her interlocutors, Käthi and Wolf, as well. During the formulation of IU 10, then, the three ideas are said to be refreshed in the speaker's consciousness; and upon their being uttered, the speaker assumes that they are refreshed in the listeners' consciousness as well. Note that *da* and *wir* are attenuated forms: both are unaccented proforms. On the other hand, *gestern* does bear the primary accent of the IU—more on this in a later section (see *Explaining linguistic form*).

In IU 11 two ideas—that of a drunkard and that of Aldi—can be identified. The first, expressed by the noun phrase *'n Besoffener 'a drunk', constitutes new information, while the second, expressed in the prepositional phrase *im Aldi*, is given. That is, before the formulation of this IU,

---

events; and 'state', any static situation or property (cf. Lyons 1977: 483).
the idea of Aldi is taken to be already active in Tamina’s, Käthi’s, and Wolf’s consciousness on the strength of the immediately prior mention. In contrast, the idea of a drunkard, being expressed here for the first time, has not yet been activated. The process unfolds as follows: at some point during which Tamina was formulating the utterance, she retrieved the idea of a drunkard from memory, and thereupon it was active in her consciousness. Then, after she had uttered the IU, she can assume that she has introduced the new idea into Käthi’s and Wolf’s consciousness, as well as refreshed the already active one. Here, both ideas are expressed by phrases that contain accented nouns (again, more on this below).

In IU 12 there are again two ideas: those of the drunkard and of a cheese roll. In the same way that the idea of Aldi was refreshed, and the idea of a drunkard was retrieved for the speaker and introduced for the listeners in the previous IU, here the idea of the drunkard is refreshed and the idea of a cheese roll is retrieved/introduced. In this case, though, the given information is expressed by an unaccented pronominal form (der), while the new is expressed with an accented full noun phrase (’ne Rolle Käse).

In the above passage, the two ideas that had to be activated were both ideas of objects. It should be noted that, as defined, ideas may also be representations of events or states. In (2) below, the continuation of (1), we see instances of such ideas being newly activated.

---

4 The form hat glossed as ‘<pAST>’ in IU 12 is the auxiliary verb haben, not the homophonous lexical verb.
(2) Corpus excerpt from ‘Aldi drunk’ [Skt: 13-8].

13 [T]: .. ~so 'ne .. beschëu=erte Rölle Hårzer Käse \wà /+ like a crummy roll Harz cheese ok

14 für [siebzig] Pfènnig \+
for seventy cents

15 (K):  
\[ <p \ mhm \ p > \]

16 nich' mal \+
not once

17 (H) (%) .. «.7» ~hat er irg'ndwie ãu=gemacht \+ }
has he somehow opened

18 das Papier \} wa \+ }
the paper ok

13 [T]: a crummy roll of Harz cheese,
14 for seventy cents;
15 (K): mhm
16 not even [that];
17 somehow he opened;
18 the wrapper;

In IU 13 the property expressed by the adjective beschëuerte ‘crummy’, which is important to the understanding of the story, is activated. Prior to the planning and uttering of this IU, the idea was inactive for the speaker and listeners, but after the IU has been uttered, it has an active status. In IU 17 the idea of an event of opening (something), expressed by the past participle aufgemacht ‘opened’, is activated. Again, prior to the planning and uttering of this IU, the idea was inactive, but afterwards it is active. Thus, activation for ideas of states and events, expressed by adjectives and verbs, is the same as for ideas of objects, expressed by nominals. Although I will occasionally make reference to the activation of ideas of nonobjects in the following exposition, the majority of the questions to be answered

286
about information flow in the corpus have to do with the distribution of ideas of objects.

There is another important aspect of activation state in Chafe's model that has yet to be discussed, namely the value 'accessible'. It is assumed that the nature of consciousness is such that only a small number of ideas can be in an active state at any one time (Chafe 1987: 22; 1993: 52). A major consequence of this assumption is that ideas cannot remain given indefinitely—or even during the course of a conversation—but must continually be refreshed through subsequent mention, as illustrated above, or they gradually fade from consciousness. When ideas are no longer in the focus of consciousness, they need not be considered inactive, however. They may still be at the periphery of consciousness, in a 'semiactive' state. Such ideas are termed 'accessible' by Chafe. Example (3) below, a further continuation of the story in (2) and (1), illustrates one situation where this notion comes into play.

In IU 19 there is the mention 'ne Verkäuferin 'a salesclerk', expressing the idea of a salesclerk, which by IU 20 is active for the listeners. The next mention of the salesclerk does not occur, however, until IU 43. By this time, the idea of the salesclerk can no longer be in the focus of consciousness of the speaker (the intervening IUs have concerned actions of the drunk man and his interaction with the narrator), and the speaker can assume that it is no longer in the listeners' focus of consciousness either. The idea, having now faded from consciousness, is not 'given' at this point, but—activated twenty-three IUs previously—it is not 'new', either.
(3) Corpus excerpt from 'Aldi drunk' [Skat: 19-23 & 43-9].

19 [T]: #{ «und so 'ne Verkäuferin hat das gesëh'n /+ }
         and like a salesclerk has that seen
         #/ und hat /+ }
         and has
21         #{ «Q "hahaha Q" /+ }
22         #{[ was Sie hier aufmachen /+ ]}
         what you here up.open
         müsse S'e auch A› kàufen" /+ }
         must you also buy

43 [T]: #{ «A und die Verkäuferin /+ }
         and the salesclerk
         {{ die ihn vorher angemacht hat /+ }}
         who him before accosted has
         sitzt dèrweil A› neb'n ihm an 'er Kàsse "+ }
         sits next to him at a/the register
         ...(7) 'ne Kasse wèiter /+ }
         a register further
         #{ und springt sofort äuf /+ }
         and springs immediately up
         #{ und schrèit /+ }
         and shouts
         #{ "den Hàrzer Kàse /+
            the Harz cheese

19 [T]: and a salesclerk saw it,
20 and said,
21 "aha",
22 whatever you open here,
23 you have to buy too,

43 [T]: and the salesclerk,
44 who'd accosted him before,
45 is sitting next to him at the register;
46 a register away,
47 and jumps right up,
48 and shouts,
49 "the Harz cheese,
Instead, it has the third, intermediate activation status 'accessible'. This pattern is the most common source of cognitively semiactive ideas (Chafe 1993: 82). Other sources will be discussed below in §4.3.2 in conjunction with an examination of particularly difficult cases.

To sum up, there are three states of activation that mental representations ('ideas') of objects, events, and states can be in with respect to the speaker's conception of the listener's consciousness prior to uttering an IU: active ('given information'), semiactive ('accessible information'), or inactive ('new information'). The speaker's assumptions regarding the listener's consciousness affects the linguistic form of the expressions used, with given information generally taking an attenuated form and new information being unattenuated.

4.2.3 Identifiability

A second, independent component of Chafe's conception of information flow is the notion of 'identifiability'. Identifiability is a property of referents (i.e. mental representations of objects),\(^5\) which is independent of activation state and other discourse properties. In simple, intuitive terms, identifiability has to do with whether or not an intended referent can be picked out from other similar ones, or as (Chafe 1993: 88) succinctly puts it, '... an identifiable referent is one the speaker assumes the listener will be able to identify'. Thus—as with activation state—identifiability is a

---

\(^5\) More specifically, it a property of the relation between a referent and a reference; see Du Bois (1980: 218) for a discussion.
matter of the speaker estimating the listener’s knowledge, i.e. ‘[t]he speaker judges identifiability with respect to the hearer’ (Du Bois 1980: 217).

There is much confusion and obfuscation in the literature among the use of the terms ‘given’, ‘identifiable’, and ‘definite’. Givenness is sometimes used in the sense of ‘known’ (e.g. Clark & Haviland 1977), which is most closely related to the value ‘identifiable’ here. Identifiability and definiteness are often used interchangeably. In Chafe’s model, all three notions are differentiated. ‘Given’ and ‘identifiable’ are particular values associated with different discourse-pragmatic categories, while the definite/indefinite distinction is strictly used to refer to differences in the linguistic form of noun phrases (cf. Du Bois 1980). Formally definite nominal expressions in German, for instance, include noun phrases beginning with the definite article, demonstrative determiners, or possessive pronouns, and the personal pronouns. Of course, identifiability and definiteness are linked, in that in phrases for which the contrast is relevant, identifiable referents are expressed with definite nominals, and nonidentifiable referents are expressed with indefinite ones.

In order to analyze identifiability in discourse adequately, Chafe (1993) proposes decomposing the notion. He defines identifiability in terms of three components. A referent is identifiable when it is:

---

6 Cf. Hentschel & Weydt (1990: 205): ‘Die Identifizierbarkeit bezieht sich ... auf den von der Sprecherin angenommenen Kenntnisstand des Hörers’ (‘Identifiability refers to the speaker’s assumption of the state of the listener’s knowledge’).
(a) assumed already shared, directly or indirectly, by the listener
(b) expressed in a sufficiently identifying way
(c) contextually salient

The property of being shared has to do with the listener's knowledge. A referent that is directly shared is considered known to the listener, either from previous, common experience or from the current context of the talk, while the identity of an indirectly shared referent is assumed to be inferable. But being shared does not suffice. The shared referent must also be sufficiently identified and it must be contextually salient. That is, the expression used in mentioning the referent must delineate the category of the referent from other shared categories, and it must allow the referent to be distinguished from other referents in the same category. The salience of one referent as opposed to another may be due to textual, situational, social, or universal factors.

As before, let us examine some discourse from the primary corpus to see how Chafe's conception of identifiability operates. Consider IU 11 in (1) above. There are two referents expressed in this IU, 'a drunk man' and 'Aldi supermarket'. The first, a new character in the story, is neither directly nor indirectly shared by the listener, and so is nonidentifiable at this point. (Note that the noun phrase is indefinite as well.) On the other hand, the second referent is identifiable in IU 11. First, it had been mentioned three IUs earlier, and so is certainly shared knowledge at this point. Second, the use of the name Aldi is sufficiently identifying. As Du Bois (1980: 218) points out, proper nouns usually presuppose identifiability by themselves. Third, the larger context of the speech act (in this case,
knowledge of the habits of the narrator) serves to make a particular Aldi supermarket more salient than the others in town. The mention thus satisfies all three criteria for identifiability.

Now consider IU 12 in (1) above. Again there are two referents, 'the drunk man' and 'a cheese roll'. At this point, the first is identifiable. By virtue of its mention in the previous IU, 'the drunk man' is shared knowledge. The reference using the proform der is sufficiently identifying, as pronouns typically are (Chafe 1993: 93). In contrast, the second referent in the IU is nonidentifiable, as knowledge of it is unshared at this time. In IU 12 the idea of a 'cheese roll' is introduced as an object in the story.

In sum, there are two statuses of identifiability that referents (i.e. mental representations, or ideas, of objects) can have with respect to the speaker's conception of the listener's knowledge prior to uttering an IU: identifiable and nonidentifiable. Identifiable referents are those which the speaker can assume that the listener shares (directly or indirectly), which are expressed in a way that both establishes category membership and locates the referent within the category. Identifiable referents are generally expressed by means of definite nominals. In contrast, nonidentifiable referents are expressed using indefinite nominals.

4.2.4 Other discourse properties

There are numerous other associated semantic or pragmatic properties that have been attributed to discourse referents besides activation state
and identifiability that might be considered important in developing a concept of information flow. Logicians and semanticists alike have commonly distinguished 'singular' vs. 'general', 'opaque' vs. 'transparent', 'specific' vs. 'nonspecific', and 'generic' vs. 'particular' reference (Lyons 1977: 178-97), and discourse pragmaticists have paid due attention to these additional categories as well. Du Bois (1980), for example, draws attention to the roles that genericness and specificity play in the definiteness of referents in spontaneous speech. Chen (1986) also makes use of genericness and specificity in discussing referent introduction and tracking. Du Bois & Thompson (1989) include a broadened notion of specificity in their outline of nominal information flow categories. These categories, however, play little role in Chafe's (1993) conception of information flow. The notion of genericness does surface in the context of discussing the applicability of identifiability—for generic (and nonreferential) referents the identifiable/nonidentifiable distinction is irrelevant (98)—and specificity is mentioned in passing (100), but these categories remain in the background.

On the other hand, the notions of 'contrastiveness', 'starting points', 'referential importance', and 'discourse topic' play an important role in Chafe's information flow model, as all are independent of activation state and identifiability, and have considerable effects on the interpretation of linguistic structures in discourse. Although I do not systematically discuss any of them in the analysis of my corpus (for reasons to be discussed in §4.4), there are occasions when reference to one or more of these categories is necessary. For this reason, I sketch their content below.
Contrastiveness, which is primarily expressed by means of accent and word order, is concerned with the explicit highlighting of a choice made by the speaker. In essence, the speaker asserts that ‘... a certain focus item rather than other possible ones is correct ...’ (Chafe 1976: 35). Contrastiveness is especially necessary as an independent category to account fully for the distribution of accented phrases. It can be viewed as contributing an additional layer to the prosodic structure of an IU, which must be factored out before the contribution of activation state to that structure can be determined.

The notion of starting point constitutes a functional role concerning the 'point of departure' of a clause, that is, ‘... a referent from which it moves on to provide its own new contribution’ (Chafe 1993: 78). In languages such as English or German, it involves the syntactic structure of clauses directly, as starting points are generally expressed as grammatical subjects. The interaction between starting points and activation state gives rise to the postulation of certain constraints on linguistic expression, which I will discuss below in §4.6.

Referential importance concerns the relative conceptual contribution

---

7 Cf. Weil (1879: 20): 'Il y a donc un point de départ, une notion initiale, qui est également présente et à celui qui parle et à celui qui écoute, qui forme comme le lieu où les deux intelligences se rencontrent' ('Thus there is a point of departure, an initial notion, which is equally present both to the speaker and listener, which forms the place where the consciousness of both meet'); and Mathesius (1975 [1929]: 467): ‘... the point of departure of an utterance [is] that which is known or at least understood in a given situation and from which the speaker starts off ...'.
among referents in relation to the discourse as a whole, or put simply, '... their importance to the subject matter being verbalized' (Chafe 1993: 83). While referential importance may not be coded directly in the grammar (of German), there often seems to be a certain iconicity between the number of mentions a referent receives and its importance in the discourse. Again, the interaction of this category with other discourse properties plays a part in the formulation of constraints on expression.

Discourse topic describes what the talk is about. It is, unlike the discourse categories just discussed, not a property of referents, but of information in a larger, propositionally oriented sense (cf. Keenan & Schieffelin 1976: 338). For Chafe, discourse topic provides a label for the totality of ideas that are semiactive in a speaker's consciousness. It would seem, then, that the notion of discourse topic should be important to the determination of activation state, if not other discourse properties. In principle, this is true; however, as Chafe points out, ideas which are semiactive for the speaker are inactive for the listener until they are expressed. Since Chafe's information flow model seeks to represent the speaker's conception of the state of the listener's consciousness, discourse topic will play only a subsidiary role in the determination of activation state values.

4.3 Probing fundamental notions of information flow

In the previous section I outlined component parts of Chafe's model of information flow. In this section I examine some of the notions that underlie these parts, with the aim of clarifying the scope of their application. Fundamental in this regard is explicating the relationship between
the notions 'concept' and 'referent', and their relevance to discourse-pragmatic categories. I attend to these issues in the next subsection. In the subsection thereafter, I deal with other practical matters that arise when the model is applied to conversational discourse.

4.3.1 Concept and referent

Understanding what information flow is depends in no small part on understanding what exactly it applies to. In §4.2 I reported that activation state, identifiability, and other discourse-pragmatic categories are associated with what Chafe calls 'ideas', i.e. mental representations of objects (of all kinds), events (in a broad sense), and states (or properties). Somewhat more specifically, an idea is '... a cognitive unit that may be verbalized by a segment of language' (1992: 270-1). Thus, these ideas of objects, events, and states are generally expressed, respectively, by means of noun, verb, and adjective phrases, and consequently, it is through such expressions that ideas are apprehended.

Still, what kind of mental entities are ideas? What Chafe has in mind here is not entirely clear, as the picture he paints is unconnected in places.

---

8 This is the term used in Chafe (1976; 1992; 1993); cf. also (1977b: 227). In Chafe (1974; 1987) the term 'concept' is used instead. As will become clear in this section, I am reserving this latter term for another use. It may be difficult for the reader to keep in mind the technical sense of such an ordinary word as 'idea'. I wish to emphasize here that throughout this chapter idea always corresponds to Chafe's narrower definition.

9 Chafe uses the term 'verbalization' in all his writings to mean '... all those processes by which nonverbal knowledge is turned into language' (1977a: 41); cf. also 1977b. Note in particular that 'verbalization' in this use does not necessarily involve lexical verbs.
On the one hand, ideas are conceived of as representations of particular objects, events, and states (1976: 28; cf. 1993: 63-5). Concerning object ideas, Chafe (1976: 29) explains: 'It is the constant idea of the individual rather than the shifting words that may occupy one or several of the statuses listed above'. By this, he means that discourse-pragmatic properties are to be associated with a mental entity, not a linguistic expression (i.e. a mention), which may take a variety of different forms in referring to the same individual. On the other hand, the activation process seems not to be limited to ideas of individuals. Chafe (1987: 28) observes that '... when a particular instance of a category has been activated, all other instances of the category are simultaneously activated too'.\(^{10}\) Again, a mental entity is involved here, but one that is quite different from an idea of an individual. If this is the case, then presumably it is not only the referent as the idea of an individual entity that is activated in consciousness, but also the general 'concept'—whatever that turns out to be—that has been evoked with it. Although the activation of concepts is seemingly incongruous with the first statement cited, there need not be a problem here. The real problem is that a notion of concept has not really been integrated into the information flow model—nowhere else is mention even made of it.\(^{11}\) To address this shortcoming, I will in this subsection explore what concepts can do to enhance Chafe’s model and how they can play a central role in accounting

---

\(^{10}\) See also Chafe (1974: 125-27) and (1976: 32), which use constructed examples to illustrate the point. Chafe (1993: 75) does not discuss this process, but refers the reader to his 1974 article.

\(^{11}\) Chafe (1976: 42-3) approaches the problem in the above-mentioned context, but does not embrace the systematic solution I am proposing here.
for activation state values. I will suggest that by fully incorporating a notion of concept into the model, we are better able to appreciate the individual contributions made by concepts and ideas to information flow.

Concepts are undeniably important to human thought and discourse. In the introduction to their review of approaches to the structure of categories, Smith & Medin (1981: 1) open with some of the reasons:

Without concepts, mental life would be chaotic. If we perceived each entity as unique, we would be overwhelmed by the sheer diversity of what we experience and unable to remember more than a fraction of what we encounter. And if each individual entity needed a distinct name, our language would be staggeringly complex and communication would be virtually impossible. Fortunately, though, we do not perceive, remember, and talk about each object and event as unique, but rather as an instance of a class or concept that we already know something about.

In characterizing the categorization function of concepts (which is the focus of their work), Smith & Medin then state: 'To have a concept X is to know something about the properties of entities that belong to the class of X …’ (8). Thus, under this view, concepts can be conceived of as some kind of mental representations of classes of objects, events, or states, as opposed to the individual entities in those classes.12 Unfortunately, because linguists have been late to realize—and then, not without considerable

12 Cf. Medin & Smith (1984: 114): ‘... a mental representation of a simple class (i.e. a class denoted by a single word)’. See also the helpful discussion in Droste (1987).
controversy—the importance of concepts as a means of relating linguistic expressions with the way in which humans categorize the world about them, the nature of concepts has been studied primarily by psychologists and philosophers, mostly without concern for natural discourse. Still, recent research has uncovered a number of interesting characteristics of concepts that might be more closely scrutinized in a discourse context. Among the more interesting (but not necessarily compatible) proposals I count the difference between context-independent and context-dependent properties of concepts (Barsalou 1982), the emergence of thematic and functional (as opposed to object) concepts from representations of events (Fivush 1987), and the insufficiency of similarity-based approaches to concepts (Medin & Wattenmaker 1987).

While the developmental and structural aspects of concepts are, of course, of great importance, the immediate significance of such studies for the discourse-pragmatic issues under consideration here is that there must be two sides to the mental entities that we say linguistic mentions reflect: a generalized semantic frame that provides the necessary relational network for language use (i.e. the concept), and a locus for the organization of instantiated properties (i.e. the individualized idea); both of which may be activated in consciousness.

There is an obvious parallel to this concept/idea relationship that can be found in discussions of linguistic meaning, that of ‘sense’ vs. ‘designation’, which has constituted an integral part of philosophical conceptions of semantic reference at least since Frege (1892) proposed his ‘Sinn’ vs.
'Bedeutung' dichotomy. With the notions of both concept and sense, the focus is on characteristic properties of meaning, while the notions of both (individualized) idea and designation have to do with individuals associated with the concept or sense. In fact, the connection has been explicitly drawn. Rey (1983: 246), for example, equates sense and concept. But as I view matters, this is not quite correct. Concept, as it is used here, is paired with the DISCOURSE-PRAGMATIC notion of (individualized) idea; sense, on the other hand, is a SEMANTIC notion, and so cannot be directly equated. There is an essential connection, however. A concept can be construed as an abstract of the stereotypic knowledge linked to a linguistic form. Such an abstract would include sense relations and denotata, or at least prototypical ones. In this regard, it may also be possible to consider the interdependent semantic notions of sense and designation, taken together, as forming the basis of Barsalou's (1982) 'context-independent' side of concepts mentioned above.

Having sketched one view of 'concept' and its relation to Chafe's view of 'idea', I will now turn to discussing several specific reasons to include a notion of concept as an integrated component of information flow, especially of activation state. The two most important arguments in this regard have to do with the general scope of activation state and with the cases of divergence in the activation status of concepts and ideas.

---

13 Frege's opposition is usually translated 'sense' vs. 'reference'. As the latter is apt to lead to confusion in the context of Chafe's use of 'referent' for object idea, I have substituted 'designation' after Lyons (1977: 199).
The first point to make concerns the fact that activation state must apply to ideas of events and states as well as to referents. Under a simplified conception of activation—where only the idea of the individual is available to bear the activation state value—certain descriptive difficulties arise. Consider the following example:

(4) Corpus excerpt from ‘Ambassador’ [Uli’s visit: 342-50].

342 [A]:  #\{ ... und denn hätten sie /+
               and then had they
343            .. ’n .. Pakèt an den Köch geschickt /+
            a package to the cook sent
344            #\{ ...(9) die hätten uh% d%- --
            they had uh
345            #\{ .. ~es gibt ja auch von /+
            it gives yes also of
346 (T):       ... (N)
347            von uh= .. den M- Medikamènten /+
            of uh the medications
348            un’ sò was /+
            and thus something
349            ~gibt ’s ja [ überhaupt nichts ] /+
            gives it yes absolutely nothing
350 T:          #\{ [ ja= { es ist gànz schlimm } dort /+
               yes it is completely bad there
               and then they (sent),
343               a package to the cook (,)
344               they had uh –
345               there’s also,
346 (T):       (breathes)
347               of medications,
348               and such,
349               there’s absolutely nothing,
350 T:          yeah it’s really bad there,
The clause construction of interest here is formed in IUs 342-3. The predicate, *hatten geschickt* 'had sent', occurs with three arguments, the subject *sie* 'they', the accusative object *'n Paket* 'a package', and the oblique *an den Koch* 'to the cook'. While the referents conveyed by both the mentions *sie* and *den Koch* have been previously expressed, the idea of a package is introduced here for the first time, and so it would seem that this referent would bear the activation state value 'new' in IU 343. There is reason to believe, however, that the mention *'n Paket* was not intended to express an individual entity of its own, but rather was intended to form part of the predicate. Note that there is no mention of a package in the subsequent talk. It is not the package as an object that plays a role here; instead, it is the consequence of sending a package (i.e. receiving medicine) that is of significance. Assuming this to be the case, then it is the idea as a whole expressed by the collocation *hatten *'n Paket geschickt* 'sent a package' that is activated and thus 'new'.

14 Now consider the similar grammatical construction in example (5):

---

14 In considering this and similar examples, the familiar notion of incorporation may be relevant to understanding how an object + verb expression could work together to convey a single idea. The above collocation does not exhibit the grammatical features expected in various types of noun incorporation (see Mithun 1984). But surely there are degrees of incorporation—even German exhibits (orthographically, at least!) a range of N + V collocations/compounding types, where the noun (or noun phrase) is more or less morphosyntactically independent. Compare with the above case: *rudführen* 'ride a bike' (one orthographic word, no capitalization of noun, no article), *Bесheid sagen* 'inform' (two orthographic words, capitalized noun, no article), and *eine Rolle spielen* 'play a part' (two orthographic words, capitalized noun, only indefinite article possible). Collocations of the type in IU 343 may simply lie at one end of the continuum.
(5) Corpus excerpt from 'Ambassador' [Uli's visit: 279-81].

279 A: #\{ [ die hat mir ] mal 'n Fóto gežëigt /+ ]
\hspace{1cm}
\textit{she has to me once a photo shown}

280 (T): [ (N) ]

281 #\{ ... uh hätte ihr Männ âufgenomm'm \+ }\hspace{1cm}
\textit{uh had her man taken}

279 A: she showed me a photo,

280 (T): (breathes)

281 uh her husband took <it>;

The clause construction in IU 279 again contains three arguments, the subject \textit{die} 'she', the accusative object \textit{'n Foto} 'a photo', and the dative object \textit{mir} 'to me'. And again, the referents conveyed by both the mentions \textit{die} and \textit{mir} have been previously expressed, while the idea of a photo is new in IU 279. In this case, however, the subsequent talk suggests that the speaker intended the idea of the photo to have an independent conceptual status. Note that in the following IU she refers to it once more (via what may be called 'maximal attenuation of the prefield', i.e. a 'zero anaphor'), implying that there is an individuated, identifiable photo. If this is so, then the idea of the photo itself bears an activation state value.

The information flow contrast between the two examples is evident. But it is difficult to explain how a verb + object collocation can express one mental representation in (4), but two separate ones in (5), if we have only ideas of individual objects and events (and states) to ascribe activation state to. There is, after all, some notion of 'package' involved in the idea 'sent a package'. By adding a separate notion of concept, however, we can appeal to the difference between basic and complex concepts to de-
scribe the contrast, and the account becomes more plausible. It is well established that basic concepts combine to form complex ones (Smith & Medin 1981: 6). Given that lexical items express basic concepts—Rey’s (1983: 237) characterization of them is ‘... those conceptual contents associated with “open class” items such as nouns, verbs, and adjectives’—the verb + object collocations in (4) and (5) each express two basic concepts. The difference between (4) and (5), then, amounts to a difference in the complexity of the concepts the ideas are linked to. In (4) the basic concepts «PACKAGE» and «SEND» are combined into the complex concept «SEND A PACKAGE», and the idea expressed by hatten ’n Paket geschickt is linked to it.\(^{15}\) On the other hand, in (5) the concepts «PHOTO» and «SHOW» are not combined, thus ’n Foto and hat gezeigt express separate ideas.\(^{16}\) In support of Chafe, what matters to information flow distinctions is not the internal conceptual composition of an idea—the number of basic concepts linked to a single idea is not important—but the number of ideas in an utterance. Whether concepts are linked to separate ideas or are combined and linked to one depends on processing limitations, on assumptions about the state of the listener’s consciousness, and to some degree on the intents and purposes of the speaker—it is up to the speaker to express the

\(^{15}\) Note that in this chapter and the next, when discussing particular examples, I take care to distinguish between the actual mention in the exchange and the concept I hold to underlie the mention: mentions are given in *italics*, while concepts are rendered using the format «SMALL CAPITALS».

\(^{16}\) ‘Information’ is thus restricted to basic concepts and their combination in the model. Other informational aspects, such as certain concept enhancements (e.g. quantification, intensification) and relationships between concepts fall outside the domain of information flow (cf. Chafe 1993: 114).
content of the utterance in one way or another (Chafe 1976: 28). This is analytically, of course, a very slippery matter which requires a great deal of attention. I will in fact return to it in the next subsection, but what I have said here suffices to drawn attention to the problems that a notion of concept might mitigate.

The second argument illustrating the necessity of both concepts and ideas in an account of information flow concerns situations where one notion is activated separately. Activation of concept and idea normally happens in tandem. But there are sometimes points in the course of speaking where one or the other is already active (or semiactive), and not clearly distinguishing the two notions can lead to logical contradictions.

One such situation is where a concept happens to be associated with more than one entity at different points in the discourse, the situation cited at the beginning of this subsection as a clue to the necessity of concepts. In conjunction with his observation regarding the activation of a class of entities, Chafe (1987) gives the following example: Speaker A has (apparently) been talking about big undergraduate classes. Speaker B begins a story, and in the third IU uses the noun phrase a big undergraduate class that I had. For Chafe, the referent is given, due to the immediately preceding conversation: ‘Here the speaker assumed ... that everybody was already thinking about big undergraduate classes’ (28). Note—in support of this view—that the full noun phrase is weakly accented, except for the contrastive accent on I. But note in addition that the referent is also non-identifiable. Presumably, Speaker B has had many undergraduate classes
as a student, and he cannot assume that the listeners will know which one he means on first mention; rather, it is just this first mention that serves to establish the referent—to use an effective metaphor, to create a new mental 'file' in the listeners' consciousness (Du Bois 1980: 220). Yet if we follow these two threads to their logical conclusion, we are left with a curious state of affairs. Speaker B's referent is, as explained, given, i.e. assumed to be already in the minds of the interlocutors, but it is also nonidentifiable, i.e. assumed not to be identifiable to them.

In contrast, in an information flow account integrating concepts, there is no such contradiction. What is 'given' is a complex concept associated with the phrase a big undergraduate class that I had, not the referent. The speaker could only assume that «BIG UNDERGRADUATE CLASS» was still in the listeners' consciousness, at least peripherally. The idea of the particular undergraduate class clearly cannot be assumed to be in the listeners' consciousness until after the utterance, since this mention refers to an entity different from the one(s) expressed in the previous conversation. Thus, as the referent is 'new', there is no contradiction with it being simultaneously nonidentifiable.

It is important to realize that the communicative context which leads to this situation is not aberrant, but occurs anytime there has been talk of one thing, and subsequent talk of another instance of the same kind of thing. Indeed, there are several instances of this type of activation mismatch in my corpus. One example is given in (6):

306
(6) Corpus excerpt from 'Gas stove' [Morning visit: 2].

2 T: #\{ vielleicht will Théo ja den änder'n Gâsherdl hab'n (N) \* \}
maybe wants T yes the other gas.stove have

2 T: maybe Theo would want the other stove.

Prior to IU 2, Käthi has been on the phone talking to a third party, but within hearing distance. On the phone, she has just mentioned something about a gas stove, so the concept «GAS STOVE» is currently in the interlocutors' consciousness.\(^{17}\) Upon Käthi's return, Tamina makes reference to another gas stove in IU 2, which in this case she believes is identifiable to Käthi. Although the concept is given, the referent expressed by the mention den änder'n Gascherdl is new. Tamina cannot assume that Käthi is currently thinking about the other stove.

The converse situation—the referent expressed by a mention is given, but the concept associated with the mention is new—can also be found. Such activation mismatches involve full noun phrase expressions of the type illustrated in (7) below.

In IU 69 Tamina uses the expression von ihrer Beute, which refers to the drunk man, the main character of the story (see examples (1)-(3) above). Mention of him was last made in IU 62, so that the idea of the drunk, i.e. the referent, can be held to be at least semiactive. However, the concept «PRIZE», which is conveyed by the use of the word Beute in IU 69, is introduced into the discourse for the first time at this point. In other

\(^{17}\) Käthi's phone conversation is not part of my corpus (but it is on the tape recording), thus I cannot give the IU indicating this in (6).
words, the concept associated with this mention is previously inactive, i.e. it is new. Again, the activation states of the concept and the referent are not the same. Such a mention, of course, serves to enrich the story, adding a characterization that would not be conveyed by simply reexpressing the original concept «DRUNKARD» in repeating Besoffener, or by using a pronominal form. Again, situations like these cannot be considered in any way aberrant, and without a notion of concept, this aspect of the information flow in these IUs remains unaccounted for.

(7) Corpus excerpt from ‘Aldi drunk’ [Skat: 66-70].

66 [T]: ...(.7) und denn \+
and then

67 .. die Verkäuferin \+
the salesclerk

68 die hätt’ echt im Lèb’m \+\+
she would have really in the life

69 nich’ mehr so von ihrer Béute abgelassen \ wà /+
not more like from her booty let go ok

70 .. echt nich’ \-
really not

66 [T]: and then;
67 the salesclerk;
68 she wouldn’t for her life;
69 let go of her prize,
70 really.

There are additional situations where reference to separate notions of concept and idea seems necessary. Except in the case discussed above, when a new referent is introduced, both the idea of the individual and the concept or concepts underlying that idea are activated. The mechanics
of activation, as we have seen, operates on an IU-by-IU basis. The establishment of a referent in discourse, however, may extend beyond the scope of a single IU, and in fact, may involve contributions from more than one speech act participant. Example (8) illustrates one configuration:

(8) Corpus excerpt from ‘Getting by’ [Henrika’s visit: 20-24].

20 [H]: # { (H) .. «.9» un=d éin Ëhepaar+ /+
    and one couple

21 ... (1.0) von diesen Ḑhroposóphen-Leuten da= /+
    from these anthroposophy people there

22 { ...(7) (tsk) ... «1.0» die hàben ühm= /+
    they have uh

23 {{ ... die also àuch bei dem Ærz' waren /+ }}
    who so also by the doctor were

24 { die hab’um drèi Kinder /+ }
    they have three kids

In the clause construction in (8), the referent in question is expressed in a fairly involved way, as it is distributed over three IUs (not counting the false start in IU 22). The sufficiently identifying expression is: ein Ëhepaar von diesen Anthroposophen-Leuten da die also auch bei dem Arz’ waren ‘one couple from those anthroposophy people who also saw that doctor’.

---

18 Numerous researchers have discussed the ‘incremental’ (Redeker 1986) or ‘collaborative’ nature of referent introduction (Clark & Wilkes-Gibbs 1986; Geluykens 1988).
Presumably, the concepts involved are activated over the course of the current utterance. That is, the concept «MARRIED COUPLE» is activated by the end of IU 20, «ANTHROPOSOPHY PEOPLE» by the end of IU 21, and «DOCTOR» by the end of IU 23. In contrast, it seems that the referent is not fully established until the beginning of IU 24. Given the complication of the false start in IU 22 and the subsequent addition of a relative clause to further specify the referent in the next IU (Henrika has apparently reevaluated what is sufficiently identifying), it may be proper to hold that the noun phrase *ein Ehepaar von diesen Anthroposophen-Leuten da* expresses the idea of the individual. At the earliest, then, the referent is activated in the listener’s consciousness at the end of IU 21.

This variance in activation may also occur with referents that are assumed to be identifiable. Consider the case of the referent introduction in (9) below.

In IU 54 the referent ‘Berta’ is introduced into the conversation for the first time. The mention *Berta* apparently does not suffice to establish the referent, however, and Henrika adds the phrase *diese Freundin von mir* in the next IU to establish a connection for her listeners (which presumably allows them to distinguish this Berta from any others they know). She also takes the opportunity to amplify this with the clause *die Zucker hat*, constructed in IUs 57-8. It is not until IU 59 that the referent introduction is complete and the story continues. Again, the introduction of the concept («BERTA») in one IU contrasts with the establishment of the referent (‘Berta’) over more than one.
(9) Corpus excerpt from ‘Getting by’ [Henrika’s visit: 54-60].

54 [H]: #{ (H) .. «1.5» und Berta /+
and B
also diese Fréundin von mir /+
so this friend [FEM] of me
56 (I): ... mhm \+
57 {{ ... ~die dies’n --
who this
58 {{ .. die Zûcker hat /+ }}
who sugar has
59 { .. die war auch bei dè=m \ } ne /+ }
she was also by him ok
60 #{ .. ~die is’ auch zu dem =gekomm’n \+ }
she is also to him come

54 [H]: and Berta,
55 this friend of mine,
56 (I): mhm;
57 who this --
58 who has diabetes,
59 she visited him too,
60 she also came to see him;

Finally, consider a case where a new referent is clearly introduced, but
it is not at all obvious what concept, if any, is involved and what its acti-
vation status might be:

(10) Corpus excerpt from ‘Visa for India’ [Doppelkopf a: 37-8].

37 [D]: #{ (H) «.6» ~un’ dann èiner /+
and then one
38 { .. dèr hat sofort àufgele=gt /+ }
he has immediately hung up

37 [D]: and then one ‘guy’,
38 he hang up right away,
In IU 37 Diana has clearly introduced a new referent—the introduction is effected by means of the mention *einer* 'one' in an IU prior to its deployment in the story. But the referent, who plays no further role in the discourse, is expressed in a semantically minimal way in both IU 37 and 38. (Apparently, Diana only needs the referent in order to attribute to it one action—the hanging up—in a series of events she is describing; see example (20) in Chapter 3 for more context.) It may well be the case, then, that the only concept being expressed here is the ontologically basic one of «(MALE) PERSON» (cf. Keil 1979), which may always be semiactive in consciousness. Once more, the activation of concept and idea diverge.

In sum, separate notions of 'concept' and 'idea' are needed to account for information flow in discourse. A particular mention may be used to focus on the concept or the idea, or to express both. The activation of linked concepts and ideas in consciousness may occur at different points in time. Both notions should be equally incorporated into an enhanced model of information flow based on Chafe's work.

4.3.2 The determination of activation state

However clear Chafe's model of activation state is in principle, some uncertainties arise in its application, even after clarification of the distinction between concept and idea. In this subsection I will raise numerous questions regarding the determination of activation state, questions that at this time unfortunately do not have solid answers. It is hoped that by raising them in the first place, specific points of inquiry will have been outlined for later research. Still, a closer examination aids in a better un-
derstanding of the linguistic significance of information flow in the conversational discourse that constitutes the present corpus.

A brief point of clarification addressing the issue of circularity should preface this examination. The analysis of discourse-pragmatic properties is particularly susceptible to the charge of circular reasoning. It is interpretive work, with few certain formal configurations. The main linguistic correlates of activation state, for example—accent and pronominalization—are multiply determined, at best. Yet there is reason to insist on the interpretive methodology. As argued in Chapter 2 for the auditory determination of IUs from an unfiltered signal, when the analyst is able to use all available resources, as the listener is privileged to, then a plausible and consistent interpretation is possible. Language use is not mechanistic, and the blind application of any algorithm, though ‘objective’, can hardly lead to a valid analysis.

Each concept/idea expressed in discourse can be thought of as having its own ‘activation history’ (Chafe 1992: 270). That is, at different points in time each concept/idea will be in one of three activation states—active, semiactive, or inactive—and the sum of the particular states for any concept/idea provide a history of that concept/idea in consciousness. Given that there are three activation state values, there are then six possible activation state transitions, as well as three steady-state (i.e. unchanging) combinations, as listed in (11):
(11)  
a) semiactive  ⇒  active  
b) inactive  ⇒  active  
c) active  ⇒  semiactive  
d) inactive  ⇒  semiactive  
e) [active  ⇒  inactive]  
f) semiactive  ⇒  inactive  
g) active  =  active  
h) semiactive  =  semiactive  
i) inactive  =  inactive  

transitional  

steady-state

Some of the configurations in (11) are more interesting than others with respect to their linguistic ramifications, and some are more problematic from an analytic viewpoint. Let us survey the list briefly before delving into the details. Transitions (a) and (b) constitute activation proper, and transition (c) is part of the process of ‘deactivation’ (Chafe 1987: 28). Transitions (d) and (f) describe peripheral activation and deactivation, that is, concepts/ideas going in and out of a semiactive state. Transition (e), i.e. the direct transition from an active to an inactive state, is hypothesized not to occur (Chafe 1987: 29). When a concept/idea is no longer in consciousness, it first exists in a semiactive state before it becomes fully inactive. That is, deactivation follows the path (c), then (f). The steady states (g), (h), and (i) describe concepts/ideas that retain their activation status across IUs.

The configurations that have the most direct linguistic relevance involve the status ‘active’, i.e. transitions (a), (b), and (c), and steady state (g), listed in boldface type in (11). My focus in the following discussion will thus be on these four. As indicated, transition (e) does not have to be con-
sidered for theoretical reasons; the remaining four configurations will receive only minor consideration here.

I will begin with the steady state (g), as it is usually the easiest to interpret, and especially in narrative or narrative-like discourse, it covers a substantial number of mentions. When a concept/idea is repeatedly expressed in a sequence of IU's, it remains active in consciousness over that stretch of talk. The most salient linguistic features of mentions conveying continuing given information are said to be pronominalization and lack of pitch accent. Consider the longer excerpt in (12):

(12) Corpus excerpt from ‘Aldi drunk’ [Skat: 25-42].

25 [T]: #\{ (H) .. <<6> un' dann hat er die in 'n Wàg'n gelegt /+\ }  
and then has he in the cart laid

26 ... und noch so 'ne Büchse Bier dazu \+\ }  
and still like a can beer there.to

27 für .. fùm'ndréßig Pfennig \·\ }  
for thirty-five pennies

28 #\{ (H) «.3» und hat die .. Ràlle= .. Kàse so= /+◊\ }  
and has the roll cheese like

29 .. bèi=m /+◊\ }  
by.the

30 .. Nachvòrnegeh'n aufgegess'n \+\ }  
forward-going up.eaten

25 [T]: and then he put it «the cheese roll» in the cart,
26 and also a can of beer;
27 for thirty-five cents.
28 and (ate up) the cheese roll,
29 on the,
30 way to the checkstand (;)
and he was really blasted;
I stood,
in front of him in line;
and he kept on,
stumbling into me;
well,
(K): mhm;
anyway he ate it,
and the last bit of cheese,
<he> put in his pocket,
ok;
then he paid for his beer,
Example (12) comprises the intervening IUs that were left out of example (3). The protagonist of the story is the drunk man, who is referred to throughout the above passage. Mentions of him function as the grammatical subject in five of the eight clause constructions in (12)—IUs 25, 31, 34, 38, and 42. In two clause constructions he is the implied subject (IUs 28 and 39). And in the one remaining clause, he appears in a mention functioning as the object of a preposition (IU 33). The concept/idea of the drunk is active from IU to IU. All overt mentions are made with pronominal forms, as expected. Two of the forms (in IUs 31 and 34) bear a pitch accent, which is unexpected from information flow considerations. It can be noted, however, that there may be a straightforward rhythmic explanation. In this passage, each IU contains an initial pitch accent in the first three syllables. Just in those instances where the pronominal mentions referring to the drunk occur in this initial rhythmic field are they accented. The three unaccented mentions each occur later in the IU. The general applicability of this explanation for accented active mentions remains to be investigated, but the example does raise the question of the distribution of accent and its role in determining activation state.

Before considering this point further, let us look at the other case of a continuing active concept/idea in (12). The item of contention in the story is a small cheese roll that the drunk eats while shopping. In (12) it is referred to in IUs 25, 28, and 38. Although there are two intervening IUs, it is probable that the mention in IU 28 exemplifies the steady-state configuration (g). The clause constructions containing the mentions in IUs 25 and 28 are adjacent, and IUs 26-7 are short additions to the on-going
clause construction, each containing only one (complex) concept/idea. The active mention in IU 28 referring to the cheese roll is not attenuated, however. Neither is it pronominalized nor unaccented. Again, this is unexpected from typical activation state patterns, but other factors appear to be involved. Note that there is a potentially confounding referent in IU 26, expressed as 'ne Büchse Bier 'a can of beer', a referent which does receive a mention later in the story (IU 42). The details of this kind of referential interference are beyond the scope of this subsection; see Givón (1983a) and Fox (1987: 48-62) for discussions. However, the example serves to point out the possible contribution of factors unrelated to activation state in explaining pronominalization, or the lack of it.

As indicated, the mention of the cheese roll in IU 28 is also accented. This fact points back to the issue of accentuation. Let us now examine pitch accent on noun phrases in the above excerpt and in the primary corpus. It will be noticed that all of the full noun phrases in (12) bear one or more pitch accents. (In fact, save Bifşchen 'bit' in IU 39, all lexical nouns are accented.) This pattern is by far the prevalent one in the primary corpus, regardless of activation status. Out of 348 noun phrases headed by a lexical noun, only 13 (4%) bear no pitch accent. Moreover, only 4 of these 13 clearly express the idea of an individualized object (as opposed to a basic concept in a larger complex). In other words, when a concept/idea is expressed using a full noun phrase, the mention will in all likelihood be accented, even if it conveys given information. There may again be rhythmic reasons for many of these pitch accents—notice that all but two of the accented noun phrases in (12) come late in the IU—nonetheless,
pitch accent (or the lack of it) does not appear to be a very helpful feature for identifying active concepts/ideas.

I will now turn to a discussion of activation configuration (c), the deactivation of active information. In considering the instances of repeated reference to an idea in discourse—which keeps a concept/idea active—the question naturally arises as to when a concept/idea has faded from consciousness and so has left an active state and entered a semiactive one. How many intervening IUs (or clauses or concepts) must be counted between mentions for the analyst to conclude that the second mention is no longer given, but merely accessible? There can be, of course, no absolute answer. Activation state is dependent at least in part on speaker choice. As Chafe (1976: 32) states, ‘... it is fundamentally a matter of the speaker’s belief that the item is in the addressee’s consciousness ...’, and the speaker may choose to believe one way or the other. Highly constraining cognitive (processing) factors may predominate most of the time, but it is reasonable to assume that some interspeaker variation will occur, that varying pressures on the working memory of a speaker will impact on the way a focus of consciousness is expressed, and that speakers may exploit opportunities to package the contents of their utterances for rhetorical effects. Speakers may also make faulty evaluations about the state of the listener's consciousness. And not all lapses in communication can be expected to produce overt verbal responses.

Although activation transition points are not exactly discernible, it may be possible to set up a reasonable MAXIMUM limit, whereby one can be
fairly confident that a concept/idea has receded from consciousness. An
operationalization of this sort is encountered in Du Bois (1987: 816). Du
Bois suggests that a concept/idea can be assumed to have changed from a
given to an accessible state after twenty IUs have been uttered without a
mention of the concept/idea. For statistical calculations, such limits
(even though more or less arbitrary) can help raise the reliability of activa-
tion state measurements. However, for conversational discourse, as
opposed to elicited narrative, the limits can be expected to differ. Thus,
while there can be little question that idea of the salesclerk has faded from
consciousness by IU 43 in example (3)—the last prior (implicit) mention
occurred at IU 20—most cases of deactivation are not as obvious. The idea
of the beer expressed at IU 42 in example (12)—last prior mention fifteen
IUs earlier—is perhaps still comfortably semiactive at the beginning of
that utterance. But the comfort level decreases rapidly for referents such
as the salesclerk at IU 67 in example (7) and the cheese roll at IU 38 in (12).
The last prior mention of the salesclerk occurred eleven IUs earlier (in an
implicit reference at IU 56); the last prior mention of the cheese roll oc-
curred at IU 28, i.e. nine IUs earlier. Are these referents still active at these
points? Or have they receded into peripheral consciousness? The idea of
the cheese roll is expressed using an accented proform, while that of the

19 Cf. Givón (1983b), who uses a limit of twenty in another way. Incidentally, Chafe does
not state anywhere what he considers to be a reasonable limit, but it is his view that ideas
become semiactive only after ‘a few IUs’ (p.c).

20 For simplicity’s sake, I give straight counts of all intervening IUs by the current
speaker. A serious attempt at incorporating counts of IUs would have to differentiate IU
type. For a basic typology, see Chafe (1993: 61-2).
salesclerk finds expression in an accented full noun phrase—and in a syntactic position outside the inner bracketing of the clause. Obviously, we would like to claim here that the cheese roll referent is still in consciousness, but that the salesclerk referent is merely semiactive. The difference in the counts alone do not permit that conclusion.

The relevant conjecture to make with regard to the above-described uncertainty is that there is quite possibly a modest range during which speakers have a good deal of latitude in the assumptions they make regarding the state of their listeners' consciousness. To take somewhat motivated numbers, I would say that for mentions separated by between three and fifteen IUs, distance is only suggestive, and that other parameters must be given more weight.

I turn now to activation proper, configuration (a), the activation of semiactive information. In order to understand the transition from a semiactive to an active state, it is necessary to discuss the various sources for accessible concepts/ideas. One source has already been examined: deactivation. Chafe (1993: 82) gives two other sources: the situational (i.e. non-linguistic) context (what I will call 'contextual availability'), and direct association from another concept/idea mentioned in the discourse ('co-textual availability'). These latter two sources can be described as results of transition (d), inactive information becoming semiactive at some point before it is fully activated. Transition (a) can thus be seen to embody two quite different processes.

The issue of deactivation—transition configuration (c)—was just dis-
cussed. As an input for transition (a), the consequences of transition (f), i.e. semiactive concepts/ideas receding to an inactive state, also need to be briefly considered. The main point to make in this regard is that concepts/ideas can remain semiactive for quite some time. Example (13) contains an instance of a referent persisting in a semiactive state for about fifty IUs:

(13) Corpus excerpt from ‘Ambassador’ [Uli’s visit: 56-9 & 116].

56 [A]: #\{ .. ~da woll'n s'e dann hin /+ \}
      there want they then thither

57 #\{ und komm'n diese Woche wieder z- .. nächst- --
     and come this week again b- next-

58 .. [ kommende Woche wieder ] zurück \* \}
     coming week again back

59 (U): [ «P uh ja= P» ] \*
      uh yes

116 A: #\{ (0) sie waren zuèrst in Montréux \+ \}
      they were first in M

56 [A]: then they want to go there,
57 and come (back) this week next- --
58 the coming week ().  
59 (U): uh yeah.

116 A: first they were in Montreux;

In IU 56 there is direct reference to the protagonists of the current story with the attenuated mention s’è ‘they’. An implicit reference then follows in the next IU (as the implied subject of the verb kommen ‘come’). The next reference—direct or implicit—comes in IU 116, where mention is once again made with the personal pronoun sie (this time unattenuated,
but also unaccented). In the intervening discourse, Uli has uttered forty-one IUs, Anna eight, and there have been six speaker changes. In trying to determine the cases of reactivation in discourse, one must be aware that transition (f) may not occur for long stretches of talk. Chafe (1987: 42-3; 1993: 115) proposes that major shifts of the ideas in peripheral consciousness occur at paragraph and discourse topic boundaries—presumably this would include both transitions (d) and (f)—but example (13) suggests that additional factors are involved.

Both kinds of availability are also somewhat problematic, but for different reasons. When a concept/idea is somehow present in the immediate environment, it can be difficult to distinguish given and accessible referents. Consider example (14):

(14) Corpus excerpt from ‘Weavers’ fair’ [Doppelkopf b: 128-30].

128 [M]: #1 dann hat er so getå=n \+ }
then has he thus done

129  #1 .. so gezèigt mit ’n Hånd’n \+ } «gestures»
thus pointed with the hands

130   #1 also hier sind die Wånde \ ne /+ }  
so here are the walls ok

128 [M]: then he went like this;
129  outlining with his hands;
130  here are the walls,

The referent in question is expressed by die Wånde ‘the walls’ at IU 130 and visually by a corresponding gesture.21 Assuming that the speaker

21 The notation «gestures» at the end of IU 129 does not signify that the gesture occurred precisely at this point, but only that there was a relevant gesture associated with the dis-
could believe the listeners perceived the gesture iconically, it is plausible to assume also that she could hold the idea of the walls to be in their consciousness, at least peripherally. Yet which case is it, given or accessible? The fact that the mention takes the form of an accented lexical noun phrase suggests the latter possibility, but there is still room for doubt.

Determining the activation state of direct references to the speaker and listener constitutes another related issue. It is usually assumed that the speech act participants as referents always constitute given information. Chafe (1993: 82-3) argues, however, that a first person mention conveys accessible information when the mention bears a noncontrastive accent. An instance of this kind occurs at IU 32 in example (12), where the pronoun ich 'I' is accented. The plausibility of the idea of the speaker being semiactive just prior to this mention is high. The last explicit mention the speaker made to herself occurred at IU 10 using the pronoun wir 'we' (see example (1) for the co-text). Still, taking the rhythmic factors discussed previously into consideration, some doubt must remain whether the accent can be attributed to the information flow.

In contrast, co-textual availability is problematic because it is sometimes difficult to distinguish a mention based on direct association from one based on an inference. In the former case, the referent is semiactive at the time of mention; in the latter, inactive. (I discuss inferred referents below.) Co-textual availability seems to require an explicit linguistic connec-

course. As a speech event participant, I recall that the gesture indicating the walls began prior to the mention.
tion between the referents. Chafe (1993: 82) cites the idea of your back being semiactive through direct association with a previous mention of backaches, apparently by virtue of the commonality of the form back, which of course expresses the concept «BACK». As a clear source of semiactive referents, co-textual availability is rare in my German conversational discourse. One good example occurs in (15):

(15) Corpus excerpt from 'Vacation travel' [Uli's visit: 457-65].

457 A:  #  [ ich .. bin ãbends ] hier /+
            I am in the evening here
458     .. ~wêg } mi 'm Schlàfwagenzug /+ }
            away with the sleeping car train
459     #{ .. hâttten w'r unser Schlàfwagenabteil /+ }
            had we our sleeping car compartment
460     #{ .. schlàfen kann ich nich' in dem Æbtei=l /+ }
            sleep can I not in the compartment
461     #{ .. [ weil ] ich .. weil ich --
            because I because I
462 (U):  [ mhm= ] /+
463     #{ .. man liegt ja quãrz zur Fãhrrichtung \ } nich' /+ }
            one lies yes crosswise to the travel direction not
464 U:  #{ jà= /+ }
            yes
465 A:  #{ .. und .. an jeder Kûrve hebt sich der Wàg'n /+ }
            and at every curve lifts self the car

457 A:  I left here in the evening,
458     with the sleeper,
459     we had our sleeping compartment,
460     I can never sleep in the compartment,
461     because I because I–
462 (U):  mhm,
463     you lie perpendicular to the travel direction, right,
464 U:  yes,
465 A:  and at every curve the car swings,
The referent in question is expressed as *der Wag'n* 'the car' in IU 465. Earlier mentions had referred to the whole train, *'m Schlafwagenzug* in IU 458, and to one of the compartments in the car, *unser Schlafwagenabteil* in IU 459, but not to the car itself. The form *Wag(e)n* expressing the concept «CAR» is, however, contained in both of the earlier mentions, so that conditions for assuming semiactive status of the car at IU 465 seem to obtain.

Of course, borderline cases of co-textual availability also exist, one of which may be found in (16):

(16) Corpus excerpt from ‘Washing machine’ [Morning visit: 62-3 & 144].

62  (T):  #{ .. [bloß das is' ja auch> blöde ] \+
       only that is yes also stupid

63  K:  #{ [ ~<was is' denn> mit der Wäsche машине \*
       what is then with the washing machine

   ↓

144 [T]:  #{ ich schmeiß' %uh die Wäsche rein /+
       I throw uh the wash in

62  (T):  [but that's also stupid;
63  K:  [how's the washing machine.

   ↓

144 [T]:  I throw the wash in,

At IU 63 Käthi asks about Tamina's washing machine, and two talk about it for a few IUs, before the topic shifts at IU 73. At IU 144 Tamina resumes talking about doing the wash, which she expresses as *die Wäsche* 'the wash'. This mention may well convey accessible information, but it is unclear whether or not the linguistic association of the noun *Wäsche* with the noun *Waschmaschine* 'washing machine' is direct enough to have ef-
fected the conjectured activation transition.

A perhaps more common problem involves reference to an individual which has been derived from a prior reference to a group. While the association is not as direct in this case, it does appear that semiaactive concepts/ideas are involved. Consider, for instance, example (17):

(17) Corpus excerpt from ‘Ambassador’ [Uli’s visit: 126-8].

126 A:    # { (0) jà = /+
            yes
127       { .. d% - di- die kômmen <denn> nächste Woche zurück /+ } 
          th - the - they come then next week back
128       # { <A denn èr muß wieder in den Dienst A> /+ } 
          for he must again in the service

126 A:    yeah,
127       th - the - they are coming back next week,
128       because he has to go back into the service,

Up to IU 128 Anna has referred to the daughter of her friend and the daughter’s husband using plural proforms, as in IU 127 (die ‘they’). In IU 128, the husband is referred to as an individual for the first time, and the form of the mention is an accented personal pronoun (er). Speakers can readily switch from plural to singular reference (and back again), but the implications for information flow are not always clear.

All in all, the significance of the problems arising from co-textual availability are, however, relatively minor, since the frequency of this source of semiaactive concepts/ideas appears to be quite low. The other sources of accessible information present more formidable challenges. The major difficulty in this regard is distinguishing active from semiac-
tive concepts/ideas, i.e. activation configurations (a) from (g).

Finally, I will discuss activation configuration (b), the activation of inactive information. This transitional configuration constitutes the second part of activation proper. In theoretical terms, it should differ little from transition (a) just discussed. This can be ascribed to the observation that expressions of semiactive concepts/ideas pattern linguistically more like new information than given (Chafe 1993: 81). That is, such mentions are more likely to use lexical noun phrases or bear an accent, or be provided with an enhanced degree of specification, such as a relative clause. In practice, however, there are significant differences between transitions (a) and (b), and when all the classes of cases that fall under (b) are considered, it becomes evident that this transition is the least problematic.

The first class of cases I will refer to as ‘textbook referent introduction’, as these cases include those instances of transitional configuration (b) that are typically thought of as representative. It involves the introduction of a concept/idea into the listener’s consciousness that is new for the discourse and that is assumed by the speaker not to be identifiable. This class is thus similar to Prince’s (1981) ‘brand-new’ and Chafe’s (1993: 91) ‘unshared’ referents. An example of this type of activation is illustrated in (12) above, where the concept/idea ‘ne Büchse Bier ‘a can of beer’ is introduced in IU 26.

Another class of cases constitutes a counterpart of the first. It involves the introduction of a concept/idea into the listener’s consciousness that again is new for the discourse, but that is instead assumed by the speaker
to be identifiable to the listener. This is Chafe’s (1993: 91) ‘shared’ referent. Typical instances involve the introduction of referents using proper noun mentions, where identifiability can usually be assumed (Du Bois 1980).

A third class of cases involves what I call ‘framing mentions’, that is, referent introductions that depend on stereotypic ‘frames’. Framing mentions are used to anchor referents to the discourse context, much as explicit elaborations (e.g. with relative clauses) do. As with the case of new, but identifiable mentions just discussed, the speaker is appealing to the listener’s knowledge to make the connection. In the preceding case, it was (group-specific) mutual knowledge (i.e. that shared by a limited number of people) that had to be activated, while with framing mentions, it is knowledge that virtually any native speaker can be expected to possess. It is important to emphasize that concepts/ideas derived from frames associated with another concept/idea that has previously been (or is currently) active in general constitute new—not accessible—information. As was explained in §4.2, the speaker usually cannot assume the listener knows exactly what concepts/ideas are semiactive for the speaker, and s/he must present them as inactive. Example (12) above contains several mentions that can be linked to frames—or as I prefer to view the process—it contains several expressions which are used to frame a concept/idea. The first such mention is ‘n Wag’n ‘a cart’ in IU 25, a framing mention because shopping carts can be expected to be found in a su-

22 See Tannen (1979) for a comparative review of the numerous related terms. See also Tannen (1993) for examples of recent linguistic research using this notion.
permarket. The speaker presents the concept expressed in the nominalization *beim Nachvornegeh'n*, literally 'on going forward', as well as that in the mention *in 'er Reihe* 'in line' as part of the same frame.

I have now examined the most important activation configurations outlined in (11), using examples from the primary corpus to survey problems encountered in the application of the information flow model. As a kind of interim summary, the following rules of thumb—in no way to be read as absolute statements—which have become evident from my survey of the German conversational examples can be taken as guidelines for determining activation state in similar discourse data.

(18) General guidelines for determining activation state:
   a) The second of two clause-adjacent mentions is given.
   b) Pronominal mentions are non-new.
   c) Accented pronominal mentions may be given or accessible.
   d) Accented full NP mentions may be given, accessible, or new.
   e) Unaccented full NP mentions are non-new.

A final point relating to the determination of activation state should be made. Much past confusion in the determination of activation state (or other similar categories incorporating a given/new distinction) seems to hinge on the interaction of activation state with identifiability. In principle, the two properties are independent. In practice, there are strong tendencies for certain patterns, although it must be stressed that under my conception of the two discourse properties, all of the activation state/identifiability configurations are possible and do occur. The patterns that I have observed in my corpus of German conversation and have illus-
trated in this section are summarized in Table 4.1 below.

<table>
<thead>
<tr>
<th>Mention</th>
<th>Activation state</th>
<th>Identifiability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>New</td>
<td>-</td>
</tr>
<tr>
<td>2nd</td>
<td>Given</td>
<td>+</td>
</tr>
<tr>
<td>Later</td>
<td>Accessible</td>
<td>+</td>
</tr>
<tr>
<td>1st</td>
<td>New</td>
<td>+</td>
</tr>
<tr>
<td>1st</td>
<td>New</td>
<td>+</td>
</tr>
<tr>
<td>1st</td>
<td>Given/Acc.</td>
<td>±</td>
</tr>
<tr>
<td>1st</td>
<td>Given/Acc.</td>
<td>±</td>
</tr>
<tr>
<td>1st</td>
<td>Given/Acc.</td>
<td>±</td>
</tr>
</tbody>
</table>

Table 4.1. Interaction of activation state and identifiability in tracking referents.

Interestingly, Chafe states regarding the interaction of activation state and identifiability (1993: 91): 'Given and accessible ideas are necessarily shared, since by definition the speaker believes them to be already active or semiactive in the listener’s mind'. Chafe appears to be influenced by Prince here, as he attributes to her (Prince 1981: 91) the observation that only new ideas can be unshared. But Prince has a different conception of information flow. While they do seem to agree on the major component of identifiability—cf. 'hearer-old' vs. 'hearer-new' in Prince (1992: 301-3) with 'shared' vs. 'unshared' in Chafe (1993: 89-91)—givenness for Prince (1992: 303-4) is a property of entities in a discourse model, not of ideas in the consciousness of speakers. Statements made in one model do not necessarily transfer to the other. Thus, for Prince it is always true (and not contradictory) to say that when a referent is 'discourse-old' (given), it is

331
necessarily ‘hearer-old’ (identifiable), because ‘... hearers are expected to remember what they have been told’ (1992: 303); but for Chafe this implication is valid only when the referent is unique, or at least homophoric.\textsuperscript{23} It is hard to see how this implication can be maintained when the notions of concept and referent are collapsed.

In outlining Chafe’s model of information flow in §4.2, it was reported that the properties of concept activation state and identifiability are said to directly affect the linguistic form of the expressions used, the key statement being that given information generally effects an attenuated form, while new information is unattenuated. In this section, I have pointed out numerous exceptions to this generalization. Despite the difficulties that I have outlined, I have gone forward with an interpretation of activation state in the primary corpus. It happens that several rules of thumb can be written for eliminating various activation configurations for certain mentions. Some configurations—notably those involving semiaactive concepts/ideas—are particularly troublesome. Others, notably that involving the activation of new information, are more amenable to interpretation. By focusing on the patterns associated with the clearer configurations, a substantial amount of the uncertainty surrounding the determination of activation state can be avoided. In the following two sections, I will thus concentrate on the activation of inactive referents.

\textsuperscript{23} Lambrecht (1988: 144–6), who adopts activation state together with the lack of a referent/concept distinction from Chafe, also explicitly states that ‘unidentifiable’ implies ‘inactive’.
4.4 The distribution of new mentions and referential tracking

In previous sections, I have provided a glimpse of the discourse-pragmatic characteristics of German discourse through the presentation of several examples from the primary corpus. In this section, I will present a brief, though systematic, overview of these characteristics, with a focus on new tracking mentions.

Out of more than 2000 nominals in the primary corpus, only 332 (about 17%) can be classified as new mentions, and a much smaller number, 170 (about 8%), were judged to be accessible mentions. That is, the vast majority of nominals in my sample of German discourse express given information. Perhaps one consequence of this pattern for German—as for English—grammar is the lack of functionality of marking activation state directly. Prince (1992) notes that except for the inferences that can be drawn in the use of pronominals, English grammar does not seem to encode activation state. A similar statement could be made for German (although possible word order correlations with information flow need closer investigation). If the vast majority of mentions are always given, then there is little need for elaborate marking of this fact.

This pattern of low frequency for new mentions carries over, of course, to intonation units and to clause constructions. Only 329 IUs out of 1736 (19%) contain a new nominal mention. If it is true that prosodic phrases normally conveys one unit of new information (Halliday 1967b; cf. Givón 1975), then the majority of IUs apparently express new information through other means. Similarly, out of 973 clause constructions, only
286 (29%) contain a new nominal mention.

For the remainder of this section, I will focus specifically on the characteristics of new mentions which serve to introduce individuated referents that play an intuitively significant role in the discourse, and so will typically receive additional subsequent mentions. The expression of these referents I term 'tracking mentions’, following Du Bois & Thompson’s (1989) distinction between ‘tracked’ and ‘nontracked’ referents. Tracking mentions may be contrasted with those that serve either a predicing or an orienting discourse referential function. (Predicating mentions include predicate nominals and grammatical objects that do not express individuated referents. Orienting mentions provide temporal, locative, causal, or possessive specifications.) Of interest is the fact that the introduction of new tracking referents clearly does not constitute the primary function of new nominal mentions in discourse. As Table 4.2 shows, most new mentions (71%) are nontracking. The majority serve nonreferential predicating or orienting functions.

<table>
<thead>
<tr>
<th>Tracking</th>
<th>Predicating</th>
<th>Orienting</th>
</tr>
</thead>
<tbody>
<tr>
<td>95 (29%)</td>
<td>112 (34%)</td>
<td>125 (38%)</td>
</tr>
</tbody>
</table>

Table 4.2. Referential function for new nominal mentions in corpus (N=332).

I will report the identifiability status, and the syntactic and prosodic integration of new tracking mentions. With respect to identifiability, the referents of over half (52/95) the new tracking mentions are identifiable. Again, this pattern is counter to the expectations that the textbook case of
referent introduction evokes. (See Table 4.1.) However, all instances of identifiable new mentions are accounted for by an appeal to the notions of assumed specific prior knowledge ('shared referent') and assumed general prior knowledge ('framed referent') as discussed in §4.3.2. This pattern implies that a majority of new referents are expressed in a way that directly anchors them in the listener's existing framework of the discourse. Very few new referents require the listener to create a new mental representation.

With respect to their integration in clause constructions, the vast majority of new tracking mentions are fully integrated. Three fourths (72/95) occur in the inner bracketing of their clause construction; only a quarter occur in peripheral or syntactically isolated positions. This distribution is consonant with the fact that two thirds (64/95) of the mentions are expressed as grammatical arguments. Note as well that two thirds (65/95) of the new tracking mentions occur in the midfield of the clause or later. Specifically, the prefield in the German clause is dispreferred as a locus for new referent introductions, although instances do occur (cf. Dutra 1987 for Brazilian Portuguese).

Given that new tracking mentions are so well integrated into the syntactic structure, it is somewhat surprising to find that two thirds (59/89) of the clause constructions containing these new mentions are constructed over more than one IU. Recall that the majority (54%) of clauses are constructed within the scope of a single IU. However, in half (30/59) of the cases where the clause construction extends over more than one IU, the
noun phrase expressing the new tracking mention occurs in a separate IU. These few cases constitute instances of so-called 'fancy syntax' (Prince 1985).

4.5 Evaluating hypotheses about information flow

In the preceding section I have shown that mentions expressing referents with the activation state value 'new' are patterned in certain ways with respect to prosodic and syntactic structures. Similar patternings have been reported for other languages. Thus, it is natural that general explanations for these regularities have been sought. Several researchers have recently proposed specific constraints on the distribution of new information as a partial account for recurrent linguistic patterns. Du Bois (1985; 1987) and Chafe (1987; 1993) are especially notable in this regard. Earlier works by Halliday (1967b), Givón (1975), and Pawley & Syder (1977; 1983) also bear directly on the question of generalized constraints in discourse (and the quotation at the head of Chapter 6 from Ammann (1928) provides one indication that those linguists attuned to the nature of discourse have been thinking in general terms about related issues much earlier still).

Having outlined the discourse-pragmatic structures of the present primary corpus, we are now in a position to examine a number of the hypothesized constraints which crucially involve the discourse-pragmatic dimension. In this section I evaluate four such hypotheses—originally put forth on the basis of other languages—with respect to the conversational German discourse analyzed in this study. Two of the hypotheses
discussed here are to be found in Du Bois (1987) and two in Chafe (1993). I begin with the most specific constraint and end with the most general.

4.5.1 New information and transitive subjects

The first constraint I will examine is proposed by Du Bois (1985; 1987) and concerns grammatical subjects of transitive verbs. Called the 'Given A Constraint', it is stated as a preference for surface A-role arguments in clauses to be non-new, or as Du Bois writes: 'Avoid new A's' (1987: 827). 'A' here refers to the grammatical role 'transitive subject', in contrast to 'S', 'intransitive subject' (the third core grammatical role is 'O', '(transitive) object'), following Dixon (1979); cf. Comrie (1981). In other words, it is hypothesized that transitive subjects do not function to introduce new referents in discourse.

A tally of the transitive subjects in my corpus of German conversational discourse supports the Given A Constraint. In this corpus, there are a total of 273 A-role arguments. Of these, only 2 (<1%) express referents with an activation state value 'new'. The obvious conclusion is that new A-role arguments are indeed avoided. Even after taking into consideration the fact that one third of the A-role subjects are first and second person mentions—which will always be non-new—the number of exceptional cases is indeed small.

As strong as this evidence is, more support for the Given A Constraint can be adduced. The interaction of prosody, syntax, and discourse pragmatics in the two exceptional cases provides additional insight into the
scope of the constraint, so I will now turn to a discussion of them. As it happens, both instances of new transitive subjects occur within a short stretch of speech by Uli, which I give in (19) below. (Referents with the status ‘new’ are underlined in this and the other examples in this section.) Uli has just started to relate an experience he once had where various family members were supposed to rendezvous at the Frankfurt international airport.

(19) Corpus excerpt from ‘Airport chaos’ [Uli’s visit 597-613].

597 [U]:  #\( und da [ =hätte ] -- \)
and there had
598 (A):
[ mh ] \*
599
\(.8\) mèine --
my
600 \(.8\) meine Große Mutter hatte das auch initie=rt /+ }
my grandmother had that also initiated
601 #\( .. h%-- ... und die wär schon -- \)
and she was already
602 #\( .. vörher -- \)
before
603 #\( die kam aus Nördeutschland /+ } \)
she came out of Northern Germany
604 #\( .. (H) „.5” un’ wir söllten uns dort =trèffen /+ } \)
and we should ourselves there meet

597 [U]:  and <PAST> --
598 (A):

mhm.
599
my --
600
my grandmother organized it all,
601
and she already --
602
before --
603
she arrived from Northern Germany,
604
and we were supposed to meet there [Frankfurt],
The first exceptional case concerns the clause construction in IUs 597-600. The construction begins with a common conjunction + adverb + auxiliary structure und da hatte. After a substantial hesitation, Uli con-

24 I consider the clause fragment in IU 597 to be associated with the structure in IUs 599-600. Other analyses of this sequence are possible.
continues briefly with the clause in IU 599, then hesitates again. Following this second hesitation, he completes the construction in IU 600, introducing a positional switch for the grammatical subject. That is, when the clause construction began, the projected syntactic position of the subject was in the midfield, following the auxiliary. Indeed, the subject does follow the hatten of IU 597, but in IU 600 hatten is repeated after meine Großmutter, effectively resulting in a prefield position for the subject. Thus, this clause construction has undergone a significant reformulation over the course of its development, and this reformulation, with the hesitations and repetition, suggests that the Given A Constraint may apply with less force here. As the inner bracketing of the clause appears to be constructed over the course of three IUs (there are no peripheral elements in the construction), we have overt evidence that the speaker has taken extra time to express his current thought. With more planning and mental editing, cognitive processing constraints of the type considered here may not hold to the same degree.

This example highlights the relevance of clause constructions when considering information flow questions. Had the ‘repaired’ material in IUs 597 and 599 simply been ignored and the boundaries of the clause been taken to coincide with IU 600, then the example would have constituted a blatant violation of the Given A Constraint. The facts are, as explained, more subtle, and in light of the larger, emergent structure, the violation is mitigated.

The second exceptional case concerns the stretch of speech in IUs 609-
10. In this case, the syntactic structure under consideration clearly develops over two IUs. The transitive subject of the clause construction, *der i-ce*, is expressed in IU 609, while the grammatical object and finite verb are expressed in the following IU. There is no repetition or reformulation as in the previous example, but a 1.0 second hesitation does occur. In light of this prosodic break, it is quite conceivable that Uli did not yet have the entire clause completely formulated at IU 609. (Note also that the clause construction extends beyond IU 610.) The idea to be conveyed at that point may only have been this: ‘we were late because of the train’. Moreover, a variety of completions would have been possible in IU 610, including a formulation with an intransitive verb (e.g. *spät war* ‘was late’). In other words, at IU 609 the projected subject *der i-ce* is not yet obviously associated with a transitive verb. It is only after IU 610 has been uttered that this structure emerges. Although *der i-ce* is indeed a transitive subject in the clause construction viewed in its entirety, at the point in time where it is expressed, it has not yet acquired that role. Thus, with a temporal component added to its evaluation, the Given A Constraint cannot be said to be seriously violated here, either.

In sum, viewed from the point of the syntax alone, there are few exceptional cases to Du Bois’ Given A Constraint in my corpus of German conversational discourse. Viewed from the point of view of emergent structures in discourse, there are no hard exceptions. This fact can be taken as support for this constraint.

341
4.5.2 New information and grammatical subjects

The second hypothesis to be examined is proposed by Chafe (1993) as the 'Light Subject Constraint', whereby '[grammatical] subjects carry a light information load, as is appropriate for starting points' (81). I gave Chafe's definition of 'starting point' in §4.2; what it is to be 'light' remains to be explained. With respect to grammatical subjects, 'light' is said to signify one of two properties: (a) bearing a non-new activation state value, i.e. given or accessible, or (b) introducing a new referent which receives no further mention in the discourse (called 'trivial' by Chafe). Reformulated in terms of the referential function property I discussed in §4.4, the Light Subject Constraint would read: tracking new mentions may not occur in subject position. Note that this constraint, which applies to both S-role and A-role surface subjects, substantially overlaps (but does not entirely subsume) the Given A Constraint just discussed. Since it has already been demonstrated that new A-role subjects rarely occur, additional violations will be due only to subjects in an S-role.

A strong trend for light subjects is clearly indicated by the German conversational discourse of the primary corpus. Of 799 grammatical subjects, only 22 (3%) have an activation state value 'new'; and of these, 12 subjects are 'trivial' in Chafe's sense. The Light Subject Constraint is thus supported, but as a tendency like the Given A Constraint, not an absolute prohibition.

An examination of the remaining 10 exceptional cases reveals an important, striking regularity. In the majority of cases, the new subject oc-
curs late in the clause, rather than in the first constituent position (prefield). Example (20) illustrates:

(20) Corpus excerpt from 'Aldi drunk' [Skat: 57-8].

57 [T]:  #{ (H)..<7> un' nun =kommt auch gleich der Geschäftsführer /+}  
        and now comes also immediately the boss
58    #{ und .. "wäs" /+}  
        and what

57 [T]:  and then the boss comes over,
58     and "what",

The relevant clause construction here is coextensive with IU 57. The adverb nun occupies the prefield, while the new (and nontrivial) subject der Geschäftsführer occurs after the finite verb kommt, as the last constituent of the clause.

The first transitive subject in (19) provides an additional (though less obvious) example. (The second transitive subject, while new, is trivial, and so is exempted by criterion (b) of the Light Subject Constraint.) Although the (nontrivial) subject meine Großmutter appears to occur in the prefield when only IU 600 is taken into account, its initial projected position is clearly postverbal, i.e. in the midfield, as the adverb da occupies the prefield. In light of the auxiliary switch which is effected later in the clause construction, the subject is not unambiguously in either clause field.

Interestingly, this subject-late pattern is not limited to the 10 exceptional cases. Indeed, it can be observed that out of all 22 new subjects, 19 (86%) do not occur in first position before the clause-second finite verb. In
other words, the clause prefield rarely contains new subjects. Instead, they occur postverbally (in the midfield), in verb-late clauses, or in reformulations with auxiliary switch, as in the example from (19) just mentioned. This pattern is reminiscent of that found by Bentivoglio & Weber (1986) for spoken Spanish, where new subjects typically occur postverbally.

The fact that word order considerations account for almost all of the exceptional cases in the corpus (and for most of the ‘trivial’ subjects as well) raises the possibility that there may be a dimension of the Light Subject Constraint yet to be explored. In particular, the word order factor suggests that the relationship between starting point and subject might need to be reconsidered for a language like German. In this regard, it is important to note that Chafe’s hypothesis is more accurately referred to as the ‘Starting Point Constraint’ (Chafe 1993: 87). That is, the constraint is actually meant to apply to this functional notion, rather than to surface subjects. If starting points in German are not (fully) grammaticized as subjects, then it can be expected that not all subjects will conform to the Light Subject Constraint. Only those expressing starting points would fall under its scope.

On the other hand, evidence can be inferred for grammaticized starting points in German. Of the tracking new mentions that do not occur in the inner bracketing of clause constructions, 8 out of 12 are correferential with the subject of the clause. That is, in two thirds of these cases, the mention introducing a new referent occurs at the clause periphery (usually in a separate IU), allowing for the correferential mention occur-
ring as the grammatical subject of the clause to be given. Thus, many peripheral new mentions appear to function (in part) as a way of keeping the subject from expressing new information.

The relative roles of referential importance and word order variation in the definition of 'light' need further investigation in German. In any case, the instances of exceptions to the Light Subject Constraint are few, so, in sum, the hypothesis is in essence upheld.

4.5.3 New information and clausal arguments

The third constraint I wish to examine—the second in the discourse-pragmatic dimension proposed by Du Bois (1987)—has to do with the quantity of new information in clauses viewed as wholes, rather than in any particular grammatical role. Du Bois again formulates the constraint as a tendency, thus: 'Avoid more than one new argument per clause' (826). It is important to note that this 'One New Argument Constraint' is meant to apply only to nonoblique nominals, and in particular only to A, S, and O-role arguments. However, given the facts of German grammar, a modification of this limitation in the form of an addition seems in order here. The addition I have in mind concerns verbal arguments marked with the dative case. Because German has a dative object that is indicated morphologically like the other arguments, I will include these 'D-role' arguments in my evaluation of the constraint. Therefore, the (revised) constraint can be reinterpreted for German discourse as follows: 'Avoid more than one new case-marked verbal argument per clause'.
In the German conversational discourse of the primary corpus, a clear preference for clauses to be limited to one new grammatical argument exists. In the primary corpus, 302 clauses have more than one case-marked verbal argument. Of these, only 3 (1%) clauses occur with two arguments having an activation state value ‘new’. As with Du Bois’ first constraint, the One New Argument Constraint is thus strongly supported.

Again, in order to better understand the scope of the constraint, it is instructive to look at the exceptional cases. Two instances occur in the following conversational excerpt, in which Marlies is telling about the actions of a mime she had been watching:

(21) Corpus excerpt from ‘Weavers’ fair’ [Doppelkopf b: 34-46].

34 [M]: #\{ ... ~mein’tweg’n dann ging er=% .. h\in= /+ \} 
    because.of.me then went he thither

35 #\{ (H) «4» n\åhm dem %èinen /+ 
    took from.the one

36 {{{ ~w\ås wei\ß ich /+ }}
    what know I

37 aus ’m P\üblikum= \+
    out.of.the public

38 uh= .. «5» das Fèuerzeug weg /+
    uh the lighter away

39 #\{ g\åb das irgendjemand’m an- am ånder’n [Ènde ] /+
    gave that to.someone at.at.the other end

34 [M]:  so then he went,
35 took (away the lighter) from one [person],
36 let’s see,
37 in the audience;
38 ( ),
39 gave it to someone at the other end,
The two exceptions to the One New Argument Constraint in (21) involve the clause constructions in IUs 35-8 and 41-3. In both cases, an A-role subject is not overtly expressed, but the clauses contain O and D-role arguments, with both arguments bearing an activation state value ‘new’. (The third exceptional case, which can be found in the longer stretch of discourse analyzed in Chapter 5—see IUs 137-44 there—also has new O and D-role arguments.)

Since I demonstrated above that A-role arguments in German hardly ever introduce new referents, it is not surprising that the three exceptional cases involve dative objects. On the other hand, I showed in Schuetze-Coburn (1987b) that dative objects do not normally express new
information, either. The occurrence of two neighboring clauses containing new D-role as well as new O-role arguments thus points not to any discourse regularity, but to the particular event structure of the conversational story and the need to convey a series of similar sequential actions in a short stretch of talk. In other words, a configuration of ‘high information pressure’ (Du Bois 1987: 834) is at work here. The mime—the agent in the story—repeatedly takes personal items from random members of the audience and gives them to other random onlookers. Each act requires the mention of a new referent, which, when expressed as an argument of a verb of giving or taking, will appear as a noun phrase in the dative case (semantically, a source with the verb *wegnehmen* ‘take away’ and a goal with the verb *geben* ‘give’). The repetition of taking an object and then giving it results in a sequence of four adjacent new dative mentions, two of which cooccur with the new O-role object mentions.

This situation of (near) maximum information pressure allows us to gain a fuller view of the ‘verbalization’ process. With regard to the One New Argument Constraint, a number of possible implications are suggested. First, it may be the case the extension of the hypothesis to include D-role arguments is not well motivated. If this were true, then the constraint would have no counterexamples in the present corpus. Second, it may be the case that some notion of referential importance should be incorporated into the constraint, as is found in the Light Subject Constraint. Note that each of the new dative mentions in (21)—*dem einen* in IU 35, *irgendjemand’m am andern Ende* in IU 39, *‘ner andern Frau* in IU 41, and *jemand anders* in IU 46—expresses a referent that is of no further
consequence to the story. (This is not true for the referent expressed by the exceptional dative mention in Chapter 5, however.) Third, it may be the case that the constraint is too narrowly formulated in another way. Up until this point, constraints on information flow have been phrased only in terms of grammatical relations, and while my conception of clauses in discourse is sensitive to the prosodic structure of utterances, the constraints themselves have not incorporated this dimension of discourse production. It is to this issue that I now turn in examining the last constraint.

4.5.4 New information and intonation units

The final constraint that I will examine here is the ‘One New Idea Constraint’. Chafe (1987; 1993) puts forward the hypothesis that ‘... an intonation unit can express no more than one new idea’, or viewed in terms of the process of making inactive ideas active in consciousness, that ‘... thought, or at least language, proceeds in terms of one such activation at a time, and each activation applies to a single referent, event, or state, but never to more than one’ (1993: 105). Accordingly, there should be at most one mention in an IU that has the activation state value ‘new’. It is obvious that this constraint differs substantially from the first three I considered. First, it makes no claim about specific grammatical roles. It is indifferent as to the syntactic structure in its domain, the IU. Second, it is much more general in nature. It applies not only to mentions that express referents, but to the ‘verbalization’ of all ‘ideas’. That is, it covers all major linguistic units, not only noun phrases.
As I have mainly been concerned with new referent mentions in this chapter, I will focus my evaluation of the One New Idea Constraint on IUs which involve them. That is, I will consider only one aspect of the hypothesis here, though undoubtedly the one that includes the most cases. (Other investigations might focus on IUs which contain more than one clause, i.e. more than one lexical verb, or on those containing adjectives.) Specifically, I will examine the activation state values of referents, events and states in IUs known already to express a new referent. The expression of any previously inactive event or state concept, as well as the presence of an additional inactive referent, in such an IU would constitute a violation of the constraint.

In the primary corpus of German conversational discourse, speakers by and large adhere to the One New Idea Constraint. Of the 327 IUs (out of 1736) that contain new referent mentions, only 27 (2% of all IUs) show a possible violation. (What I mean by ‘possible’ here will be explained below.) Thus, as a probabilistic statement on the distribution of new information, the hypothesis is confirmed.

Two striking examples of one effect of the One New Idea Constraint can be found in excerpt (21) above. As discussed, the excerpt illustrates a brief stretch of talk having a particularly high degree of information pressure. Two of the clause constructions each contain two new referent mentions. But in both instances, the new referents are expressed in separate IUs: the new D-role object dem einen does not appear in the same IU as the new O-role object das Feuerzeug, and the new D-role object ‘ner
andern Frau does not appear in the same IU as the new O-role object *den Regenschirm*. In contrast, note that when the O-role objects express given information in IUs 39 and 46 (expressed as *das* ‘that’), they do occur in the same IU as the new D-role object. This pattern seems to be the norm in the German data. In 39 out of the 42 (93%) clause constructions containing two new referent mentions, the new referents are expressed in separate IUs.

The remaining exceptional cases involve new event or state concepts cooccurring with a new referent. Interestingly, the excerpt just discussed also exhibits two of the violations of the One New Idea Constraint considered under this broader application. IUs 35 and 39 each contain not only a new referent, but a previously inactive event concept as well. In IU 35 *nahm*, the finite part of the verb *wegnehmen* expressing the concept «TAKE AWAY», occurs with the new referent mention *dem einen*, and in IU 39 *gab*, expressing the concept «GIVE», occurs with the new referent mention *irgendjemand’m am andern Ende*. In these two cases, it is not clear how the violations can be explained.

On the other hand, many of the exceptional cases seem to be amenable to various kinds of post hoc explanations. In a few instances, it may be the case that what I have analyzed as separate object and event concepts was intended by the speaker as one complex concept.\(^{25}\) More significantly, almost three quarters of the apparent violations involve new but identifiable mentions. Due to the identifiability of the referent, the speaker may

\(^{25}\) Chafe (1993: 106-10) discusses the interpretation of these predicate types in detail.

351
simply have misedjudged the referent's activation status in the listener's consciousness, assuming it to be semiactive rather than inactive. Such a misjudgement would lead to a violation of the One New Idea Constraint from the viewpoint of the analyst.

Another type of special circumstance is illustrated in example (22):

(22) Corpus excerpt from 'Propane tank' [Morning visit: 263-4].

263 * ...(5.8)
264 T: #{ die lètzte Gásflasche hab'n w'r zu Fúß gehòlt gehàbt \* }
the last propane tank have we to foot picked.up had

263 * (pause)
264 T: the last propane tank we carried home.

IU 264 contains the two newly activated concepts «LAST PROPANE TANK» and «GET BY FOOT». It can be argued, however, that the contents of this turn-initial IU have been expressed less spontaneously than that of most other IUs. Tamina has had almost six seconds of silence to formulate the clause—seemingly ample time for mental construction and editing, and time that is not generally available for the formulation of an IU. Over one third (10/27) of the exceptional cases show initial pauses of one second or longer.

Given these apparently well-motivated exceptions to Chafe's hypothesis, it would seem reasonable to allow for special cases in applying the constraint. However, rather than incorporating miscellaneous conditions into the constraint directly, it may prove more insightful to specify general characterizations of discourse where the One New Idea Constraint
is valid and where violations may surface. I have suggested a few possible considerations in light of the German data, but a thorough investigation of the matter must await further study.

In sum, I have evaluated four hypotheses concerning the distribution of new referent mentions in discourse. All four hypotheses are generally supported by the German conversational discourse in this study. Exceptional cases are limited to 1-3% of the data. Furthermore, attention to the process-oriented structure of clause constructions—which incorporates the prosodic phrasing of the utterances—lessens their impact in many instances.

4.6 Summary

In this chapter I examined issues concerning one topic of discourse pragmatics, namely 'information flow'. I began the chapter by sketching the distinction between propositional and discourse pragmatics, and explaining the notion of information flow. I discussed the uses of 'given' and 'new' in relation to information flow, and I outlined Chafe's model of it, where I paid particular attention to the properties 'activation state' and 'identifiability'. In the third section I critically discussed Chafe's notion of 'idea', proposing that a separate notion of 'concept' is necessary to account for cases where the activation state of an individual and that of a class differ. I also discussed the determination of activation state from an analytic viewpoint, concluding that identifying the 'new' mentions is in many respects easier than identifying 'given' or 'accessible' ones. I presented an information flow analysis of the corpus of conversational Ger-
man with a focus on new mentions, and I compared the distribution of
the new mentions with the prosodic and clausal structures of the corpus.
Finally, I evaluated four hypothesized constraints put forth by Du Bois
and Chafe regarding the interaction of prosody, syntax, and information
flow. It was found that although a few exceptional cases for each con-
straint did occur in the corpus, the hypotheses were in essence well sup-
ported.
5. An Analysis of Conversational Discourse

Die Umgangsprache im Gegensatze zur Schriftsprache ist bis jetzt immer nur mit einzelnen Ausserlichkeiten gestreift worden ...
In contrast to the written language, ordinary colloquial speech up until now has been touched only on the surface ...

Hermann Wunderlich
*Unsere Umgangsprache*, 1894, p.VII

In the previous three chapters, I have presented the mechanics for an analysis of the prosody, syntax, and discourse pragmatics of natural spoken discourse, as well as the quantitative results of a comparison of basic features of these dimensions for a corpus of conversational German. In this chapter I present a detailed sample analysis illustrating the application of the categories and procedures developed. The main purpose for taking the reader through the following extended excerpt of a single conversational exchange is to provide a context for the better appreciation of the particular strategies that I have employed. While it is useful to have short, specific examples culled from the corpus when attention is devoted to a specific topic, it is equally important to be able to see how all the pieces of the puzzle fit together. Examining an extended and challenging conversational excerpt in detail is one way to achieve this goal. Along the way it is hoped that the reader will also gain a clearer picture of the benefits—and the demands—of whole-text analysis.

My example follows the lead of Chafe (1987) and Du Bois (forthcoming), but I devote additional attention to explaining the basic prosodic and
syntactic structures discussed in earlier chapters. In the first section I present the glossed and translated conversational discourse as a whole. As has been the case throughout this work, the excerpt is arranged by intonation units (IUs), reflecting, among other things, the sequential production of the talk. In the subsequent sections, I turn to the analysis and commentary. The order of these sections is intended to reflect the organization of Chapters 2-4. That is, I undertake a step-by-step analysis of the discourse, beginning with the prosodic structure and ending with the information flow, but in each case building on the analysis of the previous dimension. Note that a few IUs before and after the ‘story’ division (see §1.4) are also presented in order to place it in the conversation as a whole. These IUs, however, are generally not included in the analysis.

Since it is fundamental to the analysis, the prosodic phrasing of the excerpt is discussed first, in the second section. For each IU, the prosodic features covered in Chapter 2 that have been used to segment the sequence of utterances are spelled out as carefully as possible. Transcription conventions for prosodic features have been explained extensively in §2.3, but are also listed in Appendix 1 for convenience. In lieu of an interactive digitization of the audio recording embedded in the text of §5.1 (which no doubt someday will be a reality), fundamental frequency (F₀) tracings of most IUs in the sample exchange have been provided in Figures 5.1-5.6 for the reader’s inspection.¹ References to contours are largely based on

¹ The recording was digitized at 8 kHz and analyzed on a Sun SparcStation using Waves at the Phonetics Lab, University of Pennsylvania. I thank Cindie McLemore and Felicia Hurewitz for their help obtaining the F₀ tracings.
these figures. Exact \( F_0 \) values for specific points, however, have been derived from separate calculations, in which the inverse of the time between glottal pulses (determined from separate waveforms) was taken.\(^2\)

In the third section the clause structure of the passage is outlined, as it is incrementally constructed, and its relationship to the prosodic structure is explained. As discussed in Chapter 3, the notion of 'clause' employed here is somewhat broader than is traditional, and I use the term 'clause construction' to set my usage apart; see §3.2 for my particular interpretation. Conventions for delineating clause constructions have been discussed in §3.3, but I will outline them here. Briefly, construction units are indicated by a number sign ('#'), usually at the beginning of a line. (The subscripted integer following a number sign is an ad hoc means of indexing clauses for ease of reference in this chapter.) Within every clause construction, each instance of an initial clause boundary is also marked with an open curly bracket ('{'), each instance of a final clause boundary with a close curly bracket ('}'), with the exception that embedded clause boundaries are indicated instead with double curly brackets ('{{' or '}}'). Brackets do not necessarily pair up.

In the fourth section aspects of the discourse pragmatics—the information flow—of the passage are laid out. The focus is on referents which are 'tracked' through the 'story', with special attention given to mentions

\(^2\) Waveforms were obtained from speech digitized with MacRecorder (also 8 kHz); display and measurement from SoundEdit. Sorensen & Cooper (1980: 405-6) nicely illustrate this method of \( F_0 \) calculation.
which introduce new concepts/ideas. The main categories included in the analysis are activation state and identifiability. The principles behind these categories are not recapitulated here; they have been discussed in §§4.2.2-3. An additional category of 'generality', briefly mentioned in §4.2, is also included, which I do outline below. The mentions under investigation are highlighted in boldface type in the transcript. Specific information statuses are not marked, but are given instead in the commentary portion of this chapter. Particular attention is paid to the distribution of referents vis-à-vis the prosodic and syntactic structures in which they are expressed.

The sample excerpt is part of an extended casual interaction between two close female friends, Elena and Sabine, about twenty years of age. In the recorded portion of this interaction (about one hour), the interlocutors both recount recent experiences and comment on topics that were relevant in their lives at the time of their conversation. The verbal interaction thus consists of a combination of more narrative-like portions, where one speaker makes most of the contributions, and portions where change of speakers is frequent. The selected passage presented here is more narrative-like in this respect.

Prior to the prosodic analysis, the exchanged was divided into speaker turns. Assigning a turn status to virtually all of the utterances in the excerpt was unproblematic. Elena (Speaker E) clearly is the main speaker throughout the passage, and most of Sabine’s (Speaker S) contributions make no claim on the floor, i.e. they are backchannel utterances (see §1.5).
Turns are indicated by the placement of speaker labels. For example, the use of the label 'E' in IU 130 (see §5.1 below) signals that this is the beginning of Elena's turn. As long as Elena has the floor, the label is not repeated. Backchannel utterances are indicated by enclosing the speaker label in parentheses.
5.1 The conversational story 'Dirges' [After school tea: 135-192]

130 E: #{ ... glàub' ich nich' \+ }
    believe I not
131 #{ ... ((noise)) (H) «.9» ^ ~àußerdem nù=r (%) —
    besides only
132 { ^ ich héirate Rùdiger nùr /+
    I marry R only
133 {{ ^ «A das hàb' ich ihm schon gesàgàt A> /+ }}
    that have I to him already said
134 { (H) «.2» wenn er gànz ganz =rèich wird \* }
    when he quite quite rich becomes
135 #1 { ... (H) «.5» ~wèiùt 'e [ wàs= ] /• }
    know you something
136 (S): [ (@@ ] @ «.3»
137 #2 { ^ ~ér [[ will so dèn /+
    he wants like to the
138 (S): [[ @ @
139 { .. ^ den rèichen /+
    to the rich
140 (H) «.2» ^ Amerikàñern /+
    Americans
141 [ ~will er ] --
    wants he
142 (S): [(COUGH)]
143 (H) ((blows out candle)) «1.3»

130 E: I don’t think so,
131 besides only --
132 I’ll only marry Rüdiger,
133 I’ve already told him,
134 after he’s very very rich.
135 you know something?
136 (S): (laughs)
137 he wants «to sell» to the,
138 (S): (laughs)
139 to the rich,
140 Americans,
141 [he wants --
142 (S): ((coughs)
143 (blows out candle)
will er Träuermäsche verkäufen  •  

wants he dirges sell

#3{ ... (H) «.6» ~er wûrde dann sò = (%) +
he would then like

.. hündert --
hundred

.. nèin ◊ .. zèhn'tausand Dòllar /
no ten thousand dollars

für èin'n Träuermarsch nèhm'n / • }

for a dirge take

#4{ ... (H) «.7» ^~un' d'n [wûrd' er ihn'n ] +
and it would he to them to them

(S): [ «p mein Gott p » ] +
my God

.. ^~ihn'n .. ihn'n ◊ gànz speziell schriëiben 

to them quite specially write

#5{ ~%also er wûrde er so~ +
so he would he like

(H) «.3» ^ vörher so 'n hàlbes Jàh=r /
before like a half year

auf so % äh % «.2» --
on like uh

^ bei dem lèb'n müssën / • }
by him live must

#6{ .. (H) «.4» ^ ~dmit er sie richtig kœnn'nlernt /
so that he them correctly gets to know

#7{ ^ bevor die Leute stèrben /
before the people die

he wants to sell dirges.
he would (take) then like,
a hundred --
no ten thousand dollars,
per dirge ()?
and he would (write) it for them,
(S): my God,
for them for them specially (),
so he would 'have to' like,
half a year before,
at like uh --
have to live at his 'house'?
so that he really gets to know them,
before the people die,
158  

-#8\{ ^\text{das wei\ss} \ [ja\ \text{man}]\ ja\ m\=e\text{-stens} /+ \} 
that\ knows\ yes\ one\ yes\ mostly

159  \((S)\): 

-\(\text{(T)\}}\)

160  \(9\{\ ..\ \text{bevor\ man\ so\ stirbt} /+\}\)
before\ one\ like\ dies

161  \(-\text{jedenfalls\ bei}\ (%)\ 0.\ \text{Leuten} 10\ \{\text{dieso\ l\=anger\ k\=rank}\ \text{[sind]}\ \text{n\=e}] /+\}
\text{in\ anyway\ case\ by\ people\ that\ like\ longer\ ill\ are\ ok}

162  \((S)\): 

-\(\text{ja}\) \(+\)
\text{yes}

163  \(\#11\{\ \text{(H) \ll\text{.7}\ ~also\ n\=ur\ ganz\ ganz\ stinkr\=eiche\ Amerik\=aner} /+\}\)
so\ only\ quite\ quite\ filthy\ rich\ Americans

164  \(\sim\text{so\ so\ n\=e} /+
like\ like\ ok

165  \(^\text{so= [\text{Industrielle\ oder\ so}} /\cdot\}
like\ industrialists\ or\ like

166  \((S)\): 

-\(\text{mhm} /+

167  \(\#12\{\ \text{(H) \ll\text{.5}\ \^\text{A\ un’\ d:\mit\ er\ die\ richtig\ k\=enn’nlernt} /+}\}
\text{and\ so\ that\ he\ them\ correctly\ gets\to\ know}

168  \(\#13\{\ \text{un’\ dann\ w\=urde\ er\ ihn’\ n\gg\ ganz\ individu\=el} /+
\text{and\ then\ would\ he\ to\ them\ quite\ individual}

169  \(\ldots\ \text{’n\ Tr\=auermarsch\ schr\=eiben} /\cdot\}
\text{a\ dirge\ write}

170  \(\#14\{\ \ldots\ ^\text{un’ den\ w\=urde\ dann\ ganz\ ganz\ gr\=o\=z\ inszenier’n}= /+
\text{and\ it\ would\ he\ then\ quite\ large\ produce}

171  \(\text{mit\ br\=enn’\den\ F\=ackeln} /+\}
\text{with\ burning\ torches}

158  \(\text{you\ usually\ know,}\)
159  \((S)\): 

-\(\text{clears\ throat}\)
160  \(\text{before\ one\ dies,}\)
161  \(\text{with\ people\ that\ have\ been\ ill\ a\ long\ time\ anyway,}\)
162  \((S)\): 

-\(\text{yeah}\)
163  \(\text{so\ only\ filthy\ rich\ Americans,}\)
164  \(\text{like\ like,}\)
165  \(\text{like\ industrialists\ or\ something,}\)
166  \((S)\): 

-\(\text{mhm,}\)
167  \(\text{and\ so\ he\ really\ gets\ to\ know\ them,}\)
168  \(\text{and\ then\ he\ would\ (write)\ for\ them\ individually,}\)
169  \(\text{a\ dirge}\ ()\)
170  \(\text{and\ then\ he\ would\ produce\ it\ grandly,}\)
171  \(\text{with\ flaming\ torches,}\)
right and left
and so forth,
and with that he would --
(S): (laughs)
he would he would become rich,
he told me.
isn't that wonderful?
S: I think it's great.
E: he can already write beautiful dirges,
and he compo--
composes so well.
[it's so wonderful,
(S): [really? s
suddenly he jumps out of bed,
186  #23{  ... (N) «.6» un' rënnt zum @ Klavier @@; /+ }  
and runs to the piano

187  [ @=

188  (S):  [ @= «1.0»

189  S:  #24{  [2 @ der komponiert @> /+ ]
          he composes

190  (E):  [2 @=

191  E:  #25{  [ @ und [ da wà−r ihm wieder was éingefall'n @> \+ ]
            and there was to him again something occurred

192  S:  #26{  [ das find' ich ja süper @> \+ ]
            that find I yes super

193  E:  #{  .. ja= /+

194  {  ~er schrèibt ja auch die gänzen Lieder /+ }
          he writes yes also the entire songs

195  für die= Bànd \ } .. nè /+ }
          for the band ok

196  S:  #\{  .. èch' /•
          really

197  \{  .. die [ schrèibt er /• }
          they writes he

198  E:  #\{  [ un' die sind ja genià=1 \ }  <@ pe> /+ }
          and they are yes brilliant ok

186  and runs to the piano,
187  [(laughs)
188  (S):  [(laughs)
189  S:  [2 he composes,
190  (E):  [2 (laughs)
191  E:  and there was something that occurred to him,
192  S:  I think that's super,
193  E:  yeah,
194  he also writes all the songs,
195  for the band,
196  S:  really?
197  he writes them?
198  E:  and they're brilliant,
5.2 Prosodic analysis of 'Dirges'

In the analysis in this section, I consider each IU boundary (and any other possible phrase-level boundaries) one at a time, noting the specific prosodic features which play a role in signaling the boundary. I begin with the juncture marking the start of IU 135, the first IU of the story 'Dirges'.

The prosodic break between IUs 134-5 is a substantial one, consonant with the beginning of a new discourse topic, or subtopic, as is the case here. The final syllables of IU 134 are lengthened (the approximate onset indicated by '≈'). The extent of IU 134 is also indicated by the pitch contour (see Figure 5.1 below, top).³ With the exception of a final excursion on reich (at time 133), which constitutes the primary accent of the unit, the pitch falls continuously after the first instance of ganz (at time 132.5) all the way to the end of howd, to below 200 Hz (F₀ on howd does not appear in the tracing, but is measurable from the waveform). The fall to low is marked in the transcript by \\.

³ In the figures, F₀ is plotted in hertz (Hz) as a function of time in seconds. Octave errors have been adjusted manually. Text alignment is approximate. Initial IU boundaries with the IU sequence numbers have been placed above (in a few cases, below) the corresponding contours. 

365
Figure 5.1. $F_0$ tracings for IUs 134-144.
best determined on unaccented syllables), it is clear that IU 135 constitutes a separate intonation contour, as there is a primary accent on *was*. Furthermore, two completely different contours can be discerned: in IU 134, a 'final fall'; in IU 135, a 'high rise' (transcribed ‘/•’).

IU 135 is overlapped somewhat, and it is separated from IU 137, by three pulses of intervening backchannel laughter from Sabine (IU 136). The break thus created between IUs 135 and 137 is fairly short (about 0.3 second), but there is, additionally, lengthening on the final segment of IU 135 (transcribed ‘=’). Moreover, the pitch juncture is unequivocal. After producing the rise to high pitch in IU 135, Elena resets her pitch at the beginning of IU 137 (transcribed ‘^’). That is, F₀, which is high at the end of IU 135, drops to a mid-range level for the speaker at the beginning of the next unit (*er* starts at 216 Hz).

IUs 137 and 139 are separated by a slight pause (transcribed ‘..’). More seems to be involved here. The pitch patterns of the units are likely to be relevant, but evaluating the intonation contours of IUs 137 and 139 acoustically is complicated by the overlapping backchannel laughter in IU 138. Auditorily there is a short, but completed nonfalling, continuing pattern on IU 137 (transcribed ‘/+’); the primary accent falls on *den*. The pitch of the following IU (i.e. the repetition of *den*) begins at a higher level, dropping (reaching a minimum at time 136), and then rises rapidly, with the primary accent on *reich-*. Thus, whether a reset is present is unclear, but there do appear to be two contours.

The course of pitch at the boundary between IUs 139-40 is in this case
indicative (see Figure 5.1, top and bottom). Following the large pitch excursion on *reichen*, the initial pitch on *Amerikanern* (at time 136.5) is substantially lower, indicating a pitch reset. The reset is expected in this case, as there is an inhalation between IUs. A primary accent falls on *-kan*, again suggesting that two IUs are involved.

The break between IUs 140-1 can be attributed largely to a change in tempo at the boundary. As the intonation contour on *Amerikanern* instantiates a typical nonfalling, continuing pattern, the marked increase in the rate of articulation starting with *will* at the beginning of IU 141 (transcribed `~`) indicates that Elena is beginning a new IU.

IU 141 is truncated (not completed) and is thus transcribed `~`. Perceptually, there is no primary accent in this short stretch of speech. After the overlapping cough by Sabine, Elena inhales and commences to blow out a candle in the room, thus interrupting the flow of speech to the degree that a new IU is virtually certain, even though there does not seem to be a significant change in pitch across the break.

The juncture between IUs 144-5 is signaled by pitch, tempo, and pause features. IU 144 ends in a final fall contour (again, falling below 200 Hz). Although there is no large reset at the beginning of IU 145 that would create an obvious boundary (see Figure 5.2 below, top), the forms of the contours indicates that a boundary is to be drawn. The final fall of IU 144 occurs after the primary accent on *Trau-*, signaling a juncture, and the rise on *würde* (at time 141) in IU 145 to a high level contour constitutes a separate contour. Together with the combination of pause and inhalation,
as well as the accelerated speech at the beginning of IU 145, the prosodic break is unambiguous.

The end of IU 145 is marked by final prosodic lengthening and a glottal stop (transcribed ‘%’), terminating the phonation of the lengthened vowel, which also bears the primary accent of the contour. Given that there is a pitch drop at the very end of the IU, it might be expected that the initial pitch on *hundert* (at 256 Hz) would constitute a pitch reset. The perceptual status of this kind of transition is not entirely clear, though. The last quasiperiodic glottal pulse of IU 145 registers at 268 Hz, so that in some sense continuity has been maintained. Furthermore, there is a prominence on the initial syllable of IU 146, which complicates the interpretation. IUs 145-6 are also separated by a slight pause.

The prosody of IU 146 is interrupted so that the unit is rendered uncompleted. While *hundert* is fully articulated, and bears an accent, it contrasts sharply with the immediately preceding IU. Although IU 145 ends in a glottal stop—often indicating a truncated unit—it nevertheless sounds complete, as the internal structure of the IU is developed: it exhibits accelerated speech at the beginning of the unit, a significant contour through the unit, and prosodic lengthening at the end. On the other hand, the prosodic structure of IU 146 is not fully developed. The perception of IU 146 as truncated implies that a new IU boundary follows. Between IUs 146-7 there is also a slight pause.

IU 147 contains a unit-internal break, a slight pause, at which point an IU boundary is conceivable (transcribed ‘⁰’). There is a discontinuity in
pitch here: $F_0$ on *nein* falls from 234-196 Hz (at time 142.5); $F_0$ on *zehn-* begins at 268 Hz. Yet, *zehn-* is accented, and the temporal organization of the sequence is such that *nein* is cohesive with the following unit.\(^4\)

The prosodic break between IUs 147-8 is minimal. There is no reset, no pause, no phrase-final lengthening. However, a sudden increase in intensity on *für* is perceived, which has a delimiting effect, much as an accelerated speech rate does. Also, the intonational pattern of IU 147 seems to constitute a ‘coherent contour’ by itself, i.e. the prominence on *Dollar*—a major $F_0$ excursion (at time 143.25)—is perceived as a primary accent in addition to the primary accent on *nehm-* in IU 148. Given that the features appealed to here are somewhat tenuous, the boundary is marked as uncertain (transcribed ‘0’), relative to other boundaries.

In contrast, the boundary between IUs 148-9 is signaled by several prosodic features. IU 148 ends in a high rise (off the scale to 476 Hz), and IU 149 begins with the pitch markedly lower (*und* registers at 250 Hz, see Figure 5.2, bottom); that is, the pitch is reset. There is in addition a distinct pause and inhalation before IU 149. And *und den würde er* is articulated quite rapidly (only three syllabic centers are perceptible), so that an acceleration in speech rate is indicated.

IU 151 is separated from IU 149 by a slight pause, a pitch reset, laryngealization, and an increase in tempo. The shift down in pitch is evident

\(^4\) The difference in pause durations surely plays some role in this regard. The slight pause at the beginning of IU 147 is about 125 msec, while the unit-internal pause is more of a ‘rhythm break’, with a duration of less than 20 msec.
despite the overlapping backchannel (IU 150), which has high pitch, but low volume. IU 149 ends above 340 Hz (at time 146), while IU 151 begins at 275 Hz. The first repetition of *ihn[e]n* in IU 151 has slight laryngealization, and both repetitions are uttered at an accelerated rate of speech, which also sets them off from the first instance in IU 149.

Following the repetitions of *ihn[e]n* in IU 151, a possible intonation break is indicated. At this point, the pitch drops somewhat, but as in IU 147, an accented syllable follows the juncture, casting some doubt on the status of this pitch shift as a boundary cue. Also playing a role is the perception of the primary accent on *schreib-*, not on either repetition of *ihn[e]n*. In any case, the juncture is uncertain enough so that the transcription of a separate IU is not warranted.

IU 151 is characterized by a pitch contour which falls at the end, but not low enough to be perceived as 'final' (cf. other falling contours, e.g. IU 144). Thus it is transcribed ‘/\+’, or falling continuing. The intonational break between IUs 151-2 does not register in the F0 tracing (see Figure 5.2., bottom, and Figure 5.3 below, top), where it appears that the pitch is maintained across units. A new contour is, however, delimited by the presence of laryngealization: the first pitch pulses of IU 152 are well below 150 Hz. The juncture is also highlighted by a marked increase in tempo which characterizes the entire unit.

The juncture between IUs 152-3 consists of pitch reset (see Figure 5.3., top) and a break for inhalation, a fairly clear boundary. (The allegro speech that extends through IU 152 does not sound so rapid in IU 153,
perhaps also contributing to the perception of a boundary.)

In contrast, the boundary between IUs 153-4 is less certain, although strong enough to indicate a new IU. Although there is no pause and little pitch movement takes place at the boundary (in fact, the pitch remains quite level, see Figure 5.3, top, at time 150.25), significant lengthening occurs on the final syllable of IU 153, which also bears the primary accent, so that the end of an intonation contour is perceived.

IU 154 is truncated; again, there is no primary accent. At the end of the unit, the hesitation particle āh falls into creaky voice ($F_0$ is very low pitch—just above 100 Hz—and does not appear in the tracing), so that a prosodic break with the subsequent utterance is indicated. (In the transcript, the IU boundary is marked as having a pitch reset. While it is true that there is a jump in pitch at the juncture, the pitch returns to the level it had before the pitch drop on the hesitation marker. Whether this kind of phenomenon actually should be included under the category of ‘reset’ is a question open to further investigation.)

The boundary between IUs 155-6 is especially strong. IU 155 finishes in a rise to very high pitch, ending at 422 Hz (see Figure 5.3, bottom), while IU 156 begins at 244 Hz. There is a slight pause between the units, inhalation, and also an obvious acceleration in speech rate.

The boundaries between IUs 156-7 and IUs 157-8 are perceptually quite salient, even though the prosodic features marking them are few. There is no pause or significant tempo variation, in each case only a very large
rise in pitch at the end of the first unit (specifically, a large pitch excursion on the final accented syllable, the primary accent), with a corresponding pitch reset in the second (see Figure 5.3, bottom, and Figure 5.4 below, top). In IU 158 the initial stop and vowel of the word das is voiceless, so that no F0 registers, but the reset is apparent on weiß. In both instances, note that the rise in pitch does not continue all the way to the end of the unit (as it does in IU 155), but dips slightly. This is one type of rising, continuing contour (i.e. ‘/+’), as opposed to a high rising, final one (i.e. /•).

The boundary separating IUs 158 and 160 is marked by a slight pause and timing cue. Notably, at the end of IU 158 there is lengthening of the final prominent syllable. In addition, the intensity at the end of the unit drops off sharply (the final syllable is largely voiceless), which contributes to the perception of a juncture after meistens. Pitch reset may also be a small contributing factor to the perception of the boundary. IU 158 ends at 220 Hz, while IU 160 begins at 250 Hz.

The boundary separating IUs 160-1 is similar. Following the primary accent on stirbt, there is a slight pause and timing cue. In this instance, the tempo shift occurs at the beginning of IU 161, with the accelerated speech rate on jedenfalls.

Within IU 161 there is a point which could be interpreted as an IU boundary, namely between bei and Leuten, where a glottal stop occurs and a slight pause is found. However, the pitch throughout this stretch of speech remains fairly level, so that a continuous contour is perceived, in
Figure 5.4. F0 tracings for IUs 157-161.
spite of the timing and voice quality variation. In such a case, it may well be a judgment call as to whether a boundary should be marked or not.

The boundary between IUs 161 and 163 is very clear, despite the overlapping backchannel (IU 162) at the end of IU 161. The rising pitch on *ne* (at time 158.5) unmistakably closes the contour, following the primary accent on *krank*. In addition, there is a rather long inhalation between units, and IU 163 begins with accelerated speech.

IU 163 is separated from IU 164 primarily by a change in tempo. There is no pause, and the pitch remains constant across the juncture. (Figures for IUs 164-6 are not reproduced, as no $F_0$ registered for this stretch of speech. The pitch is fairly low, and the volume is low as well.) The acceleration in rate of speaking at the beginning of IU 164 is quite noticeable, however.

At the juncture of IUs 164-5 there is a significant shift down in pitch. $F_0$ rises about 100 Hz on *ne* at the end of IU 164, but returns to the pre-rise level at the beginning of IU 165. This shift is interpreted as a pitch reset.

At the beginning of IU 167 there is the familiar combination of inhalation and pitch reset, which follows the high rise, separating this IU from IU 165. (The reset is not evident in Figure 5.5 below, bottom, as the tracing for IU 165 is not reproduced.) There is also accelerated speech at the beginning of IU 167, which, incidentally, extends over most of two IUs.

The juncture following IU 167 requires more careful consideration. The speech is rapid over this stretch, and there is no pause, no laryngeal
Figure 5.5. F₀ tracings for IU 163 and IUs 167-8.
ization, no tempo shift, no reset (in fact no change in $F_0$ across the boundary can be measured). However, in Figure 5.5, bottom, it can be seen that a very large and extended pitch excursion occurs in conjunction with the accent on *kenn(e)lnert*. The pitch prominence on *kenn-* together with the maintenance of a relatively high pitch on -*lernt* produces the perception of a primary accent plus boundary tone, i.e. an IU boundary should be placed after -*lernt*. This perception is reinforced when it becomes clear that IU 168 bears a complete contour of its own, with the primary accent on *ganz*.

Between IUs 168-9 there is a pause following the rise in pitch associated with the boundary tone at the end of IU 168, effectively segmenting this stretch of speech.

IU 169 falls to low, and the pitch appears to be reset at IU 170 (see Figure 5.6 below, top; evaluation is hampered by background noise at this point). Between IUs 169-70 there is also a pause and a tempo shift.

The junctures between IUs 170-1, IUs 171-2, and IUs 172-3 are similar to that between IUs 167-8. The main feature identifying the IU boundary in each case is the configuration established by a large and extended $F_0$ excursion—the primary accent—at the end of the unit. In the first and third instances, there is also noticeable lengthening of the final syllable at the boundary.

Between IUs 173-4 there is a break for inhalation, and the final, accented segment of IU 173 has marked lengthening. Following the in
Figure 5.6. $F_0$ tracings for IU's 169-176.
breath, there is possibly a small reset.

IU 174 itself is truncated, one sign of which is the glottal stop as transcribed. The juncture with IU 176 is reinforced by the slight pause and the accelerated tempo on the repetitions of *würd‘er*.

The primary prosodic feature separating IU 176-7 is a tempo shift. There is no pause, and the pitch across the juncture is more or less continuous. The pitch drops to low following the primary accent on *reich*, and continues at a low level over IU 177. No $F_0$ tracing is available for the stretch of speech following IU 176, as the pitch and intensity are both too low, but measurements from the waveforms could be made. However, at IU 177 the speech is also accelerated so that a boundary is readily perceived.

The boundary between IU 177-8 is indicated by the final fall at the end of IU 177, with a corresponding lessening in intensity, and by the slight pause at the juncture.

At IU 178-9 and IU 179-80 there is a change of speaker, which by itself has a segmenting effect.

IU 180-1 are separated by a slight pause and a tempo shift. The pitch falls to a fairly low level (204 Hz) at the end of IU 180, but IU 181-2 continue on low pitch, so that it is the increase in speech rate on *und* following the pause that indicates the juncture.

IU 181 is truncated; again, there is no primary accent. In this case,
Elena breaks off her articulation of the IU in the middle of a word (probably due to a mispronunciation), pauses briefly, and then resumes. While the pitch changes are not significant, the rhythm of speech was interrupted severely enough for an IU boundary to be perceived at the break.

The boundary between IUs 182-3 is marked by the fall to low pitch (186 Hz) at the end of IU 182, and by the inhalation break. The pitch reset cannot be instrumentally confirmed in this case, due to the overlapping backchannel (IU 184), but it is audible. In this situation—after a final fall and inhalation—a reset is certainly expected.

The boundary between IUs 183 and 185 is indicated by the extra lengthening of the syllable bearing the primary accent at the end of IU 183, and by the slight pause.

IUs 185-6 are separated by a combination of pause and inhalation following the primary accent on Bett.

At IU 187 laughter ensues, which breaks the prosodic flow. (IU 186 is a completed unit, even though Elena starts to laugh during the articulation of Klavier.) And finally, at IUs 189 and 191 and at IUs 191-2 there is again a change of speaker.

The analysis of the conversational exchange into IUs—together with their arrangement on the page one IU at a time, however short or long—can be seen as reflecting the natural development of the discourse, bit by bit, with prosodically salient transition points prominently placed, not
buried midline. In the sample excerpt, there are a total of 45 of these IU transition points, plus 3 other possible boundary points. In the above examination of prosodic features associated with IUs and their boundaries, we have seen a range of feature configurations used in marking IU boundaries. Some boundaries are perceptually quite strong and have numerous, or particularly salient, features associated with them. Others are perceptually rather weak, with a corresponding lack of prosodic events. Most, however, fall somewhere in the middle. Not counting the transition points involving a change in speaker, the majority of boundaries in the excerpt (68%) can be associated with two or more features. The prosodic structure so outlined provides an optimal framework for subsequent analyses in other structural dimensions.

5.3 Clausal analysis of ‘Dirges’

In this section, I consider the speakers’ construction of clauses from the point of view of the preceding prosodic analysis. The initial and final boundaries of each clausal sequence are determined individually, and their location is compared in turn to the location of IU boundaries. Backchannel utterances are ignored in the syntactic analysis. Again, I begin with IU 135.

The selection opens with a syntactic ‘yes-no’ question at IU 135: the clause construction has a verb-first structure, i.e. the finite verb is the first constituent of the clause. The clause beginning is marked in the transcript by the number sign plus open curly bracket combination (‘#’). The number sign separates two clause constructions; in other words, it indicates
that the structure in IU 134 is finished and that a completely new clause construction starts at this point. The open curly bracket indicates the syntactic boundary of the current clause. At the end of IU 135 there is a close curly bracket ('\)'), marking a point of syntactic closure (but not necessarily the end of the construction). In other words, at this break in the flow of speech, the syntactic construction so far produced can be interpreted as a clause by itself. That the construction in addition contains everything Elena intended to put into the clause is indicated by the number sign at the beginning of the next IU. That is, the number sign at the beginning of IU 137 serves to separate completely the clause in IU 135 from the new construction begun at IU 137. The location of the clause boundaries being identical to the location of the prosodic boundaries, IU 135 thus consists of exactly one clause construction. The one-to-one relationship illustrated here between basic prosodic and syntactic structures is quite common, and it is the norm for syntactic questions in my data. (About 80% of clauses with this form in the corpus are contained within the scope of one IU.)

The second clause construction of the story begins at IU 137 and ends with IU 144; it contains instances of repetition and syntactic reformulation. Thus, it is an interesting construction distributed over many IUs, with internal structure at the clause level. Presumably, the speaker's formulation of the clausal constituents is made more difficult by the non-verbal action she is attending to, as evidenced by IU 143. There is also a potentially interruptive laughter by the listener at IU 138 and a cough at IU 142. And there are information flow considerations, which will be taken up in §5.4. The construction opens with a subject pronoun (er), a fi-
nite auxiliary (*will*), a particle (*so*), and an article (*den*) in IU 137. In IU 139 the article is repeated and the adjective *reichen* is added. In the next IU the noun *Amerikanern* is uttered, completing the noun phrase, whose grammatical status—dative plural—now becomes clear. In IU 141 the subject pronoun and finite auxiliary are repeated, this time in inverted order (i.e. the subject becomes post-verbal), so that the dative argument expressed in IUs 139-40 is to be reconstructed as the first constituent of the clause. Therefore, there is an open bracket placed at the beginning of IU 139. Following the break of IU 143, Elena repeats the finite auxiliary and subject pronoun in inverted order, thereby reinforcing the reformulated construction, and she completes the clause with the NP argument *Trauermärscbe* and infinitive *verkaufen*. Syntactic closure (but not necessarily the absolute end of the construction) is certain when a nonfinite verb form occurs. It is interesting to note that from a prosodic point of view, the clause is, in effect, constructed as nonsubject NP argument plus clause remainder. That is, taking just the resultant reformulation and ignoring the repetition of *will er* in IU 141, the clausal constituents fall into two parts: the preverbal NP argument (distributed over two IUs), and the rest of the clause. The significance of this construction is addressed in §5.4 when information flow is taken into account.

The third clause construction is also distributed over several IUs, and contains an instance of lexical replacement. This construction begins, much as the second one does, with subject pronoun, finite auxiliary, and adverbials in IU 145. The next two IUs contain an NP argument with the lexical replacement. The original modifying expression (*hundert*) occurs
in IU 146. IU 147 contains the repair marker (*nein*), the replacement expression (*zehntausend*), and the noun which completes the phrase (*Dollar*). In IU 148 the clause is completed with a prepositional phrase and infinitive verb form. In contrast with the second construction, there are no internal clause boundaries to be marked.

The fourth clause construction begins with the conjunction *und* at IU 149 and ends with the infinitive *schreiben* in IU 151. As the accusative object pronoun (*den*) is the first constituent of the clause, the subject pronoun (*er*) is post-verbal, i.e. it follows the finite auxiliary *würde*, which is in second position. Repetition of the dative pronoun *ihn(e)n* occurs across the IU boundary (in fact, one could argue that the IU boundary is due in part to the repetition), but this is a lexical matter that does not build new clausal structure.

The fifth clause is also constructed over the course of several IUs, beginning at IU 152 and continuing through IU 155. It begins with an adverb + subject pronoun + finite auxiliary (*also* *er* *würde*) and ends with a double infinitive (*leb[e]n müssen*). The clause contains two reformulations, neither of which, however, introduces an internal clause boundary. In the first instance, the subject pronoun *er* is repeated in IU 152 in post-verbal position. Normatively, the pronoun would follow the auxiliary if the adverb *also* is construed as the first constituent of the clause. With the pronoun in pre-verbal position, a prosodic break would be expected to intervene between adverb and pronoun. Thus, one interpretation of this construction is that Elena repaired the sequence *also* *er* *würde*—with the
pronoun pre-verbal, but without the prosodic break—by repeating the
pronoun in its normatively correct place. Alternately, there may be some
structural influence from the preceding clauses, one which has the pro-
noun pre-verbally, the other post-verbally. In any case, no clause restart is
made. In IU 154 an apparent prepositional phrase is broken off and is re-
paired in the next IU: auf so is the original expression, üh the hesitation
marker, and bei dem the repairing expression. This replacement consti-
tutes a group-level reformulation and does not affect the internal struc-
ture of the clause.

The sixth clause construction begins with the conjunction damit at
IU 156. Thus, it bears verb-late (dependent) word order, i.e. the finite verb
is expected at the end of the clause. And indeed, the verb form
kennenlerned occurs at the end of the IU. Given that a new construction
immediately follows in the next IU, the clause is coextensive with its IU.

The seventh clause construction begins with the conjunction bevor,
and so is expected to bear dependent word order as well. It ends with the
finite verb form stirbt. Again, given that a new construction immediately
follows, the clause is thus coextensive with IU 157.

The construction of the eighth clause, like a number of the previous
clauses, extends over more than one IU, from IU 158 to the end of IU 161.
It is, however, the most complex construction up to this point, containing
instances of clausal embedding and addition. The clause starts off with a
typical sequence. IU 158 contains what appears to be a complete clause-
level unit, with an NP_object+ V + NP_subject structure. (Thus, a point of syn-
tactic closure is marked at the end of the IU.) However, from what follows in subsequent IUs, it becomes clear that Elena has not fully completed the construction, but builds upon the existing structure. Skipping over IU 160 for the moment, let us look at IU 161. The expression in this IU does not constitute a clause by itself: it consists only of an adverb, a prepositional phrase (albeit a complex one), and a particle. Rather, it adds to the structure of a clause which has not yet been definitively closed. (Semantically, it modifies the scope of the intended referent in IU 160.) In other words, there is no new clause or clause restart (thus no new open bracket) with this expression, but only the construction of additional syntax for an existing clause. Following the addition, however, another point of syntactic closure is reached, thus a close bracket is in order.

Returning now to IU 160, we see that the sequence in this IU begins with the conjunction bevor and ends with the finite verb form stirbt. Thus, it constitutes a dependent clause itself; this is the ninth clause construction. In this case, however, it is not located at the periphery of other clauses (as are the dependent clauses in IU 156 and IU 157), but rather it falls within the boundaries of clause eight. It is thus sequentially embedded, and is so marked by the double brackets.

In IU 161 there is another type of embedding illustrated. Following the noun Leute, a dependent (relative) clause beginning with the pronoun die and ending with the finite verb form sind has been constructed. This is the tenth clause construction of the story. The noun + relative clause sequence constitutes the NP argument of the prepositional phrase in the
IU, and so is syntactically embedded. This status is indicated by the double bracketing.

Clause construction eleven is distributed over three IUs, and it stands apart from other construction as it lacks a verbal expression. Consequently, it is a 'clause equivalent' structure, consisting primarily of nominal expressions. In IU 163 there is the particle also and a noun phrase. IUs 164-5 contain several instances of the particle so, an instance of the particle ne, a noun, and the phrase oder so. The expression in IU 163 is bracketed as a separate unit, with the following two IUs bracketed as clause additions. IUs 163-5 are taken to constitute a clause by themselves, as there is no existing structure which they can be sensibly integrated into. Unlike the prepositional phrase in IU 161 which can be syntactically (and semantically) linked to the preceding constructions, the nominals in IUs 163-5 cannot. The reason for this has to do with the structure of the story. Constructions 7-10 constitute a coherent digression from the main story line. The expressions in IUs 163-5 do not belong to this digression. Instead, they function in part to help return to the point where the speaker left off.

The twelfth clause construction at IU 167 begins with the conjunctions und damit. It has dependent constituent order, ending with the finite verb form kenn[e]nternt, and is coextensive with the IU.

Clause construction thirteen, which begins at IU 168, extends over two IUs. The beginning of the clause is indicated, like many above, by a sequence of conjunction + nominal or adverbial constituent + finite auxil-
iary. Its end boundary is signaled, again like many of the above clauses, by an infinitive verb form.

Clause fourteen is constructed over the course of four IUs, and exhibits a type of clause addition. It begins at IU 170 with the familiar sequence of conjunction + nominal or adverbial + finite auxiliary, and a point of syntactic closure is reached at the end of the same IU with the infinitive inszenieren. The next three IUs, however, add incrementally to this clause. IU 171 provides a prepositional phrase argument, IU 172 a pair of descriptive adverbs, and IU 173 a generalizing expression. After each addition, a point of syntactic closure is reached, which is marked in the transcript. Only with the commencement of the next clause in IU 174 is it clear that the speaker intends to add no more to the construction begun at IU 170.

Clauses fifteen and sixteen form a complex construction, beginning at IU 174 and ending with IU 177. IUs 174 and 176 constitute one clause, which starts with the conjunction und and ends with the double infinitive werden wollen. There is a repetition of the finite auxiliary + subject pronoun würden [er a]t the beginning of the second IU. This clause is syntactically embedded in the following clause, functioning as the first constituent of the latter, as evidenced by the position of the finite auxiliary hat in IU 177. The clause expressed in IUs 174 and 176 is thus marked with double brackets.

For the rest of the story, from IU 178 to IU 192, the clause structure is quite simple and does not require detailed commentary. The remaining ten constructions (17-26) are each coextensive with an IU, except for con-
struction twenty in IUs 181-2, which contains a lexical correction. Moreover, there are no clause restarts or additions, and the syntax is for the most part unremarkable. Two points are worth mentioning, however. One interesting phenomenon is the maximal attenuation of the prefield in IU 179 (here, 'object ellipsis'), which I treat in detail elsewhere (Schuetze-Coburn 1989, in progress b). Secondly, attention should be drawn to the instance of subject ellipsis in IU 186. The important point here is that the lack of an overt subject in a clausal sequence does not preclude its treatment as a separate clause construction.

In the sample of ordinary conversational discourse, we have seen evidence for twenty-six clauses. The syntactic analysis of Chapter 3 identifies a variety of clause construction types. The prosodic length of the clauses varies from less than 1 IU (the embedded relative clause in IU 161) up to 6 IUs (the construction beginning at IU 139). The fact that a number of clauses are produced over the course of more than one IU highlights the incremental production of syntax in discourse. Verb-first, verb-second, and verb-late clauses are all represented in the excerpt, as are simple, complex, and embedded clause types. There are numerous instances of lexical repetition or replacement, and some syntactic reformulation is made. Finally, there are interesting cases of clause addition present in the sample, also illustrating that clause construction is a dynamic process.

5.4 Discourse-pragmatic analysis of 'Dirges'

In this section I examine the discourse pragmatics of referents in the sample excerpt as laid out in Chapter 4, and I discuss the distribution of
these referents in terms of the prosodic and syntactic analyses provided in the previous two sections. My focus will be on mentions which lead to referential tracking in the story and on the concepts that underlie them. Roughly speaking, this means I will examine only those referents which have a central role in the story, and which the speakers thus make repeated reference to. Expressions such as *[ei]n halbes Jahr* 'half a year' in IU 153 or *aus [de]m Bett* 'out of bed' in IU 185, which occur only once, and which are in some sense incidental elements—they are more likely to refer to concepts that form part of the predication than individuated entities—are not discussed here.

Apart from the two speech event participants themselves, who are, of course, both referred to throughout the larger conversation, three referents play an extended role in the story: Rüdiger, rich Americans, and dirges. From the point of view of 'what the story is about', these three referents clearly constitute the major subjects of the story, which concerns Rüdiger writing dirges for rich Americans as a means of becoming rich himself. I begin now, as before, with the first IU of the story.

IU 135 contains an instance of the second person singular familiar subject pronoun *du* 'you' (here in a highly attenuated form 'e [ə]'). In most similar situations, because the mention is phonologically almost fully reduced (the next step would be complete elision), and because the additional property of referents, following Du Bois & Thompson (1989), will be useful in the analysis of 'Dirges'. Individualized referents are said to be 'particular'; nonindividualized referents are 'general'. Generic and nonspecific reference is subsumed under the 'general' label.

---

5 One additional property of referents, following Du Bois & Thompson (1989), will be useful in the analysis of 'Dirges'. Individualized referents are said to be 'particular'; nonindividualized referents are 'general'. Generic and nonspecific reference is subsumed under the 'general' label.
dressee clearly constitutes an 'identifiable' and 'particular' entity in the speech context, the concept/idea of the addressee, to which the mention would refer, would have to be taken to have the status 'active'. In this particular case, however, there are additional considerations. A second possibility is that the reference is not 'particular', but 'general'; that is, it may have the sense of *man* 'one', much like the similar use of *you* in English. A third possibility is in this case the most likely. The clause *weißt *'e was is most certainly formulaic (cf. Pawley 1986), and thus it is less directly associated with any of the speech event participants—or even with a general referent 'one'—than its component grammatical elements would imply. In other words, 'e may not be intended to refer at all. If the mention is nonreferential, then it need not be further discussed here, as I am focusing on tracking mentions.

The next mention to be examined is the subject pronoun *er* 'he' of clause two in IU 137, which refers to the concept/idea of 'Rüdiger'. As the human protagonist of the story who is well known to the listener, this referent is both 'identifiable' and 'particular'. Since reference is achieved with a pronominal form without accent, and prior reference was made in IU 134, the concept/idea of 'Rüdiger' must be already 'active' at this point. It is useful to observe that the referential tracking of 'Rüdiger' that extends through the present story begins with the mention *Rüdiger* in IU 132, and that the last mention of him prior to this mention took place in IU 111, twelve clauses removed. Seeing that reference to 'Rüdiger' in IU 132 is made with a proper noun that bears an accent, we may surmise that just before its expression there, the concept/idea probably had been

393
displaced from consciousness, leaving it in a ‘semiactive’ state. In subsequent references to ‘Rüdiger’, we need be concerned only with the activation state of this concept/idea. The semantic status of the referent (human, individuated), his role as agent/actor, and his prior establishment in the story mean that only exceptional mentions will be anything other than ‘identifiable’ and ‘particular’.

Clause two expresses another concept/idea that finds subsequent mention as the story unfolds: ‘rich American’. Several properties of this mention are interesting with regard to information flow. First of all, there is the prospect to be considered here that the activation state of the concept and the idea differ prior to the mention. The idea of the rich Americans, i.e. the referent is clearly ‘inactive’. It has not been expressed before, and nothing in the context warrants assigning a ‘semiactive’ status. On the other hand, the (complex) concept «RICH AMERICAN» may be semiactive. The concepts «AMERICA» and «AMERICAN» have already been expressed earlier in the conversation (but not in the transcribed excerpt), and may still be ‘semiactive’. The concept «RICH» is clearly still ‘active’, having been just expressed in IU 134. (Furthermore, it serves arguably as the catalyst for the ensuing story at this particular point in the conversation.) If the activation state of the individual concepts can be shown to contribute to the concept as a whole, then «RICH AMERICAN» may not be ‘inactive’ here. Still, the concept expressed in IUs 137-40, «RICH AMERICAN», is a composite one which is conveyed only by this mention. At the least, this complex concept presupposes knowledge that is not present in a simple concatenation of concepts (see below), and on this basis it might be concluded that the
concept expresses new information. This is an issue that needs more at-
tention. In any case, the expression *den reichen Amerikanern* '[to] the
rich Americans', which is a dative argument of the transitive verb
*verkaufen* 'sell', is used to introduce the idea of 'rich Americans' into the
story.

The second interesting thing about the expression is that it is dis-
tributed over more than one IU. In fact, the 'verbalization' process ap-
ppears to be sufficiently challenging here so that it is distributed over three
IUs. In IU 137 the definite article *den* occurs. In IU 139 the article is re-
peated, along with the adjective *reichen*, which, as mentioned in §5.2,
bears a primary accent. The noun *Amerikanern*, completing the mention,
occurs in IU 140 (it also bears a primary accent). From the point of view of
the prosody, it is as if a concept «RICH» is being introduced separately from
a concept «AMERICAN» (even though each concept itself is already 'active'
or 'semiactive'). The cognitive challenge in this case might have less to do
with the management of activation state and more to do with the intro-
duction of a new referent. There are other, extralinguistic factors in play
here as well: the intervening laughter in IU 138, the cough in IU 142, and
the long break in IU 143. Moreover, a syntactic reformulation is taking
place in this sequence of IUs. As explained in §5.3, the dative expression
starts out as a postverbal argument, but ends up as the preverbal con-
stituent of the clause when Elena repeats the auxiliary and subject pro-
noun in IU 141 and in IU 144. There is so much happening at once here
that it is impossible to assess to what extent cognitive vs. interactional
considerations have determined the way the referent is expressed, but
each is likely to have exerted some influence.

The third interesting property of the mention is the definite article. Under one interpretation, the use of the definite article implies that the referent is ‘identifiable’ (Hentschel & Weydt 1990: 205-6; cf. Du Bois 1980: 206 for English), and hence ‘particular’. But this is unlikely here. Although the concept «RICH» and the idea of ‘Americans’ have been referred to in the prior conversation, the separate mentions associated with these concepts/ideas could not by themselves allow the listener to identify any relevant set of individuals. A second interpretation is that the speaker fully intended to use the definite article in a meaningful way; that is, she is purposefully presenting the referent as ‘identifiable’. Although the speaker may be unsure whether the listener can adequately identify the intended referent when introduced in this way, she can exploit the grammatical contrasts of the language to attempt the reference, and hope that the listener can make the necessary inferences. The speaker does have ‘... facultative control over definiteness’ (Du Bois 1980: 219), and while it is playing fast and loose with principles of conversational interaction to do so (e.g. Grice 1975), the speaker—by forcing the listener to exert a little more cognitive effort than usual in constructing a working hypothesis—is making efficient use of available resources. This interpretation implies that the referent is both ‘identifiable’ and ‘particular’. Still, there are other possibilities. Because we are analyzing discourse data, we must be willing to accept the eventuality that the speaker erred in her introduction of the referent, and that given a chance, a ‘correction’ would be made. There is some evidence that Elena was not sure Sabine was follow-
ing her. Later, in clause eleven, Elena provides additional characteriza-
tion of the intended referent. Locally, however, only a repetition of the
definite article is made, not a lexical replacement. Thus, while speaker er-
rors do happen, evoking this interpretation should remain a last resort,
especially in the absence of fully compelling evidence.

Yet a fourth interpretation is available. In light of the clarification in
clause eleven just mentioned, it is quite possible that reference is to the
class of all rich Americans as a whole, i.e. the referent is ‘general’, not
‘particular’. Unlike most similar expressions in English, German permits
the use of definite plural noun phrases for general reference (Helbig &
Buscha 1987: 375; Eisenberg 1989: 164). The question of pragmatic status
would then be limited to one of relevance: the speaker need only assume
that the listener knows the reference demarcates a meaningful set of in-
dividuals, not any specific one. Given that phrases such as der reiche
Onkel aus Amerika ‘the rich uncle from America’ are stereotypic for
Germans, this assumption can easily be made.

Finally, the fourth interesting property of the mention under review,
which I will comment on here only in passing, is that its grammatical
form is a dative nominal. ‘Inactive’ concepts/referents are not commonly
expressed using dative arguments in spontaneous speech. As I pointed
out in §4.4, 94% of dative arguments are either ‘active’ or ‘semiactive’ in
the primary corpus (cf. Schuetze-Coburn 1987b, where 100% of dative ar-
guments were found to be ‘active’ or ‘semiactive’). While this statistic
could be used in arguing for a ‘semiactive’ status of the complex concept
«RICH AMERICAN», it can also be used to illustrate the tremendous variation that does occur in natural discourse, if but rarely.

The next mention to be discussed is the reference to the concept/idea 'dirges' in IU 144, also in clause two. As the concept/idea is newly introduced here, its activation state prior to its being expressed is 'inactive'. The expression Trauermärsche 'dirges' occurs as the accusative argument of the verb verkaufen 'sell' (of the same clause as the argument den reichen Amerikanern just discussed), is grammatically indefinite, and bears the primary accent of the IU. As such, it compares favorably to typical patterns for the introduction of new referents. Still, there are a few interesting complications to deal with. Somewhat surprising is the occurrence of both 'rich Americans' and 'dirges'—two 'inactive' referents—in the same clause. (This is, in fact, the last of three exceptional cases to the 'One New Argument Constraint' discussed in §4.5.3.) Significantly, however, the two referents are expressed in separate IUs. Indeed, the portion of the clause containing Trauermärsche is uttered after a 1.3 second break, so that there is by all accounts ample processing time for 'verbalizing' both referents. More difficult to reconcile is the occurrence of the 'inactive' concept «SELL» in the same IU as 'dirges'. Explaining this double occurrence is of special theoretical importance in view of the 'One New Idea Constraint' discussed in §4.5.4, as it appears to constitute an exceptional case. As pointed out in §4.3.1, object + verb sequences frequently function as lexicalized units, and as such express unitary (but complex) concepts. Yet while Trauermärsche verkaufen 'sell dirges' matches this grammatical pattern, the match is lexically and semantically less felici-
tous. There is no evidence that *Trauermärche* has lost any of its nominal properties in this collocation, and neither the object nor the verb is a semantically ‘bleached’ item. Nevertheless, this expression may provide an example of ‘pragmatic incorporation’. A clue to the status of this expression is to be found in the next clause. In IU 148 there is a subsequent reference to the concept «DIRGE» in the expression *für einen Trauermarsch* ‘for a dirge’. The mention is also grammatically indefinite, and this form would be expected following a prior mention only if a referent associated with the concept «DIRGE» were yet to be established (Du Bois 1980: 210). In other words, we can conclude that the first mention in IU 144 is ‘nonreferential’, the status associated with noun phrases in ‘object + verb’ collocations expressing a unitary (complex) concept, the concept being in this case «SELL DIRGES».

Clause three, articulated in IUs 145-8, contains a mention referring to ‘Rüdiger’ (the subject pronoun *er*), and a mention referring to ‘dirge’ (the prepositional object *einen Trauermarsch*). Having just been mentioned, there is nothing to add here to what has already been said about the protagonist, save to state that the concept/idea ‘Rüdiger’ is still ‘active’. An observation about the other mention, however, is in order. As noted above, the expression *für einen Trauermarsch* ‘for a dirge’ is grammatically indefinite. One interpretation of this mention is that it establishes a referent which then may be subsequently tracked in the story. The prior mention of ‘dirge’, while activating the concept, was nonreferential, and thus no referent was introduced. The mention in this clause provides a proper introduction: the first mention of the referent is presented as ‘non-
identifiable'. The referent is in this case also 'general', as no particular instance of the concept «DIRGE» is intended.

Clause four, starting at IU 149, contains mentions of all three tracked referents, all in pronominal form. The third person masculine accusative object pronoun den refers to 'dirge', and is the first constituent of the clause. This preverbal clause position is typical for a d- pronoun when it refers to a concept/idea expressed in the immediately preceding IU. The concept/idea 'dirge' is, as in the previous IU, already 'active', and the expression is noticeably attenuated (written d'n in the transcript). In a postverbal position, the subject pronoun er 'he', referring to 'Rüdiger', and the dative object pronoun ihnen 'for them', referring to 'rich Americans' follow. In both these instances, the most recent prior mention does not take place in the preceding IU, but a bit earlier. Reference to 'Rüdiger' occurs at IU 145, and to 'rich Americans' at IU 138. In terms of clauses, the referential distance is not so great: 'Rüdiger' is expressed as the subject of the previous clause; 'rich Americans' is expressed as an argument of the clause before that. Thus, although the pronominal mentions are not attenuated in this clause—indeed, ihnen bears the primary accent of the IU—the concepts/ideas 'Rüdiger' and 'rich Americans' may be taken to still be 'active'. Based on the evaluation of the discourse pragmatics in clause two, the use of ihnen constitutes a general reference. It is noteworthy that IU 149 of clause four contains all three mentions itself, but since all references are to 'active' concepts/ideas, there is no processing difficulty (or 'information bind') to be concerned with. The other IU forming the clause contains two repetitions of ihnen, which, however, do not dif-
fer in information status. Moreover, they constitute, under the view espoused in Chapter 4, a single act of referring. (In any case, there is but one dative argument in the clause.) Unlike the situation with den reichen Amerikaner in clause two, we cannot appeal to the activation state of the concepts/ideas to explain why the constituents of this clause are distributed over two IUs. The two instances of ihnen in IU 151 are merely repetitions, not repairs in the sense of correcting or reformulating, so the source of the break is not to be found in the properties of the mentions. The overlapping backchannel (IU 150) may be a factor. But this utterance is weakly articulated, and so in no way competes for the ‘acoustic space’ of the exchange. Another possibility is that Elena simply needed more time to plan the clause. The new information of the clause (ganz speziell schreiben ‘write specially’) is yet to be expressed; and while the verbal bracket constructed in clause four often requires knowledge of the (clause-final) lexical verb in order to ensure the correct form of the (clause-second) auxiliary (Schuetze-Coburn 1993), this is not the case with the auxiliary würde. That is, there is no constraint involved in the selection of the lexical verb in this instance, so that it is indeed grammatically possible for the second half of the clause to be planned after the first half has been articulated.

The four IUs of the fifth clause contain references to ‘Rüdiger’ and, seemingly, to ‘a rich American’. The two instances of the subject pronoun er ‘he’ in IU 151 refer to ‘Rüdiger’ (the repetition has already been discussed in §5.3), who thus continues as grammatical subject. The concept/idea ‘Rüdiger’ is maintained ‘active’. The dative singular pronoun
dem in the phrase bei dem ‘at his [house]’ in IU 154 presents a bit of a puzzle. The reference is, semantically, to an individual of the class «RICH AMERICAN». What Elena is explaining is that Rüdiger would have to live with a rich American in order to get to know him well enough to write a personal dirge for him. Up to this point—and, more importantly, subsequent to it—reference to ‘rich Americans’ is achieved with grammatically plural forms; the mention dem is the only reference using a singular form. A possible interpretation is that Elena shifts—momentarily—from a ‘general’ to a ‘particular’ reference. Speakers do make this shift, without any apparent difficulty. But the shift back occurs immediately, and one may wonder whether this is the correct interpretation. There is, however, another plural-to-singular-to-plural shift in IUs 157-61 (which will be brought up in the context of clause eight), so that the possibility should be entertained. An alternate, more likely interpretation is available, though. It is quite plausible that the reference is indeed plural, but that this fact is obscured by a speech error affecting the realization of the pronoun. That is, what is pronounced [dem] could be phonologically /denən/, the dative plural pronoun. (The pronoun denen would likely be attenuated to [deni] or [den], or at least to [denn], in context here anyway, so that the weak nasal at the end of /denən/ would be vulnerable to further change.) Because of the strong presence of bilabials in the double infinitive leb’ n müs sen ‘have to live’ at the end of IU 155, which is realized as [lebm mysn], the pronunciation [dem] could well have resulted from the articulatory anticipation of the bilabials. Given these two possibilities, we cannot be certain what the referent is, even if the underlying concept is clear.
(and 'active') enough.

The pronominal mentions in the next unit, IU 156, which constitutes the sixth clause, refer to 'Rüdiger' (the subject pronoun er) and to 'rich Americans' (the accusative object pronoun sie), maintaining their 'active' status. Both mentions are attenuated and without accent.

As explained in §5.3, clauses seven through ten present a short digression from the main story line. IU 157 contains the nominal expression die Leute 'the people', which is the grammatical subject of clause seven. There is more than one possibility in interpreting the intended referent of this mention. On the one hand, 'rich Americans' could be meant, taking the article die in the sense of diese 'these'. On the other hand, the expression could constitute a general reference to 'people'. This latter interpretation is bolstered somewhat by subsequent mentions in clauses eight and nine. In IU 160 (clause nine) the general pronoun man 'one' is used to reiterate the proposition expressed in IU 157 with die Leute. That is, Elena is clarifying that the proposition is intended to hold of anyone generally, not just rich Americans. The expression Leuten die so länger krank sind 'people who have been sick a long time' in IU 161 is used to qualify this generalization, but again, it demonstrates that a broader class of individuals was meant.

The phrase nur ganz ganz stinkreiche Amerikaner 'only really really filthy rich Americans' in IU 163 can be associated with the concept/idea 'rich Americans' previously expressed unequivocally in IU 156. Above, I suggested that this mention, and the one that follows it, constitutes an
added characterization of the referent already established. In other words, it does not seem to be the case that a different referent is intended, only that the original referent should be clarified. To be noted is the indefiniteness of the expression. This property may be interpreted two ways. The reference may be general, as in the previous mentions. Alternately, the expression may be functioning much as a predicate nominal would (but here without the copula). In this case, the mention is simply nonreferential. The expression *Industrielle* ‘industrialists’ in IU 164, which closes clause eleven, functions the same way. It provides an additional refinement—perhaps the prototypical exemplar—of the general class intended by the concept «RICHER AMERICAN».

IU 167, which forms clause twelve, resumes pronominal reference to ‘Rüdiger’, and to ‘rich Americans’, after the brief digression. In fact, the clause is a near repetition of clause six in IU 156. Because four clauses comprising seven IUs intervene between mentions of ‘Rüdiger’, we with good cause may ask whether the concept/idea is still ‘active’, or whether it has receded to a ‘semiactive’ state. One indication of the activation state of a concept/idea prior to its mention is the degree of attenuation of the mention itself (although, as we saw in §4.3.2, this is a delicate issue). Since IU 167 is more or less a repetition of IU 156, where it has been established that both concepts/ideas are ‘active’, we have an opportunity to compare expressions directly. If the articulation of *er* and *die* in IU 167 is less attenuated that the articulation of *er* and *sie* in IU 156, then we would have good reason to pursue this inference further. As it turns out, the mentions *er* and *sie* in IU 167 are, if anything, more attenuated than their

404
counterparts in IU 156. The duration of IU 156 (excluding the initial pause and inhalation) is 1.12 seconds, with the sequence *er sie* lasting 0.11 second. IU 167 (minus the inhalation and conjunction *und*) has a duration of 0.99 second, with *er die* approximately 0.09 second long. Proportionately, this is about equivalent. It can be noted, however, that auditorily the vowels of the mentions in IU 167 sound a little more centralized. Be that as it may, it is certainly not the case that the mentions in IU 167 are fuller forms. To the extent that form attenuation is an index of activation state, the phonetic evidence for a ‘semiactive’ status is negative. Nevertheless, the repetition of the clause may itself be indicative. What we can be certain of is that the mentions in the following IU refer to concepts/ideas that are ‘active’.

In IU 168 reference is made again to ‘Rüdiger’ and ‘rich Americans’, both concepts/ideas being ‘active’ here, as just noted. Clause thirteen also contains a mention of the concept/idea ‘dirge’, expressed as the accusative object *[eine]n Trauermarsch ‘a dirge’ in IU 169. In this instance, we can surmise that the concept/idea ‘dirge’ had been displaced to a ‘semiactive’ state sometime before this mention. The last previous mention was in IU 149, and the mention here is with a full noun phrase, notably in an IU separate from the other nominal expressions in the clause. The fact that the mention here is again indefinite (cf. IU 148) can be attributed, under one interpretation, to the use of the noun phrase in an object + predicate construction to express a complex concept «WRITE A DIRGE», much as in IU 144, without the intention of referring to an individualized object. In other words, this would be a nonreferential, predating mention.
Clause fourteen contains ‘active’ mentions of ‘dirge’ (the accusative object pronoun den) and ‘Rüdiger’ (the subject pronoun er), which continue to track the primary characters.

Clauses fifteen and sixteen contain four mentions of the ‘active’ concept/idea ‘Rüdiger’ (the pronoun er), which constitute the grammatical subjects of the two clauses. The distribution of constituents over IUs 174 and 176 is reminiscent of the way the dative argument ihnen was split over IUs 149 and 151 following a backchannel utterance. In IU 177 there is also reference to ‘Elena’, through the dative pronoun mir. It is unaccented, and so in all likelihood refers to an ‘active’ concept/idea.

The subject of clause eighteen (IU 179) is the first person pronoun ich, referring to ‘Simone’. It is not accented in this instance, and thus it is most probably ‘active’, like mir in the preceding IU.

Clause nineteen (IU 180) contains both a mention of ‘Rüdiger’ (the subject pronoun er) and of ‘dirges’ (the plural noun phrase ganz schöne Trauermärche ‘very nice dirges’). The latter expression is indefinite. As in IU 169 it can be taken to form part of the predicate, even though in this case the infinitive schreiben ‘write’ must be inferred. That is, the mention is nonreferential here, too.

Clauses twenty and twenty-two through twenty-five all contain references to ‘Rüdiger’. In the first four of these clauses, the mentions are expressed as grammatical subjects, in IUs 180 and 185 with er, and in IU 189 with der. In IU 186 (clause twenty-three) the mention is achieved by
means of maximal attenuation of the prefield (cf. §4.3.1). In clause twenty-five the reference is by means of the dative object of the verb *einfallen* 'occur to'. Throughout this sequence, the concept/idea ‘Rüdiger’ is ‘active’.

The passage ends with a self-reference to ‘Simone’ in IU 192, using the first person singular pronoun *ich* as the subject of the transitive verb *finden* ‘find’. As in IU 179, the mention is unaccented, and thus probably ‘active’.

Rüdiger is the main protagonist; he is consistently the agent or actor of events in the story. Consequently, in the majority of clauses (16/26) he is tracked with attenuated mentions, i.e. pronominally or through maximal attenuation (ellipsis). The concept/idea ‘Rüdiger’ is predictably ‘active’ throughout. The other two concepts/idea that occur repeatedly, ‘rich Americans’ and ‘dirge’, present a markedly different picture. While they are also mostly ‘active’ after their initial introduction, some mentions are to general referents, some are nonreferential. In these instances, mention is made with a full nominal expressions.

5.5 Summary

The above exposition sheds light on the way language is used by one speaker to communicate a story to another person as a means of exchanging ideas about the speaker’s world view. In terms of the interlocutors’ knowledge, the sharing of experiences broadens the common ground on which further exchanges can be based, and so contributes to the mainte-
nance of interpersonal relationships. Rendering the conversation into IU's is a first step in a linguistic understanding of the structures underlying this sharing of experiences. It is seen that the story told in the excerpt does not in general unfold clause by clause (or sentence by sentence); rather, the telling is richer and more varied than this, reflecting in part the cognitive effort that goes into recalling and 'verbalizing past experience' (Chafe 1977b). Examining the discourse pragmatics in conjunction with the prosodic and syntactic structures thus provides us with an approximation of some major linguistic processes at work in the on-line production of language. There is still a wealth of additional observations that could be made about this 55-second discourse excerpt that I have left untouched.
6. Summary and Prospects

In einem Satz läßt sich (höchstens) eine Tatsache mitteilen, in einem Satz läßt sich (höchstens) ein Erlebnis fassen, in einem Satz kann nur ein Vorgang dargestellt werden, nur ein Gegenstand oder eine Seite eines Gegenstandes beschrieben werden, in einem Satz nur eine gedankliche Synthese sich vollziehen.

In a single clause one fact (at most) can be communicated; in a single clause one experience (at most) can be expressed; in a single clause only one event can be represented, only one object or one side of an object can be described; in a single clause only one mental synthesis occurs.

Hermann Ammann

Die menschliche Rede II: Der Satz, 1928, p.93

In the past three chapters, I have dealt with numerous methodological and theoretical issues in three different domains of discourse—prosody, syntax, and information flow—and I have examined their interplay in the example of natural, unscripted, spontaneous, connected German conversation. These investigations have covered much ground, and the many observations made have not yet been fully connected to yield an overall picture. Thus, in this chapter it remains to review my most important findings and to discuss their interconnection, also at the same time offering some thoughts as to their possible significance for a theory of grammar and some suggestions for future research.

In §1.3 I framed my goals for this dissertation in terms of two avenues of investigation. First of all, I stated that numerous fundamental questions had to be tackled at the outset. I argued that for the most part we currently lack the necessary analytic tools for sophisticated and sensitive
linguistic investigations of ordinary conversational discourse. The study of discourse is still very much developing: we are just now producing the incunabula of a field that surely has yet to make its major contributions to our understanding of language. In such a situation, it seemed prudent to closely examine working assumptions and carefully weigh all analytic options.

With regard, then, to the methodological concerns, the following major points were made. First and foremost, I argued that the issue of the segmentation of discourse had to be addressed. The utterances of our corpora—the undifferentiated sequences of words produced by the speakers in our recordings—cannot be gainfully analyzed until they are broken into manageable segments. In §2.2 I presented available evidence showing that the segment appropriate for this purpose is a type of prosodic phrase, a unit based on specific cues in the prosody of spontaneous speech that has significance to both speaker and listener. I advocated the use of one prosodic phrase in particular, the ‘intonation unit’ (IU), and in §2.3 I discussed the actual cues used in its delimitation in some detail. I then proposed that a prototype representation of the intonation unit helps to address some of the problems that have been associated with units of this type.

I argued further that such intonation units used to segment recorded utterances not only reflect the prosodic organization of the discourse, but organization at other levels as well, in particular, in a syntactic dimension. In §3.1 I presented evidence from discourse that in the first instance
a syntactic analysis should be sensitive to clausal structure and should employ some notion of 'clause'. In light of the prosodic organization, I proposed in §3.3 the developed of 'clause construction', with which I sought to represent clauses as they are formed in discourse by the speaker (and sometimes the listener as well), intonation unit by intonation unit, thereby capturing the dynamic nature of this process.

In the context of a discussion of discourse pragmatics, I critically examined Chafe's conception of information flow in §4.3, discussing the properties 'activation state' and 'identifiability'. I argued that a more sophisticated treatment of the mental entities that are associated with information flow statuses leads to a more thorough account of observed linguistic phenomena. Specifically, I proposed that the notions 'concept' (approximately, a mental representation of a class) and 'referent' (approximately, a mental representation of an individual) be distinguished, and that their contribution to the activation status of mental entities be evaluated separately.

If the responsibility of the linguist for all aspects of his or her data should be complete, then the appropriate tools for meeting this responsibility must be developed. In reviewing fundamental questions of prosody, syntax, and pragmatics in the context of providing a structural description of ordinary language use, I therefore articulated a theoretical framework which will allow the linguist to achieve a 'whole-text' analysis of extended stretches of discourse.

To demonstrate the application of the linguistic methods I developed
for investigating basic structures in discourse, I submitted in Chapter 5 an extended excerpt from one of the conversational stories in my corpus of German discourse to a detailed prosodic, syntactic, and discourse-pragmatic analysis from my framework. I stepped through the excerpt intonation unit by intonation unit, clause construction by clause construction, and discourse referent by discourse referent, explaining the relevance of individual features present for the determination of the resulting structures. This exercise successfully illustrated a comprehensive ‘whole-text’ analysis.

The second avenue of investigation that I undertook consisted of a progressive tabulation of features of my German conversational corpus, and a comparative discussion of those features in the three dimensions of analysis. Specifically, in §2.4 I gave a characterization of the intonation units resulting from the prosodic analysis of my entire corpus. In §3.4 I outlined properties of the clause constructions in the corpus, and in §3.5 I compared the occurrence of these syntactic units with respect to the IUs. In §4.4 I compared the distribution of new referent mentions with the prosodic and syntactic structures.

Finally, in §4.5 I examined the interaction of the three dimensions by evaluating four hypotheses concerning the limited occurrence (or nonoccurrence) of certain configurations of prosody, grammar, and information flow in discourse. Specifically, Du Bois’ ‘Given A Constraint’ and ‘One New Argument Constraint’ were tested against the German conversational discourse in my corpus. Du Bois’ first constraint was strongly sup-
ported. Less than 1% of the transitive subjects in the corpus expressed new referents. The second constraint was also strongly supported. Again, only 1% of the clause constructions containing two or more core (i.e. case-marked) grammatical arguments expressed more than one new referent. Chafe’s ‘Light Subject Constraint’ and ‘One New Idea Constraint’ were also evaluated. Chafe’s first constraint was also supported. Just 3% of all grammatical subjects in the corpus were ‘heavy’, i.e. they expressed new referents that were tracked through the discourse. As this constraint is complex (there are multiple criteria), it may be the case that the slightly higher number of exceptional cases is a result of the interplay of the factors that the constraint embodies. The second constraint was also supported. Only 2% of the intonation units that contained one new referent also expressed another new concept. The complications that arise in the determination of the concept structure of the IUs, however, suggest that this topic needs further scrutiny, possibly under well-designed experimental conditions that respect the interactional parameters of discourse, as well as from other conversational corpora.

The specific patterns that surfaced from the tabulations should permit some general observations to be made at this point. Consider first a broad implication for sociolinguistic studies using discourse as primary data. The basic IU analysis of my corpus confirmed not only the relevance of these units for language production from a cognitive perspective, but from an interactional perspective, too. In my German conversations, where the number of speakers ranged from two to five, I found that the turn structure of the talk as I represented it meshed well with the
prosodic structure. In many respects, this result is not remarkable—indeed, it would be most remarkable if verbal interaction was not sensitive to a fundamental unit of language production. However, the synchronicity does serve to highlight the grave importance of prosodic structure to any study of conversational interaction, for while turn units and IUs mesh, they are not equivalent. IUs provide a much more fine-grained structure, the significance of which to interaction has yet to be explored.

Consider, too, the relationships between prosody and other dimensions that were made evident. Since my analysis of the corpus of conversational discourse began with a segmentation of the utterances into intonation units, and my other analyses proceeded from that, the presumption was made that syntax and information flow are dependent on prosodic structure. While there are valid methodological considerations for beginning with discourse prosody—prosody by nature is temporally grounded, and an orientation to it helps one to maintain a focus on the dynamic aspects of other language structures—it would nevertheless be deceptively simple to imply that the prosodic structure of discourse determines the others. Clearly, prosodic structure itself is constrained in certain ways. From a physiological viewpoint, humans can talk in much longer stretches than they ever do in spontaneous speech, for example. The conclusion to be drawn, then, is that the observed structures in the various dimensions are mutually constraining. We can then in a way come full circle and make an observation linking aspects of the prosodic dimension directly with that of information flow. The intonation unit may be the proper domain for information flow, but the cognitive pro-
cessing constraints underlying information flow also limit the extent of intonation units.

Where does that leave syntax? There is, finally, an observation to be made regarding the specific pressures that prosody and information flow together can exert on constraining syntax. As we have already seen in §4.5 when evaluating hypotheses concerning the placement of new referents, strong trends are found in discourse which exactly match morphosyntactic patterns. The clausal patterns outlined in §3.4 proved to be more variable; nonetheless, the fact remains that a substantial majority of clause constructions (showing no internal complexity) are coextensive with intonation units. Majority syntax is quite simple in its construction. This fact is consonant with the observation that there is a tight limit as to what can be spontaneously expressed at any given moment, both prosodically and informationally. The complexity of the syntax of an utterance, then, appears to be variable only within the confines of domains that are themselves cognitively highly constrained. Hermann Ammann expressed the idea of these mutually constraining limits in his own way some time ago. The details have yet to be worked out in many respects, but we can be certain that further exploration in this direction will undoubtedly contribute much to functional explanations for the form of grammar.
REFERENCES

Aarts, Jan & Willem Meijs. 1984 (eds.) Corpus Linguistics: Recent developments in the use of computer corpora in English language research. Amsterdam: Rodopi.


Allen, George D. 1972. The location of rhythmic stress beats in English: An Experimental study I. Language and Speech 15.72-100.

Allerton, D.J. 1969. The sentence as a linguistic unit. Lingua 22.27-46.

Allerton, D.J. 1978. The notion of ‘givenness’ and its relation to presupposition and to theme. Lingua 44.133-68.


Konstruktionen. (Linguistische Arbeiten 106.) Tübingen: Niemeyer.


Barry, William J. 1981. Prosodic functions revisited again! Phonetica
38.320-40.


Boyce, Suzanne & Lise Menn. 1979. Peaks vary, endpoints don’t: Implications for intonation theory. BLS 5.373-84.


Brazil, David; Malcolm Coulthard & Catherine Johns. 1980. Discourse
*Intonation and Language Teaching.* London: Longman.


Bresson, Daniel; Margret Sperbaum; Helmut Richter; Edeltraud Knetschke; & Werner O. Droschler. 1982 (eds.) Zur gesprochenen deutschen Umgangssprache I. (Phonai, Deutscher Reihe, Beiheft 5.) Tübingen: Max Niemeyer Verlag.


Chafe, Wallace L. 1988a. Linking intonation units in spoken English. In
Haiman & Thompson (eds.), 1-27.


Du Bois, John W.; Stephan Schuetze-Coburn; Danae Paolino & Susanna Cumming. 1992. Discourse Transcription. (Santa Barbara Papers in Linguistics 4.)


germanistischen Linguistik 18.46-93.


Ford, Cecilia E. & Sandra A. Thompson. 1993. Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the projection of turn completion. To appear in Ochs, Schegloff & Thompson (eds.).


Fuchs, Harald P. & Gerd Schank. 1975 (ed.) *Texte gesprochener deutscher Standardsprache III. "Alltagsgespräche"*. (Heutiges Deutsch II/3.)
München: Max Hueber Verlag.


Linguistik 56.1-31.


Jassem, Wiktor. 1952. Intonation of Conversational English. Prace
Wrocławskiego Towarzystwa Naukowego, Seria A, Nr. 45.


Kohler, Klaus J. 1983. Prosodic boundary signals in German. Phonetica 40.89-134.


435

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Cambridge, MA: The MIT Press.


Mann, William C. & Sandra A. Thompson. 1992 (eds.) *Discourse Description: Diverse linguistic analyses of a fund-raising text.* (Pragmatics & Beyond, New Series, 16.) Amsterdam: Benjamins.


O'Connell, Daniel C & Frank L. Slaymaker. 1984. Evidence for the


Ochs, Elinor; Emmanuel Schegloff; & Sandra A. Thompson. To appear (eds.) *Grammar and Interaction*.


Pawley, Andrew & Frances Syder. 1977. The one clause at a time hypothesis. Ms.


Pheby, John. 1975. *Intonation und Grammatik im Deutschen*. (Sammlung
Akademie-Verlag 19.) Berlin: Akademie-Verlag.


Pignataro, Vincenza. 1987. Topik und Fokus in der Sprachproduktion. (KoLiBri Arbeitsbericht 8.) Bielefeld: LiLi Fakultät, Uni Bielefeld.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Schuetze-Coburn, Stephan. In progress b. Explaining verb-initial clauses in German discourse.

Schuetze-Coburn, Stephan. In progress c. An outline of clause construction in English and German discourse.


Schwitalla, Johannes. 1976. Dialogsteuerung: Vorschläge zur

443
Untersuchung. In Berens, Jäger, Schank & Schwitalla (eds.), 73-104.


Fifth DARPA Workshop on Speech and Natural Language.


Sweet, Henry. 1877. Words, logic and grammar. Transactions of the Philological Society 1875-6, 470-503.


Thompson, Sandra A.; Patricia M. Clancy; Ryoko Suzuki; & Hongyan Tao. 1993. The conversational use of reactive tokens in three Pacific Rim languages. Submitted to *Journal of Pragmatics*.


48.819-38.


APPENDIX

Summary of transcription conventions

Each numbered line corresponds to an intonation unit (IU).

\_ = truncated (uncompleted) IU
\_\_ = possible/uncertain boundary

Phrase-terminal pitch direction:
\_ = falling
/ = nonfalling

Transitional continuity:
\* = complete
+= incomplete

Syllable prominence:
' = prominent

The phrase-final prominent syllable is most prominent unless another is indicated with ' (acute accent).

Tempo:
\* = segment/syllable lengthening
\* = final stretch of slow speech
\* = initial stretch of rapid speech
\langle A \rangle = other rapid speech
\langle L L \rangle = other rapid speech

Pause:
.. = brief pause (≤2 sec)/rhythm break
... = default pause (.3–.6 sec)
...(n) = timed pause (≥.7 sec)

Vocal noises & voice quality:
(H) = audible inhalation
(Hx) = audible exhalation
(%) = glottal stop or laryngealization
@ = laughter
(N) = nasal inhalation
(L) = audible lip noise
(T) = throat clearing
\langle QQ \rangle = 'quotation' quality
\langle PP \rangle = stretch of low volume speech
\langle FF \rangle = stretch of high volume speech
Other noises noted as (COUGH), etc.

Syntactic boundaries:
{ = clause start
} = clause end
{[ ]} = embedded clause
# = clause construction boundary

Interactional:
[ or ]\_\_ etc. = speaker overlap
(0) = 'latching', i.e. no break between
tums
A: = turn-initial IU
[A]: = nonturn-initial IU
(A): = backchannel utterance

Unlabeled IU's belong to current floor holder.

Other:
\^ = pitch reset (Chapter 5 only)
' = elided segment
<> = uncertain segment(s)
"" = quoted speech
(( )) = transcriber's comments

The length (in sec) of initial sequences of vocal noises & pauses is given in « »

450