Sound Patterns in Interaction

# Typological Studies in Language (TSL)

A companion series to the journal *Studies in Language*

Volumes in this series will be functionally and typologically oriented, covering specific topics in language by collecting together data from a wide variety of languages and language typologies. The orientation of the volumes will be substantive rather than formal, with the aim of investigating universals of human language via as broadly defined a data base as possible, leaning toward cross-linguistic, diachronic, developmental and live-discourse data.

**Volume 62**

Sound Patterns in Interaction: Cross-linguistic studies from conversation
Edited by Elizabeth Couper-Kuhlen and Cecilia E. Ford

# Sound Patterns in Interaction

Cross-linguistic studies from conversation

*Edited by*

Elizabeth Couper-Kuhlen
University of Potsdam

Cecilia E. Ford
University of Wisconsin-Madison

# Table of contents

# List of contributors

Peter Auer
Deutsches Seminar I
Albert-Ludwigs-Universität
79102 Freiburg, Germany
peter.auer@germanistik.uni-freiburg.de

Juliette Corrin
Department of Human Communication
Science
University College London
London WC1N 1PG, U.K.
juliette.corrin@btopenworld.com

Elizabeth Couper-Kuhlen
Institut für Anglistik & Amerikanistik
Universität Potsdam
14415 Potsdam, Germany
ecouper@rz.uni-potsdam.de

Traci S. Curl
Department of Language and Linguistic
Science
University of York
Heslington, York YO105DD, U.K.
tsc3@york.ac.uk

Cecilia E. Ford
Department of English
University of Wisconsin-Madison
Madison WI 53706, U.S.A.
ceford@wisc.edu

Barbara A. Fox
Department of Linguistics
University of Colorado
Boulder, Colorado, U.S.A.
Barbara.Fox@colorado.edu

Auli Hakulinen
Department of Finnish
PL 3 (Fabianinkatu 33)
00014 University of Helsinki, Finland
Auli.Hakulinen@helsinki.fi

John Hellermann
Department of Applied Linguistics
Portland State University
Post Office Box 751
Portland, OR 97207-0751, U.S.A.
jkh@pdx.edu

John Local
Department of Language and Linguistic
Science
University of York
Heslington, York YO105DD, U.K.
lang4@york.ac.uk

Richard Ogden
Department of Language and Linguistic
Science
University of York
Heslington, York YO105DD, U.K.
rao1@york.ac.uk

Barbara Rönfeldt
Nagelshof 3, 22559 Hamburg, Germany
barbara.roenfeldt@web.de

Margret Selting
Institut für Germanistik
Universität Potsdam
14415 Potsdam, Germany
selting@rz.uni-potsdam.de

Beatrice Szczepek Reed
Department of Music
University of York
Heslington, York YO105DD, U.K.
beatrice_szczepek@hotmail.com

Liisa Tainio
Department of Finnish
PL 3 (Fabianinkatu 33)
00014 University of Helsinki, Finland
liisa.tainio@helsinki.fi

Hiroko Tanaka
Department of Sociology
University of Essex
Wivenhoe Park, Essex CO4 3SQ, U.K.
htanaka@essex.ac.uk

Gareth Walker
Department of Language and Linguistic
Science
University of York
Heslington, York YO105DD, U.K.
gw115@york.ac.uk

Bill Wells
Department of Human Communication
Sciences
University of Sheffield
18/20/31 Claremont Crescent
Sheffield S10 2TA, U.K.
Bill.Wells@sheffield.ac.uk

# Introduction

# Conversation and phonetics
## Essential connections

Cecilia E. Ford and Elizabeth Couper-Kuhlen

In this introductory chapter we provide an initial conceptual backdrop for the volume. We review studies of conversational interaction which demonstrate specific attention to the phonetic design of talk and strongly invite the mutual integration of linguistic inquiry and conversation analysis (CA). And we outline the development of 'phonology for conversation', a field of linguistic inquiry responsive to and inspired by CA research. Finally we describe what is new in the volume and offer a brief preview of the contributions themselves as well as some of the ways they relate to one another.

## 1.   Why conversation and phonetics?[1]

Our title for this introductory chapter intentionally invites reflection on both theoretical and methodological issues at the intersection of these two fields. In placing "conversation" before "phonetics", we respond to the linear constraints of written language and resort to alphabetical ordering, but in fact the contributors to this volume analyze sound patterns hand in hand with interaction. Though our title separates "conversation" from "phonetics," our contributors view phonetics and interaction as most sensibly analyzed in tandem, because that is how they operate in real use.

We use the term *phonetics* – for the purposes of this chapter – to reflect analysts' attention to, and the potential interactional significance of, all audible aspects in and of speech that are produced by the human vocal apparatus. *Phonetics* as used here is thus a superordinate term encompassing at a subordinate level a broad array of sound patterns, some of which are traditionally called "segmental" or "phonetic",[2] some of which are "suprasegmental" or "prosodic" and others of which are not considered to be core to linguistic science at all and for this reason are called "paralinguistic". The methods and findings in this vol-

ume, however, contest the separation of segmental from suprasegmental and of linguistic from paralinguistic with respect to sound patterns in interaction. By using "phonetics" as a cover term here we attempt to avoid this separation.

If we recognize that languages are dynamic systems adapted to serve humans as they navigate through the recurrent and yet infinitely contingent social interactions that make up their lives, then it follows that the study of language must ultimately account for language as embodied in social interaction. The practices and resources of human languages are adapted to and for interactional functions, and as such, their study is best done simultaneously with the study of interaction. In enacted, embodied oral language use, sound patterns are fundamental, and as the contributors to this volume demonstrate, such patterns are best accounted for through methods that combine analysis of action with analysis of sound and in data from naturally occurring language use. The focus of the present collection, then, is on the interconnectedness of social action and sound patterns as documented in actual everyday use.

The studies collected here offer cutting-edge research in an innovative interdisciplinary area of scholarship aimed at broadening the domain of discourse functional linguistic theory and method to account for the organization of sound patterns in talk-in-interaction. At the same time, these studies serve to provide linguistic detail and grounding for observations regarding sound production in earlier conversation analytic research. In fact, the academic ground represented in this volume has developed into a lively area of interdisciplinary scholarship in its own right, an area populated by researchers committed to the integration of fine-grained analysis of sound patterns with fine-grained analysis of social action. The founding contributors to this interdisciplinary movement have generously contributed their most recent research to this volume, and a number of new scholars who find an intellectual home in the study of language and interaction have joined in.

Each chapter in the volume examines the organization of phonetic detail in relation to social actions in data from talk-in-interaction, and the data in each chapter are drawn from diverse languages: Japanese, English, Finnish, and German, as well as from diverse speakers: children, fluent adults and adults with language loss. Significantly, given the empirical practices traditional in phonetics research, while each study aims at rigor in its methods, none of the studies examines data produced purely for the purposes of such analysis; the talk analyzed here was real and consequential for the participants.

By acknowledging that linguistic inquiry is inseparable from the study of social practices, our contributors commit themselves to crossing traditional academic boundaries both in carrying out their research and in presenting

their findings. We cross these boundaries with the specific aim of integrating approaches, and we intend these studies to be available and comprehensible to students both of language and of human interaction regardless of academic home. Human actors do not respect the boundaries of academic disciplines in their spontaneous and skillful use of language in interaction, nor should students of human language or human interaction be limited by unproductive artifacts of the histories of our research traditions. The contributors to the volume work carefully to make their goals, methods and findings transparent and accessible across disciplines.

In this introduction, we provide an initial conceptual backdrop for the volume. We review studies of conversational interaction which demonstrate attention to sound production and strongly invite the integration of linguistic inquiry into an enterprise, conversation analysis (CA), that was initially conceived of as exclusively sociological. We then provide an overview of the development of 'phonology for conversation', a field of linguistic inquiry responsive to and inspired by CA research. And finally we offer a brief preview of the contributions themselves and some of the ways they relate to one another.

## 2.   How conversation analysis leads to phonetics

The form of linguistics represented in this volume is one that takes language to be a fundamentally situated social phenomenon. This outlook might be seen as fitting into the broader field of functional linguistics, a diverse but related set of methods and theories connecting language forms and structures to cognition, discourse, and culture (see Tomasello 1998, 2003). Such a perspective stands in contrast to one which strives to arrive at an account for language as a highly abstract system, cleanly distinct from other human activities. In idealizing language structures and forms, theories of an autonomous linguistics exclude vital aspects of the structures of languages; most relevantly to our needs here, such approaches exclude features that enact and reflect the adaptation of languages to the dynamics and contingencies of social interaction. Whether viewed ontogenetically or phylogenetically, language has its original home in social interaction. In stark contrast to autonomous theories, a functional linguistic outlook understands human social life as the environment in which language must be accounted for. Yet while the linguistics represented here is affiliated with the cross-disciplinary enterprise of functional linguistics, an exploration of social and cognitive motivations for patterns of use in diverse languages, not all of the contributors to this volume are primarily or only linguists. Conversation analy-

sis, a field originally developed by sociologists, underpins the studies of conversational phonetics offered here. Before we review the ways in which phonetics has been part of CA from its inception, let us provide an abbreviated overview of the CA approach itself for anyone not familiar with what distinguishes it from other ways of looking at discourse, language and human interaction.

Originating in a questioning of assumed categories in social science research (Heritage 1984), CA has developed into a rigorous approach aimed at accounting for the orderly and artful skills and practices that are enacted in human interaction – or more succinctly, for the work language does. Appropriately, from its inception CA research has been acutely aware of the work speakers do through the phonetic design of talk. This attention is reflected most strikingly in Gail Jefferson's original transcription system and those that have developed from it, because a concerted effort is made to reflect *all* audible details of and in everyday talk. The central role of sound production in interaction is also represented, though not elaborated, in Sacks, Schegloff and Jefferson (1974), a highly influential account of turn-taking, and sound production has been central to the elaboration of our understanding of what counts as a transition relevance place (TRP) since that time. Recent work in this area is showing the rich rewards of CA's influence on linguistics and vice versa, with phonetic and prosodic aspects of turn construction and turn transition being explored in a number of different languages.

Among the fundamental tenets of conversation analysis is that *a priori* categories – be they social (woman, man, non-native speaker of *X* language, etc.) or linguistic (subject, pronoun, verb, particle, incomplete sentence, etc.) – are at best suggestive starting points for inquiry. Traditional and taken-for-granted analytic categories demand skepticism and resistance until the grounds for their meaningfulness to humans in interaction have been demonstrated. Yet questioning categories is not an end in itself in CA; rather it is a means through which analysts step away from traditions in order to arrive at accounts for human social behavior that are representative of what its practitioners treat as real. While speakers do not, as a regular or frequent practice, explicitly state their interpretations of each other's talk, what they do offer is responsive actions which themselves entail interpretations of the talk and action that has preceded. Discovering the orientations of participants means noting how, through the timing and linguistic formulation of subsequent actions, participants in conversation enact their interpretations of prior talk. As Heritage (1984) has suggested, such displays of interpretation through sequential action are the bedrock for building and maintaining intersubjectivity, for making sense together and developing a common ground for emerging action. At the same

time they afford the analyst access to the sense that participants make of each other's actions. This yields a challenging and rewarding method for discovering the normative functions of forms of language as captured in recorded talk.

CA takes as its subject matter the emergent and negotiable production of local social order through talk-in-interaction; CA research thus focuses on the fundamental contingency inherent in talk, and on discovering human interactional practices that manage such contingencies. As nothing is pre-scripted in mundane interaction, the practices and resources we use to navigate through our social interactional lives must be specifically adapted to contingency. We do no doubt draw heavily on prefabricated and highly routinized utterances and sequences, but our skill at interaction is also fundamentally improvisational and adapted to the ever-changing particularities of interactional moments.

The sense-making practices we rely on in interaction and our constant updating of that sense as new contingencies present themselves relate to another aspect of the CA approach which is distinct from other methods of accounting for language. In contrast to contemporaneous methods of looking at language, CA research has demonstrated that there is no interactionally persuasive evidence for what linguists treat as "ungrammatical" nor even for what speech act theorists categorize as "infelicitous." Based as it is on working toward an understanding of how real interactions unfold, CA research shows that interactants are always making sense of one another and that any action will be interpreted based on the activity so far. In view of this evident behavior, there is no grammatical or ungrammatical, no felicitous or infelicitous human language use. Rather, within a context of developing interaction, any next action will be interpreted for the sense it makes in context and will be accountable, i.e. furnish grounds for inferences concerning 'Why that now?'. Thus, in responding to an utterance that seems problematic in some way, a speaker will rarely explicitly name it as wrong; instead, next speakers work to make sense of whatever has just been done. Orientations to standard prescriptive norms of language or etiquette are perhaps the closest one might come to what others (using grammatical judgments or judgments of felicity) might call acceptable or unacceptable uses of language. Whatever the source of an interactional problem, CA has found that a primary means for managing interactional "trouble" is through the ready-to-hand mechanisms and practices of repair.

CA's skeptical attitude toward traditional social and linguistic categories, and its resistance to attending to form at the expense of action, has not meant that form has been ignored – quite the contrary. Language forms as interactional practices have been a central focus of CA research from its earliest manifestations, and analysts have been extensively occupied with accounting

in detail for the formulation of actions in interaction. Noticings regarding formal aspects of talk – and in particular attention to the way words are chosen and produced – are evident at the very foundations of CA work. In Harvey Sacks' lectures ([1964–1972] 1992, volumes I–II), for instance, he observes orderliness and interactional functions in "pause distribution" in lists (I: 784), in turn-taking and the use of overlaps in turn-taking, in the artful use of laughter and *uh huh* (I: 739), in the mundane poetics of sound-sequence relationships (II: 292), and the selection of words by reference to sound (II: 308).

Evidence of exceptionally close attention to the orderliness of sound phenomena is found throughout the work of another of CA's founders, Gail Jefferson. Jefferson's acute observation of the social semiotic potential of sound manipulation is strikingly evident in her original and widely adopted system for the transcription of conversation. This detailed and infinitely adaptable system was based on the assumption that order is possible at all levels, and that no detail of sound that one might perceive can be treated, *a priori*, as irrelevant or as an error. Jefferson's system, while sometimes criticized for the cartoon-like representation of speech, grew from and supported the CA practice of "unmotivated looking," a methodological practice aimed at holding at bay the biases and assumptions one might otherwise bring to an analysis. The system has a built-in openness to further levels of detail and requires that transcribers put down whatever they can hear, regardless of whether it has been proven to be "meaningful" by any tradition or by common sense.

The consequences of Jefferson's transcription method for CA's basic discoveries can be underestimated. The practice of specifically attending to form turned out to provide powerful leverage for learning about things we never imagined were indeed orderly. Transcription in this form is as much analysis as it is the creation of a written record:[3] in fact, one is constantly updating a transcript as one combs through it and comes to note even finer levels of detail. The transcription process becomes what one early researcher called a "technology stimulating close attention to the details of conversational utterances" (Schenkein 1978: 6). This opening of the ear, so to speak, regularly leads to noticing ways in which previously unheard (and untranscribed) details play a role in the actions of the participants.

In Jefferson's own research, attention to aspects of the production of talk has led to observations on the ways that laughter tokens work (1979, 1985), on the practice of producing a token such as *nyam*, an artful melding of affirmative and negative token sounds (1978), on the social significance of a syllable that is barely uttered and then repaired (1974). Jefferson has also produced a compendium of insights on the poetics of everyday talk, which she originally

presented as a lecture in 1977 and later published as a lengthy article (1996). Here again she reveals ways in which interactants rely on sound patterns within and across speaker turns to achieve social ends. The level of detail in this work includes patterns of sound selection and repetition, as well as attention to consonant clusters and their reversal. The central role of phonetic and prosodic detail in interaction is of course also reflected, or at least gestured to, as part of what is involved in turn projection as first outlined in Jefferson's collaboration with Sacks and Schegloff. This collaboration resulted in their highly significant account of turn-taking, research that has been the touchstone for CA work ever since.

The importance of attention to phonetic detail is also noted in Emanuel Schegloff's foundational and continued contributions to CA. Schegloff has offered numerous noticings of the key role of "pitch peaks" in the projection of upcoming transition relevance places in English conversation (e.g., 1987, 1996). And he has repeatedly noted the phenomenon of "rush through", placed just as a possible point of turn completion is imminent (1979, 1996). He has also drawn attention to contrastive stress and what it may index as well as to the modulation and matching of pitch in the opening of a conversation (1998). Like Goldberg (1978), who observed the contribution of significant amplitude manipulation in displaying affiliation, disaffiliation, and sequence beginnings, Schegloff has demonstrated the use of loudness (a perceptual correlate of amplitude) in managing and resolving overlaps in talk (2000). Schegloff has also noted the ways that words that are "suppressed" for interactional reasons may "surface in the immediately following talk" (2002: 236). Signficantly, what surfaces may only be similar in sound but not in denotation or grammatical function, e.g., "mean" as an adjective indicating malevolence vs. "mean" as a verb indicating intention.

Another rich source of detailed examination of patterned sound production in the CA tradition can be found in the work of Charles and Marjorie Goodwin. These scholars have consistently integrated patterns of sound, gesture and gaze into their analyses. For example, Goodwin (1979) is a detailed (interdisciplinary) examination of sentence, turn, and sound production in relation to gaze and the contingent coordination of speaker and recipient behavior. The article draws attention to the work of what here and elsewhere in CA research is termed a "phrasal break." The break in this case is accomplished with a glottal stop followed by the syllable *uh:* produced with noticeable sound stretch. The break itself works to "request...the gaze of a recipient," while the *uh* produced with "extra length" is a means to "extend the fragment until [the recipient's] head move has been completed (1979: 108). Goodwin

and Goodwin (1987) delineates the ways that assessments can be produced, through hesitations and sound stretches, to invite overlap and offer opportunities for recipient responses. Notably, the Goodwins never separate phonetic description from the description of gesture, gaze and lexico-grammatical formulation; their research demonstrates the ways that these simultaneously unfolding aspects of talk are mutually elaborating.

In addition to drawing attention to aspects of sound production in their own studies, CA practitioners have repeatedly called upon linguists to collaborate in accounting for language practices in interaction. And linguists are beginning to hear the call. In the realm of grammar, for example, Sacks ([1964–1972] 1992), referred to "first verbs," but it is only very recently (Ford 2000; Schulze-Wenck to appear) that a more detailed understanding of what these forms might entail has been explored. In the realm of sound, the phenomenon of the "cut off" has been frequently referenced in CA studies, but its formal features were not specified until very recently (Jasperson 1998, 2002). Such a specification is essential in accounting for participants' interpretation of what linguistic/action trajectory has been cut short. "Cutting off" refers to stopping the progress of a clearly projected trajectory, most generally a word's production, at a point when the full form of the word has not been produced but where the projected trajectory is potentially identifiable. Cut-off is a common way in which conversational repair is achieved (Schegloff, Jefferson, & Sacks 1977). Not until Robert Jasperson's comprehensive phonetic analysis of cut-offs (1998, 2002) have we had access to anything beyond intuitive descriptions of this phenomenon. There is little doubt that greater facility and deeper experience with the description of language form and structure – the skill, that is, of trained linguists – can only augment CA, provided that such skill is brought to bear in a manner that takes action as basic and maintains a healthy skepticism regarding traditional categories.

While the contributors to this collection have varying backgrounds, we share a commitment to using methods from conversation analysis, a powerful and challenging approach to understanding human interaction, for the study of phonetics in interaction. To our thinking CA is as useful a tool for linguistic inquiry as it is for social inquiry. Indeed some of us resist the demarcation of strict borders between the subject matter of conversation analysis and of (functionally oriented) linguistics. CA has included the analysis of sound patterns from its inception, and the accumulation of CA findings on the basic practices of interaction clearly leads to and entails a form of phonetic inquiry for talk-in-interaction.

Having touched upon the ways that CA research has led to and contributed to a study of phonetics in talk-in-interaction, we now turn to research and findings from linguistically trained scholars.

## 3.    How phonetics leads to conversation analysis[4]

It was with the advent of structuralism in the early twentieth century that the disciplines of phonetics and phonology came into their own. In the United States the development of phonemic theory can be traced directly to the "discovery" of Native American languages and attempts to preserve them through the introduction of appropriate writing systems. Yet once the early boom in phonology had subsided, there followed an ever so slow recognition that phonemic models work successfully – if at all – only for citation forms of language. Elaborate supplements to these models were necessary to account for so-called "connected speech", and the result was less than satisfactory. As it turns out, introspection – the predominant methodology employed at the time – is notoriously limited with respect to discourse-size chunks of language. So coming to terms with the "hurly-burly"[5] of everyday language situations seemed out of the question, although far-sighted linguists called for precisely this. One of the first to do so was a phonetician, David Abercrombie, who pointed out some forty years ago that contemporary linguistic endeavor addressed little more than "spoken prose", by which he meant "essentially language organized for visual presentation" (1965: 3f.).

The impetus for a radically new approach to phonetics, one that is capable of coming to terms with everyday conversation in its own right, came out of the University of York and is enshrined in a volume by two phoneticians, John Kelly and John Local, entitled *Doing Phonology* (1989). Taking inspiration from Firthian linguists, they claim that traditional phoneme-based approaches to phonetics and phonology bring unwarranted assumptions with them, namely that the speech continuum is segmentable into discrete units, that these units are neatly sequenced in the stream of speech with little or no overlap, that allophonic variance in phonemes is uniquely determined by phonetic context, etc. Kelly and Local conclude that the phoneme – a unit developed to support writing systems – is poorly equipped to handle the most common form of language use, conversation. Moreover, they point out that so-called "suprasegmental" analysis as practiced so far also has a written-language bias. Attention is paid above all to features which are capable of orthographic representation, e.g. via dashes (pause), italics (stress) and/or punctuation (final pitch movement).

Kelly and Local argue that these and other putative suprasegmental categories such as "nuclear tone" have been hypostasized based on idealized language use: they are derived from the examination of simple sentences read aloud, rather than from genuine utterances embedded in rapid familiar conversation.

Instead of relying on *a priori* phonetic and phonological categories, Kelly and Local advocate close listening to real speech in actual situations of language use and on "impressionistic" recording. By this they mean attending to and notating every phonetic detail which a trained ear can perceive, including the "articulatory skeleton" of speech, its long-domain properties such as pitch, loudness, tempo, syllable rhythm and articulatory/phonatory settings, its resonance[6] as well as the variability and co-occurrence, relativisms and phasing of its parameters. Emphasis is placed on a parametric, dynamic and relative view of phonetic substance. Only once a careful impressionistic record has been made of speech (or of a target utterance therein) can functional analysis, or data "interpretation", follow. The latter involves looking for sound patterns and relationships and, in an interactional perspective, setting them in relation to empirically discoverable *tasks* which participants in an interaction can be shown to be addressing.

The final chapter in Kelly and Local (1989) presents a case study in which precisely this is done. The study is exemplary in showing how a conversation analytically informed approach can reveal phonetic patterns which contribute to a *phonology for conversation*. The data come from dialect survey interviews in Tyneside, specifically from sequences in which informants are asked if they recognize or use a particular dialect word. If and when they subsequently repeat the word in question, Kelly and Local show that this word-repeat turn can be interpreted as (i) a display of recognition, (ii) as an understanding check or (iii) as "mulling over". For each of these interactional tasks, distinct clusters of phonetic events recurrently accompany the word repeated. They argue that it is the phonetic patterns which steer the interviewer's interpretation of the turn and prompt him (i) to acknowledge the recognition, (ii) to repeat the word or (iii) to withhold talk, respectively, in next turn.

Kelly and Local's (1989) study is not only exemplary because of the relation it establishes between conversational interaction and phonetics. It is also exemplary because it stresses that the sound patterns identified are dialect-, situation-, sequence- and turn format-sensitive. In other words, the recurrent clusters of phonetic features which they discover have the described effect only in the Tyneside dialect, in a survey interview, following a word inquiry by the interviewer and in a turn by the interviewee formatted as a word-repeat. Else-where, with a different constellation of contextual parameters, the same pho-

netic clusters might have a different import, and/or different phonetic clusters might have the same import. The study thus demonstrates both the potentials and the limitations of generalizing about sound patterning in interaction.

Local's work, often in collaboration with colleagues and students at York, has been seminal in founding a phonology for conversation. By avoiding orthographically based phonological categories and starting directly from the audible phonetic details of speech production in interaction, his studies have revealed dimensions of patterned sound production in talk-in-interaction hitherto unfathomed. Pitch and loudness register shifts (French and Local 1983), glottal holding pauses (Local and Kelly 1986), assimilatory projection vs. articulatory disjunction (Local and Kelly 1986), pitch and loudness matching (Local 1992), "abrupt joins" (Local and Walker to appear) – all of these phonetic phenomena were virtually unknown before and certainly none were suspected of being systematically deployed in conversation. Local's studies have also shown that actions and tasks previously identified in conversation analytic work have phonetic exponents – for example, turn delimitation (Local, Kelly and Wells 1986), competition for the floor (French and Local 1983), turn holding (Local and Kelly 1986), turn continuation following suspension (Local 1992), news receipts (Local 1996). In addition, he has shown that the phonetic exponents of at least some of these tasks vary significantly within the English-speaking world. Turn delimitation ('I'm finished, it's your turn') sounds different in Tyneside (Local, Kelly and Wells 1986) from the way it does in London Jamaican (Local, Wells and Sebba 1985), and both sound different from turn delimitation in Belfast (Wells and Peppé 1996). In each case it is (varying) clusters of phonetic parameters – pitch, loudness, duration, tempo, rhythm, articulatory and phonatory settings – which serve as exponents of the conversational or interactional work being done. In each case Local and his co-workers are careful to show that the phonetic parameters identified are relevant for the participants themselves, because their behavior shows an observable orientation to them.

In recent years Local's students and his students' students, as well as other sympathetic phoneticians and prosodists, have pursued the agenda of a phonology for conversation. Couper-Kuhlen (1993) looks specifically at isochronous speech rhythm in English conversation as a means for signaling sequence organization and preference. Tarplee (1996) and Couper-Kuhlen (1996), taking off from Kelly and Local's early word-repetition study, broaden the phonetic and prosodic inquiry to other interactional contexts in which turns are repeated. Wells and Macfarlane (1998) point to hitherto unspecified "TRP projecting" pitch accents as being instrumental in signaling upcoming transition relevance in English conversation.[7] And Couper-Kuhlen (2001)

spots the systematic use of high onset as a means for marking reason-for-the-call turns in an American radio phone-in program. What all of these studies have in common is that they involve close observation of phonetic and prosodic substance in actual conversational records, without reference to pre-existing phonological categories, and that they establish recurrent correspondences between a particular set of phonetic or prosodic parameters and a particular interactional activity or task. Moreover, all show concern to warrant their analyses through the orientations of the interactants themselves.

On the Continent other linguists, most notably Selting and Auer, have been instrumental in extending Local's approach to a phonology for conversation to the study of German. Selting (1995) shows that Local's methodology lends itself equally well for the identification of phonological and prosodic units in German conversation. Starting from similar assumptions, she describes prosodic parameters for turn construction, turn-taking and specific conversational activities such as story-telling and argumentation based on German conversational data. Selting's further studies investigate prosodic dimensions of, e.g., conversational questions (1992), speech styles (1994), repair initiation (1996), unit construction (2000) and lists (2003). Likewise in Auer's work a phonology-for-conversation perspective is unmistakable: Auer (1996), for instance, is a seminal treatment of the prosody and syntax of turn continuation in German. It is work of this sort (see also Günthner 1996, 2000 and Uhmann 1996, 1997) that lays the foundation for a cross-linguistic comparison of the phonological organization of conversation. Auer, Couper-Kuhlen and Müller (1999), in fact, do just this with respect to conversational rhythm in English, German and Italian talk-in-interaction.

## 4.    What is new in this volume

Past work on phonetic design in talk-in-interaction has thus been lively, if for the most part Anglo-German.[8] This work finds its proper continuation in the present volume, in many ways with expanded scope. For one, the set of languages to come under scrutiny has been enlarged: in addition to English and German (Selting; Auer), we include a study of Japanese (Tanaka) and two of Finnish conversation (Ogden; Ogden, Hakulinen and Tainio). This selection is determined in part by accident: there simply happen to be trained phoneticians and/or trained conversation analysts who are interested in phonology for conversation in these specific languages. Yet there is every reason to believe that as time goes on the circle of conversational phonologists will enlarge and

spread to other, perhaps even more diverse languages and cultures. Second, the scope of conversants has been expanded here to include not only adults with unimpaired language ability but also children (Wells and Corrin) and aphasics (Auer and Rönfeldt). The general research questions are the same for all chapters: 'What phonetic resources are exploited in dealing with this conversational task?' or 'How is this specific interactional goal furthered by linguistic, especially phonetic, means?' But in the chapters dealing with children and aphasics, these questions are supplemented by a third question: 'How is this process affected by the special circumstances at hand?' Needless to say, any and all types of communication under special circumstances are candidates for such an inquiry, and it is to be hoped that more studies of conversation and phonetics in special communicative situations will emerge in the future.

There are two ways in which the following chapters widen the basis for a phonology for conversation. First, a larger selection of phonetic and prosodic parameters are attended to. Although some contributions display a continued interest in pitch contour (e.g. Szczepek Reed; Selting) and pitch height (e.g. Couper-Kuhlen), other contributions describe the interactionally relevant use of voice quality (e.g. Ogden), phonatory setting (e.g. Tanaka) and loudness (e.g. Auer and Rönfeldt). Above and beyond traditional long-domain or prosodic properties such as pitch and rhythm (e.g. Ogden, Hakulinen and Tainio), the manipulation of articulatory settings (e.g. Curl; Walker) and vowel quality (e.g. Local) is also shown to be consequential for participants. Second, the chapters encompass a wider range of conversational tasks and interactional goals. There has always been a keen interest in phonetic and prosodic parameters as they relate to turn-taking, and this topic continues to be a concern for, e.g., Szczepek Reed, Ogden and Tanaka. Yet we also find chapters dealing with the projection of more-to-come at turn beginnings (Ford, Fox and Hellermann), with the continuation of turns past a point of possible completion (Walker; Ford, Fox and Hellermann), with the use of repetition in other-initiated self-repair (Curl), with the disjunction of adjacent turns at talk (Local) and with the initiation of new sequences (Couper-Kuhlen).

The chapters which follow all have in common that they view the relationship between phonetics and conversation as dialect-specific (cf. e.g. Selting), as sequence type-specific (cf. e.g. Ford, Fox and Hellermann), as specific to location in particular unfolding sequences (cf. e.g. Couper-Kuhlen) and as format-specific (cf. e.g. Curl). In this sense they all follow in the path first marked out by Kelly and Local (1989). Yet in addition to the insights each study provides on single phonetic phenomena, there is a value-added dimension which comes from the deployment of a similar methodology for the investigation of similar

conversational tasks in different, and to an extent genetically and typologically unrelated languages. Upon reflection, a number of the chapters which follow will be seen to establish building blocks for a contrastive, or cross-linguistic, phonology for conversation.

Take, for example, the case of turn delimitation. Although doubts are raised about the role of final pitch movement in standard varieties of English (Szczepek Reed), we do have a body of conversation analytically grounded knowledge about the phonetics of turn delimitation in varieties such as Tyneside, London Jamaican and Belfast English. The chapters by Ogden and Tanaka now add to this body of knowledge in fascinating and unexpected ways. Ogden, for instance, shows that non-modal voice quality is used normatively in Finnish to mark the end of a turn cued as transition-ready. Compare this finding to Tanaka's study of Japanese: she examines truncated turns, i.e. those which lack the usual utterance-final elements marking termination, and finds that, among other clusters of phonetic parameters, reduced duration of the final syllable followed by a glottal stop is not an unusual pattern. Taken together, Ogden's and Tanaka's findings thus suggest that voice quality features may serve a turn delimiting function in Finnish and (under certain circumstances) in Japanese, although they do not appear to play a comparable role in English. Not only different varieites of the same language but also different languages thus deploy phonetic resources in different ways for turn delimitation. It is but a small step from this conversational task to others. Similar cross-linguistic comparison now becomes possible with respect to, e.g., turn continuation (compare Walker for English with Auer 1996 for German) and stylization (compare Ogden, Hakulinen and Tainio for Finnish with Couper-Kuhlen to appear, for English). With more and more studies of this sort, a better understanding of cross-linguistic phonetic regularities and of language-specific phonetic resources and their conversational deployment becomes possible. The present volume takes an important step in this direction.

## 5.    The chapters

While all the contributions to this volume analyze turns and sequential actions, we subdivide the chapters with respect to their special foci, trusting that the reader will see the fundamental overlapping of methods and the overarching attention to both turn and sequence. The three sections of the book reflect three major ways that interactants achieve structural orderliness in talk: transition from one turn-at-talk to the next, the construction of single- or multi-

unit turns, and the sequential organization of consecutive turns-at-talk. With respect to each type of structure the single chapters explore how phonetic and prosodic resources – together with lexical and syntactic ones – are brought to bear on the construction of talk in ways that allow conversationalists to shape and coordinate actions and to display their sense-making to each other methodically. At the same time, because the chapters examine phonetic, prosodic and paralinguistic practices in different languages, the volume makes it apparent that language-specific constraints are also at work in determining exactly which resources are deployed for a given purpose and how they articulate with one another in different cultures and speech communities.

Grouped together in the section *Practices and resources for turn transition* are papers that address the phonetic design of possibly complete turns and what young children must learn in order to master the phonetic organization of turn-taking and overlap. The evidence for a turn being possibly complete is found in what happens next in talk-in-interaction: there may be a smooth transition to next speaker or, in its absence, some indication by current speaker that the turn-so-far was ready for transition, e.g. pursuit of response. Or negative evidence may be used to show that where the putatively relevant phonetic features are absent, no orderly transition to next speaker occurs. It is in pursuing such a line of argumentation that Ogden's chapter makes a case for a shift to non-modal phonation (including creak, breathiness, whisper, voicelessness and/or exhalation) as normatively marking transition relevance in Finnish conversation. Using similar argumentation, Tanaka makes a case for specific bundles of prosodic features – e.g. lengthening and resurgence of loudness, glottal stop or turn compression – signaling transition-readiness in "truncated" Japanese turns, which lack the usual utterance- and turn-final elements. But whereas one of Ogden's points is that phonatory setting can work independently of intonation in turn delimitation, Tanaka's point is that these prosodic features, which otherwise appear to play a negligible role in Japanese turn-taking, work in tandem to mark transition relevance when the usual lexico-syntactic cues are lacking. A comparison of these two chapters thus provides a lesson in the way lexical, syntactic and phonetic/prosodic resources for turn transition can play off quite differently in genetically and areally unrelated languages.

Szczepek Reed's chapter on turn delimitation in English is based on a very simple observation concerning standard British and American talk-in-interaction: smooth speaker transitions occur after virtually any and every type of final pitch movement. This observation casts doubt on the frequently made claim that in non-regionalized accents of English final falls-to-low and rises-

to-high are turn-ending, whereas final falls-to-mid, low rises and level pitch movements are turn-holding. We note that Szczepek Reed's finding does not entail that intonation is altogether irrelevant for turn delimitation in standard English accents, but simply that pitch movement on so-called "tail" syllables, those which follow the last major accent, may be. This finding is compatible with more recent work suggesting the importance of "TRP-projecting" (on-syllable) accent types (Wells and Macfarlane 1998; Schegloff 1998).

The focus of Wells and Corrin's chapter is developmental: the data are taken from one child-mother dyad during a particularly critical stage, the latter part of the second year, when a resurgence in overlap is said to occur. Of particular interest here is the finding that the child in question appears to lack the ability to compete for the floor phonetically (by the use of high pitch and loud volume) or to resolve overlap e.g. in a curtail-and-recycle pattern. This suggests that the phonetics of turn-taking must indeed be learned and that children may not achieve adult-like mastery of e.g. turn competition or overlap resolution until after the age of two.

The second section *Projecting and expanding turns* groups together papers which address the question of how turns larger than a single unit are projected and/or constructed. The expanded turns which Walker and Auer and Rönfeldt describe, although they are produced by very different kinds of speakers – in the one case language-unimpaired speakers of English, in the other language-impaired speakers of German – come about incrementally over time. In both cases phonetic parameters are deployed to accomplish the further production of talk as a continuation of the speaker's turn. Walker examines grammatically dependent talk beyond a point of possible turn completion. Referred to as incrementing in the literature, this practice is shown to entail not only grammatical but also phonetic continuation with respect to pitch, loudness, speech rate and articulatory characteristics. By matching such features in the new bit of talk to those of prior talk, speakers are able to display that what they are producing is indeed coherent and cohesive with what precedes. Auer and Rönfeldt describe a phonetic practice known from unimpaired speech, *diminuendo* at the end of a unit followed by sudden *forte* at the beginning of a new unit. Their data, from interaction with a Wernicke aphasic patient, show this technique being skillfully co-opted to mask word-finding difficulties: the patient can thus gain time for word retrieval without losing the floor, although interlocutors may feel that they are being deprived of the right to come in.

The following two chapters, by Selting and Ford, Fox and Hellermann, deal with the projection of multi-unit turns in German and English respectively. Selting's study examines two related intonation patterns, termed "up-

wards staircase" contours, characteristic of the Berlin variety of German and used to project more to come in biographical story-telling. She finds that one is used repetitively for list-like enumerations, while the other tends to occur singly to mark the beginning of a new stage or episode in the narrative. Both appear to open up a gestalt-like structure, which is later "closed", and to be interpretable as conveying recurrentness, routineness and/or expectableness with respect to the situations being recounted.

Ford, Fox and Hellermann focus their attention on *no*, a word that may or may not constitute a possibly complete turn when used as a turn-initial token to deny, reject or disagree with a prior turn functioning as a *yes/no* question. They discover that its phonetic production patterns differently in two different interactional environments: (1) in larger telling projects, where *no* turns in response to questions by a primary speaker tend to be stand-alone *no* – versus in response to questions by the recipient, where there tends to be further talk past *no*, forming a multi-unit turn. Stand-alone tokens of *no* are longer and quieter than tokens of *no* in *no*-plus turns in this sequential environment. (2) In topic proffers, where the initiating turn proposes a topic for further talk which the recipient may or may not take up. Here stand-alone *no*s differ from *no*-plus turns by being lower in range and having a decrease in energy. The findings in Ford, Fox and Hellermann's study underline the complexity of sound patterning in interaction: sequential location, participant roles, lexical composition and phonetic production features all seem to work together to project more to come (or not).

The chapters of the final section *Implementing actions across turns* deal with sound patterns and actions and with how they are formulated as connected (or disconnected) across turns. Common to all chapters is a grounding of the action analysis in observable behavior by the participants. Ogden, Hakulinen and Tainio's chapter considers the phonetic, sequential and interactional properties of a stylized figure found in Finnish conversation which appears to mark out something in talk as obvious or not worth an undue amount of attention. Of interest here is not only the fact that intonational stylization, a phenomenon described so far primarily in Germanic languages, is now identified in Finno-Ugric (and in a language-specific fashion), but also that it is shown to have a sequentially and interactionally specific use in conversation. Both internal and external evidence is presented for this: speakers make lexical and morpho-syntactic choices in accordance with the effect of the stylized figure and co-participants show an orientation to it by producing only minimal responses and/or by subsequently shifting topic.

In a similar vein, Curl's chapter also argues that a phenomenon which has hitherto been identified as "phonetic", namely repetition, requires sequentially sensitive treatment. She examines lexical repetition in turns which are produced as repair following a next-turn repair initiating device and discovers two distinct phonetic patterns: one in which the repair expands the pitch, loudness, duration and/or alters the articulatory characteristics of the original trouble source, and one which maintains or reduces the range of these features in the original. She argues that these patterns are used systematically to display whether the repairer is treating the trouble source as fitted to prior talk or as disjunct from it.

Coherence and disjunction are relevant notions for the next chapter as well. Couper-Kuhlen is concerned to show that at possible sequential junctures in conversation, the prosodic delivery of a next turn, often together with its lexical and morpho-syntactic format, contributes to marking it as the beginning of a new sequence (with a sudden surge of pitch and loudness) or as a continuation of what went before (without a surge). It turns out that such prosodic formatting is more indicative of sequence structure than of topic organization.

Local's chapter examines the device *and uh(m)*, which he argues participants use with a stable cluster of phonetic characteristics to display that their upcoming talk is to be treated not as cohering with the immediately prior talk, but as relating to some earlier talk of theirs. *And uh(m)* produced this way proposes a return to prior action over intervening sequences which can be quite lengthy, a return which continues (rather than restarts or recapitulates) that action. One of Local's concluding remarks is particularly worthy of note. The recognizability of the *and uh(m)* device, he points out, is not due uniquely to its phonetics but is instead constituted by the totality of its design features: sequential location, position in the turn and phonetic characteristics. This, in fact, is a statement which could apply to all the work collected here.

## 6.  Closing

The present volume contributes to the growing exchange and collaboration among sociologists and linguists by offering new models for approaching language use in terms of both theoretical grounding as well as technical detail and methodology. It presents the most recent findings on phonetic design in interactional discourse available in an edited collection. We hope that the collection will stimulate further interchange between empirically oriented linguists with research interests in phonetics, prosody and grammar in use, and researchers

on talk-in-interaction whose interests extend to the linguistic underpinnings of social action and interaction.

The forms of inquiry presented in these studies are necessarily interdisciplinary and thus entail stepping across academic borders and treading into less known territory (from the perspective of the authors). Whether our academic homes are in sociology or linguistics, to name two key disciplines contributing to this endeavor, we are all on uncharted interdisciplinary turf: none of us has been trained from the outset to include both fine-grained social and fine-grained linguistic description in our empirical methods. In light of our status as non-authorities in this non-field, we invite interested scholars, especially those who are most at home in one or another of the distinct fields we traverse, to approach our work knowing that our methods will not thoroughly align with those of any one field, and to imagine the rewards to be had in exploring these borderlands together.

## Notes

**1.** We wish to thank Traci Curl, Auli Hakulinen, John Local, Richard Ogden, Margret Selting and Gareth Walker for helpful comments on a first draft of this chapter. We assume full responsibility for not always having followed their advice.

**2.** The term "phonetic" thus appears at both superordinate and subordinate levels of the category. Where it is deemed necessary to invoke the lower level, we speak of "phonetic and/or prosodic and/or paralinguistic" resources.

**3.** See Ochs (1979) for a full discussion of implications in understanding transcription as analysis.

**4.** There are currently other approaches which examine phonetics in casual, spontaneous and/or informal speech (cf. e.g. the papers in Kohler and Simpson 2001 and Docherty 2003). However, these approaches are concerned primarily with tracking and modeling phonetic variability in specific lexical items, specific classes of sound and "connected speech processes". They so far not engaged directly with the interactional work which such phonetic resources are deployed to accomplish.

**5.** We borrow Kelly and Local's (1989) picturesque term here.

**6.** A term used to refer to secondary articulations such as palatalization and velarization together with various intermediate qualities.

**7.** Schegloff (1987) notes that pitch peaks project up coming transition relevance places or TRPs, but he does not specify what a pitch peak might be in detail. Fox (2001) explores what features Schegloff might be attending to when he refers to a "pitch peak".

**8.** The first collected volume to be published (Couper-Kuhlen and Selting (Eds.) 1996) contained only one study of Romance prosody: Müller (1996).

# References

Abercrombie, D. (1965). *Studies in Phonetics and Linguistics.* London: Oxford University Press.

Auer, P. (1996). "On the prosody and syntax of turn-continuations." In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation. Interactional Studies* (pp. 57–100). Cambridge: Cambridge University Press.

Auer, P., Couper-Kuhlen, E., & Müller, F. (1999). *Language in Time. The Rhythm and Tempo of Spoken Interaction*. New York & Oxford: Oxford University Press.

Couper-Kuhlen, E. (1993). *English Speech Rhythm. Form and Function in Everyday Verbal Interaction*. Amsterdam: Benjamins.

Couper-Kuhlen, E. (1996). "The prosody of repetition: on quoting and mimicry." In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation. Interactional Studies* (pp. 366–405). Cambridge: Cambridge University Press.

Couper-Kuhlen, E. (2001). "Interactional prosody: High onsets in reason-for-the-call turns." *Language in Society, 30,* 29–53.

Couper-Kuhlen, E. (to appear). "Prosodische Stilisierungen im Gespräch." In A. Assmann, U. Gaier, & G. Trommsdorf (Eds.), *Zwischen Literatur und Anthropologie. Performanzen, Diskurse, Medien*. Tübingen: Narr.

Docherty, G. J. (2003). "Speaker, community, identity: Empirical and theoretical perspectives on sociophonetic variation." In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 11–16).

Ford, C. E. (2000). "On 'pragmatic' projection of turn trajectories." Paper presented at the EuroConference on Interactional Linguistics, Spa, Belgium. [ms. available: ceford@wisc.edu]

Fox, B. A. (2001). "An exploration of prosody and turn projection in English conversation." In M. Selting & E. Couper-Kuhlen (Eds.), *Studies in Interactional Linguistics* (pp. 287–316). Amsterdam: Benjamins.

French, P., & Local, J. (1983). "Turn-competitive incomings." *Journal of Pragmatics, 7,* 17–38.

Goldberg, J. (1978). "Amplitude shift. A mechanism for the affiliation of utterances in conversational interaction." In J. Schenkein (Ed.), *Studies in the Organization of Conversational Interaction* (pp. 199–218). New York: New York Academic Press.

Goodwin, C. (1979). "The interactive construction of a sentence in natural conversation." In G. Psathas (Ed.), *Everyday Language: Studies in Ethnomethodology* (pp. 97–121). New York: Irvington.

Goodwin, C., & Goodwin, M. H. (1987). "Concurrent operations on talk: Notes on the interactive organization of assessments." *Papers in Pragmatics, 1,* 1–54.

Günthner, S. (1996). "The prosodic contextualization of moral work: an analysis of reproaches in 'why'-formats." In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation. Interactional Studies* (pp. 366–405). Cambridge: Cambridge University Press.

Günthner, S. (1999). "Polyphony and the 'layering of voices' in reported dialogues: An analysis of the use of prosodic devices in everyday reported speech." *Journal of Pragmatics, 31,* 685–708.

Günthner, S. (2000). *Vorwurfsaktivitäten in der Alltagsinteraktion: Grammatische, proso-dische, rhetorisch-stilistische und interaktive Verfahren bei der Konstitution kommuni-kativer Muster und Gattungen.* Tübingen: Niemeyer.

Heritage, J. C. (1984). *Garfinkel and Ethnomethodology.* New York: Polity Press.

Jasperson, R. (1998). *Repair after Cut-off: Explorations in the Grammar of Focused Repair of the Turn-constructional Unit-so-far*. PhD diss., University of Colorado at Boulder.

Jasperson, R. (2002). "Some linguistic aspects of closure cut-off." In C. E. Ford, B. A. Fox, & S. A. Thompson (Eds.), *The Language of Turn and Sequence* (pp. 257–286). Oxford & New York: Oxford University Press.

Jefferson, G. (1974). "Error correction as an interactional resource." *Language in Society, 2,* 181–199.

Jefferson, G. (1978). "What's in a 'Nyam'?" *Sociology, 12,* 135–139.

Jefferson, G. (1979). "A technique for inviting laughter and its subsequent acceptance/ declination." In George Psathas (Ed.), *Everyday Language: Studies in Ethnomethodology* (pp. 79–96). New York: Irvington.

Jefferson, G. (1985). "An exercise in the transcription and analysis of laughter." In T. A. van Dijk (Ed.), *Handbook of Discourse Analysis*, Vol. 3 (pp. 25–34). London: Academic Press.

Jefferson, G. (1996). "On the poetics of ordinary talk." *Text and Performance Quarterly, 16,* 1–61.

Kelly, J., & Local, J. (1989). *Doing Phonology.* Manchester: Manchester University Press.

Kohler, K. J., & Simpson, A. P. (2001). "Patterns of speech sound in unscripted communication." *Journal of the International Phonetic Association, 31*(1).

Local, J. (1992). "Continuing and restarting." In P. Auer & A. di Luzio (Eds.), *The Contextualization of Language* (pp. 273–296). Amsterdam: Benjamins.

Local, J. (1996). "Conversational phonetics: Some aspects of news receipts in everyday talk." In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation: Interactional Studies* (pp. 177–230). Cambridge: Cambridge University Press.

Local, J., & Kelly, J. (1986). "Projection and 'silences': Notes on phonetic and conversational structure." *Human Studies, 9,* 185–204.

Local, J., & Walker, G. (to appear). "Abrupt-joins as a resource for the production of multi-unit, multi-action turns." *Journal of Pragmatics.*

Local, J., Wells, W., & Sebba, M. (1985). "Phonology for conversation: phonetic aspects of turn delimitation in London Jamaican." *Journal of Pragmatics, 9,* 309–330.

Local, J. K., Kelly, J. & Wells, W. (1986). "Towards a phonology of conversation: turn-taking in Tyneside English." *Journal of Linguistics, 22,* 411–437.

Müller, F. E. (1996). "Affiliating and disaffiliating with continuers: prosodic aspects of recipiency." In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation. Interactional Studies* (pp. 131–176). Cambridge: Cambridge University Press.

Ochs, E. (1979). "Transcription as theory." In E. Ochs & B. Schieffelin (Eds.), *Developmental Pragmatics* (pp. 43–72). New York: Academic Press.

Sacks, H. (1992). *Lectures on Conversation,* Volumes I & II. Gail Jefferson (Ed.). Oxford, UK: Blackwell.

Sacks, H., Schegloff, E., & Jefferson, G. (1974). "A simplest systematics for the organization of turn-taking for conversation." *Language, 50,* 696–735.

Schegloff, E. A. (1979). "The relevance of repair to syntax-for-conversation." In T. Givon (Ed.), *Syntax and Semantics 12: Discourse and Syntax* (pp. 261–288). New York: Academic Press.

Schegloff, E. A. (1987). "Analyzing single episodes of interaction: An exercise in conversation analysis." *Social Psychology Quarterly, 50,* 101–114.

Schegloff, E. A. (1996). "Turn organization: one intersection of grammar and interaction." In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and Grammar* (pp. 52–133). Cambridge: Cambridge University Press.

Schegloff, E. A. (1998). "Reflections on studying prosody in talk-in-interaction." *Language and Speech, 41* (3–4), 235–263.

Schegloff, E. A. (2000). "Overlapping talk and the organization of turn-taking for conversation." *Language in Society, 29*, 1–63.

Schegloff, E. A. (2002). "The surfacing of the suppressed." In P. Glenn, C. Le Baron, & J. Mandelbaum (Eds.), *Studies in Language and Social Interaction* (pp. 241–262). Mahwah, NJ: Erlbaum.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). "The preference for self-correction in the organization of repair in conversation." *Language, 53,* 361–382.

Schenkein, J. (1978). "Sketch of an analytic mentality for the study of conversational interaction." In J. Schenkein (Ed.), *Studies in the Organization of Conversational Interaction* (pp. 1–6). New York: Academic Press.

Schulze-Wenck, S. (to appear). "Form and function of 'first verbs' in talk-in-interaction." In A. Hakulinen & M. Selting (Eds.), *Syntax and Lexis in Conversation*. Amsterdam: Benjamins.

Selting, M. (1992). "Prosody in conversational questions." *Journal of Pragmatics, 17,* 315–345.

Selting, M. (1994). "Emphatic speech style – with special focus on the prosodic signaling of heightened emotive involvement in conversation." In C. Caffi & R. W. Janney (Eds.), *Involvement in Language*, *Journal of Pragmatics, 22,* 375–408.

Selting, M. (1995). *Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation*. Tübingen: Niemeyer.

Selting, M. (1996). "Prosody as an activity-type distinctive cue in conversation: The case of so-called 'astonished' questions in repair initiation." In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation: Interactional Studies* (pp. 231–270). Cambridge: Cambridge University Press.

Selting, M. (2000). "The constructions of units in conversational talk." *Language in Society, 29,* 477–517.

Selting, M. (2001). "Fragments of units as deviant cases of unit production in conversational talk." In M. Selting & E. Couper-Kuhlen (Eds.), *Studies in Interactional Linguistics* (pp. 229–258). Amsterdam: Benjamins.

Selting, M. (2003). "Lists as embedded structures and the prosody of list construction as an interactional resource." *InLiSt, 35*. http://www.uni-potsdam.de/u/inlist/

Tarplee, C. (1996). "Working on young children's utterances: Prosodic aspects of repetition during picture labelling." In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation: Interactional Studies* (pp. 406–435). Cambridge: Cambridge University Press.

Tomasello, M. (1998). "Introduction: The cognitive-functional perspective on language structure." In M. Tomasello (Ed.), *The New Psychology of Language: Cognitive and Functional approaches,* Vol. 1 (pp. 1–25). Mahwah, NJ: Lawrence Erlbaum.

Tomasello, M. (2002). "Introduction: Some surprises for psychologists." In M. Tomasello (Ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, Vol. 2 (pp. 1–14). Mahwah, NJ: Lawrence Erlbaum.

Uhmann, S. (1996). "On rhythm in everyday German conversation: Beat clashes in assessment utterances." In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation: Interactional Studies* (pp. 303–365). Cambridge: Cambridge University Press.

Uhmann, S. (1997). *Grammatische Regeln und konversationelle Strategien. Fallstudien aus Syntax und Phonologie.* Tübingen: Niemeyer.

Wells, B., & Peppé, S. (1996). "Ending up in Ulster: prosody and turn-taking in English dialects." In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation. Interactional Studies* (pp. 101–130). Cambridge: Cambridge University Press.

Wells, B., & Macfarlane, S. (1998). "Prosody as an interactional resource: turn-projection and overlap." *Language and Speech, 41,* 265–298.

# Practices and resources for turn transition

# Non-modal voice quality and turn-taking in Finnish

Richard Ogden

Non-modal voice quality (NMVQ) is used turn-finally in Finnish as part of a set of linguistic practices to mark relevant turn transition. Towards the end of a turn at talk, the current speaker regularly changes the phonatory setting to non-modal without resetting it to modal; incoming talk is regularly placed just after or in overlap with a non-modal stretch. Turn transition can occur in the absence of non-modal voice quality; and non-modal voice quality can occur without turn transition; but in both cases, there is evidence that participants orient to NMVQ as normative. The chapter also explores the relation of NMVQ and intonation. The chapter makes a contribution to the literature on the linguistic deployment of NMVQ.

## 1.  Introduction

The regulation and management of turn-taking in conversation is an issue of fundamental importance in the analysis of talk-in-interaction. In order to develop a phonology for conversation, an understanding of the phonetic resources that are available to speakers, alongside the sequential, interactional, and other linguistic resources (such as syntax and pragmatics) is essential. Much previous work assumes that intonation plays a primary role; one consequence of this assumption is that the role of voice quality has tended to be left aside. This chapter considers voice quality in the management of turn-taking in Finnish. It is argued that changes in voice quality from modal to non-modal are systematically deployed in the signaling of transition relevance.

In this section, I firstly consider non-modal voice quality from a linguistic perspective. I then look at what has been shown by conversation analysis and interactional linguists about the organisation of turn-taking; and then motivate the work described in this chapter.

In modal phonation, the vocal folds vibrate periodically along their full length due to pressure below the glottis which is higher than the pressure above the glottis. Modal phonation involves only moderate tension across the vocal folds. Although modal phonation can be given a more or less rigorous articulatory definition, it is frequently thought of as the 'normal' mode of phonation. Non-modal voice qualities (NMVQs), as is implicit in the name, have different modes of vibration, involving different degrees of tension across the vocal folds, and differences in which part of the vocal folds vibrate.

In the data discussed in this chapter, three NMVQs are particularly prominent: creak, breathiness and whisper. During creak, the frequency of the glottal pulses is very low, and frequently irregular. The mechanisms by which it is produced are a matter of debate, and do not concern us here. Breathiness involves a degree of leakage across the glottis: in the opening and closing cycle, the vocal folds do not make a complete closure. Whisper involves turbulent airflow across the glottis, and no regular vocal fold vibration.[1]

Cross-linguistically, non-modal voice qualities (NMVQ) have a variety of functions, and their phonetic extents and phonological domains likewise vary. Let us take creaky voice as an example. Its potential for lexical contrast is discussed in e.g. Ladefoged and Maddieson (1996: 317ff.) and Gordon and Ladefoged (2001). In German, it may be implicated in morphological distinctions e.g. the present vs. past tense distinction as in *können*, [kønn̩], (be able-3PL) and *könnten*, [kønn̩], (be able-SUBJ-PST-3PL) (Kohler 1999); and in many varieties of English, it is one phonetic exponent of the voicing contrast (e.g. Docherty & Foulkes 1999). Creak is frequently a variant of glottal stops, and occurs in the onset of vowel-initial words of many languages, including Finnish (Lehiste 1965; Ogden 1996) and English (Dilley et al. 1996). Thus voice quality has been shown to be linguistically contrastive in several distinct linguistic systems in a variety of languages.

Non-modal voice qualities are also frequently said to relate to sociolinguistic categories. For instance, creak has been claimed to index speaker gender (Henton & Bladon 1988) and dialect (Henton & Bladon 1988). It is also ascribed paralinguistic functions in many languages: Laver (1994: 196), Ní Chasaide and Gobl (1997: 456f.), Cruttenden (1997: 174) and Wichmann (2000) all give some indication of a relationship between voice quality and a speaker's inferred attitude or stance, or emotional state. However, the evidence for such paralinguistic functions is much harder to ascertain empirically, and it is common for analysts to draw on native speaker intuition rather than on the participants' own demonstrable orientation.

When it comes to the role of voice quality in the management of turn-taking, knowledge is rather fragmentary, even for English. Using read speech, Pierrehumbert and Talkin (1992), Pierrehumbert (1994), Dilley et al. (1996) all show that in English, creak is distributed around the margins of intonational phrases or pitch accents; there is also a consistent relation between the amount of creak and the level of the phrase, with major boundaries being signaled by stronger creak. Also using read speech, Redi and Shattuck-Hufnagel (2001) make a range of observations on the distribution of creak in American English. Among their findings are that words at the ends of utterances have a higher rate of glottalization than utterance-medially. On this basis, they argue that there is an association between utterance-finality and glottalization in American English. Likewise, Catford (1977), Laver (1980, 1994), Wells (1982), and Wichmann (2000) among others, have suggested (mostly on the basis of informal observation, though over a range of types of spoken material) that creak may be used towards the ends of utterance in British English; Laver (1994:196) goes further, and claims that some speakers use creaky phonation to signal floor-yielding in English.

With respect to languages other than English, even less is known about how NMVQ is deployed in the management of turn-taking. Turn-taking is implicitly understood to be one of the functions of intonation, but knowledge about the intonation systems of Finnish has been described as "more or less fragmentary" (Iivonen 1998:317). Iivonen et al. (1987), Iivonen (1998) and Välimaa-Blum (1993, 1999) agree that the basic pattern of Finnish is falling. According to Iivonen (1998:317) "in utterance-final positions creaky voice very often occurs", and this is associated with the "terminal intonation" of "the ends of final statements". Thus what is known about voice quality in Finnish is rather sketchy.

Linguistic research, then, provides evidence for non-modal voice qualities being potentially contrastive at a range of levels, from lexical through morphological, up to the level of the utterance. However, empirical knowledge about how voice quality might relate to turn-taking is limited, even for a well-researched language like English.

Let us now consider what is known about the linguistic resources used in the management of turn transition. Previous research has shown that transition relevance is normally produced and oriented to holistically, using syntactic, pragmatic and phonetic resources. Ford and Thomspon (1996) argue that turns at talk are produced and heard as wholes. This means that turn transition is relevant when a constellation of criteria are met: typically, there are three criteria for completion. Firstly, syntactic completion: a turn is treated

as finished if any major phrase boundaries are reached. Secondly, pragmatic completion: turns are treated as complete when the action they promote (such as greeting, assessing, responding, etc.) is done. Thirdly, prosodic completion: turns are treated as complete when they have the phonetics of finality, which for English includes such factors as slowing down and a boundary tone. Selting (2000) explores the relationship between turns, turn construction units (one or several of which may constitute a turn) and transition relevance places in more detail, focusing on the relation between syntactic and prosodic resources in German. She also argues that transition relevance is produced and oriented to holistically, and that there is a close association between the projectability of syntactic completion and prosodic completion, although syntax occasionally is decoupled from prosody, and *vice versa*. Thus there is no single factor which projects completion in a given turn at talk.

Other research concentrating more specifically on the phonetic resources available to speakers to manage turn transition has demonstrated the importance of prosodic features in signaling TRPs in various varieties of English. One major finding of this work is that typically, bundles of prosodic features, including pitch and voice quality, function together. Wells and Peppé (1996), in considering the prosodic resources used for turn-taking in Ulster English, consider pitch, tempo, loudness, duration and other factors. They compare their findings for Ulster with those of Local et al. (1985) for London Jamaican English and Local et al. (1986) for Tyneside English. They note that creaky voice is found on the last syllable of a turn in London Jamaican, but not in the other dialects. This finding shows that phonetic resources for turn-taking vary between varieties, and therefore are linguistic properties. The implication of this is that voice quality may be deployed differently in different linguistic communities.

To summarize the points made above:

– Linguistic accounts of voice quality imply that it may be a resource for managing turn-taking, but little empirical evidence for this has been given. This is at least partly because the data used in many studies is taken from monologues or read data, in neither of which turn-taking is an issue.

– The projection of relevant turn-transition is accomplished holistically, using syntactic, pragmatic and phonetic resources in tandem. An account of turn-taking should be sensitive to participants' holistic orientation to a bundle of features.

– Little is known about the deployment of phonetic resources cross-linguistically in the management of turn-taking. Work on English shows con-

siderable variation across varieties – we can therefore expect there to be substantial cross-linguistic differences.

This chapter, then, focuses on non-modal voice quality in Finnish. The claim made is that a change to NMVQ constitutes part of a normative pattern for signaling relevant turn transition in Finnish. This means that turn-finality is usually marked with a stretch of NMVQ towards the end of the turn. Where turn transition occurs without a change to NMVQ, the participants display an orientation to the normative nature of the marking of turn-finality by engaging in some other, more marked, practices.

The chapter is structured as follows: Section 2 gives an overview of the data and methodology used. Section 3 presents an overview of the findings, and then discusses some canonical cases where a change NMVQ is followed by speaker transition. Sections 4 and 5 discuss apparently deviant cases which provide evidence for participants' orientation to NMVQ as the norm for marking transition relevance: Section 4 provides examples where transition relevance is marked by NMVQ, but there is no change of current speaker; Section 5 discusses a case of speaker transition not marked by NMVQ. Section 6 briefly considers the intersection of voice quality and intonation in Finnish. Section 7 presents the conclusions.

## 2.    Data and methodology

The data in this paper are taken from radio phone-in programs broadcast on Finnish national radio and recorded in April–June 2000. Listeners call in and ask for a piece of music to be played. In most of the calls analyzed, there are two presenters (one male, one female), who encourage the callers to talk about why they have chosen that piece, and they usually develop the conversation so as to inform the listeners about the musicians or the music. Each presenter takes it in turns to take a call. Sometimes, the other presenter joins in with a call. Although each call has a similar overall structure, the content varies widely, and the corpus contains examples of many kinds of activity: complaining, telling stories, making requests, and so on.

The speakers in the calls are both males and females from all over Finland, and a range of dialects is spoken in the recordings. The heterogeneity of the data is not problematic for the analytic claims made here: there seem to be no substantial differences between the speakers in the distribution of non-modal voice qualities, despite rhythmical, intonational and morphological dif-

ferences. (Speakers' usual voice quality is variable, but despite this variability, they still modulate voice quality turn-finally.) The analysis of the relevance of voice quality in the organisation of turn-taking holds across all the speakers equally well.

Ten calls of 2:00–2:30 minutes were analyzed instrumentally and auditorily – approximately 23 minutes in total. The transcription scheme is a form of modified orthography which captures some prosodic features of spontaneous talk. In the transcriptions, P stands for the main presenter for the call, P2 for the other presenter, C for the caller. Non-modal phonation is transcribed using the conventions of ExtIPA (IPA 1999). Longer stretches of non-modal phonation are surrounded by curly braces, { }, and a capital letter is used to indicate the voice quality of that stretch: C for creak, B for breathiness, W for whisper, H for voicelessness. (The full set of conventions is presented in the Appendix.)

Non-modal phonation was transcribed when it was both instrumentally (i.e. in a waveform and/or on a spectrogram) and auditorily observable. The technique is very similar to that reported by Redi and Shattuck-Hufnagel (2001), although different kinds of creaky phonation were not distinguished. This means that it is not possible on the basis of the data presented in this paper to make any claims about the linguistic status of e.g. alternation in the shape, amplitude or duration of successive peaks (diplophonia) vs. lowering of f0 with near-total damping, which are two of the four categories which Redi and Shattuck-Hufnagel distinguished. These researchers found that there was both inter- and intra-speaker variability in the ways that creak was produced. It is not possible to comment whether the same findings would hold for Finnish on the basis of the data in this paper.

NMVQ is used turn-initially and turn-medially in Finnish. One of the functions of creak, for instance, is to mark word juncture (cf. Lehiste 1965; Ogden 1996). In this paper the data is constrained to non-modal phonation types at possible turn endings where the phonatory setting does not return to modal once NMVQ is initiated. NMVQ in this syntagmatic context, I will argue, plays a part in the turn-taking system of Finnish.

## 3.   Analysis

In this section, I present the results of the analysis of the data. I start by giving an overview of the general patterns, and then look in more detail at canonical cases, where a TRP is marked by NMVQ and is followed by speaker transition. Subsequent sections consider apparently deviant cases.