

Units of Talk – Units of Action

Studies in Language and Social Interaction (SLSI)

Studies in Language and Social Interaction is a series which continues the tradition of *Studies in Discourse and Grammar*, but with a new focus. It aims to provide a forum for research on grammar, understood broadly, in its natural home environment, spoken interaction. The assumption underlying the series is that the study of language as it is actually used in social interaction provides the foundation for understanding how the patterns and regularities we think of as grammar emerge from everyday communicative needs. The editors welcome language-related research from a range of different methodological traditions, including conversation analysis, interactional linguistics, and discourse-functional linguistics.

Editors

Sandra A. Thompson
University of California, Santa Barbara

Elizabeth Couper-Kuhlen
University of Helsinki

Editorial Board

Peter Auer
University of Freiburg

Paul Drew
York University

Cecilia E. Ford
University of Wisconsin

Barbara A. Fox
University of Colorado

Marja-Liisa Helasvuo
University of Turku

K.K. Luke
Hong Kong University/Nanyang
Technological University

Volume 25

Units of Talk – Units of Action

Edited by Beatrice Szczeppek Reed and Geoffrey Raymond

Units of Talk – Units of Action

Edited by

Beatrice Szczeppek Reed

University of York

Geoffrey Raymond

University of California at Santa Barbara

John Benjamins Publishing Company

Amsterdam / Philadelphia



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

Library of Congress Cataloging-in-Publication Data

Units of talk - units of action / Edited by Beatrice Szczepiek Reed, Geoffrey Raymond.

p. cm. (Studies in Language and Social Interaction, ISSN 1879-3983 ; v. 25)

Includes bibliographical references and index.

1. Speech acts (Linguistics)
2. Oral communication.
3. Social interaction.
4. Sociolinguistics. I. Szczepiek Reed, Beatrice, 1973- editor of compilation. II. Raymond, Geoffrey, editor of compilation.

P95.55.U55 2013

306.44--dc23

2013027064

ISBN 978 90 272 2635 8 (Hb ; alk. paper)

ISBN 978 90 272 7131 0 (Eb)

© 2013 – John Benjamins B.V.

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

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

The question of units for language, action and interaction <i>Beatrice Szczeppek Reed & Geoffrey Raymond</i>	1
PART I. Units of language revisited	
Units and/or Action Trajectories? The language of grammatical categories and the language of social action <i>Cecilia E. Ford, Barbara A. Fox & Sandra A. Thompson</i>	13
The dynamics of incrementation in utterance-building: Processes and resources <i>Per Linell</i>	57
From “intonation units” to cesuring – an alternative approach to the prosodic-phonetic structuring of talk-in-interaction <i>Dagmar Barth-Weingarten</i>	91
Perception of prosodic boundaries by untrained listeners <i>Brendan Barnwell</i>	125
PART II. Units of action and interaction	
At the intersection of turn and sequence organization: On the relevance of “slots” in type-conforming responses to polar interrogatives <i>Geoffrey Raymond</i>	169
When ‘yes’ is not enough – as an answer to a yes/no question <i>Jakob Steensig & Trine Heinemann</i>	207
Emerging units and emergent forms of participation within a unit in Japanese interaction: Local organization at a finer level of granularity <i>Shimako Iwasaki</i>	243
Phonetic resources in the construction of social actions <i>Richard Ogden & Traci Walker</i>	277

Building an instructional project: Actions as components of music masterclasses	313
<i>Darren Reed & Beatrice Szczeppek Reed</i>	
Language and the body in the construction of units in Mandarin face-to-face interaction	343
<i>Xiaoting Li</i>	
Index	377

The question of units for language, action and interaction

Beatrice Szczeppek Reed & Geoffrey Raymond

University of York / University of California at Santa Barbara

1. Introduction

This volume addresses two fundamental questions: What units of conduct are relevant for the study of (inter-)action? And are traditional linguistic units relevant for analyses of talk-in-interaction, and if so, how? The first question concerns the way in which human beings compose their conduct in recognizable ‘chunks’, how this structuring enables the sense recipients make of actions and activities, and how it features in the methods that participants, collectively, use to distribute opportunities for action in interaction. While language furnishes key practices for the formation of actions, it is only one of several resources participants draw on in composing them. As a consequence, it may be useful to distinguish between (traditionally defined) linguistic units on the one hand, and units-of-action on the other, including the elements or resources used to compose them, as well as the range of other units to which such actions contribute – such as a course of action, a “complete-able” project (Lerner 1995), or the overall structural organization of a single conversation (Schegloff 2011).

The second question concerns concepts that have been advanced in Philology and Linguistics, and have resulted in an almost universally accepted terminology of language units, such as ‘sentence’, and categories, such as ‘noun’. However, as research on language-in-interaction is beginning to show, not all of these units and categories adequately capture the empirically observed realities of language use, and many of them need to be modified, at least in part because traditional analyses have focused on their internal organization, at the expense of examining the ways in which their composition reflects various aspects of the social occasions of their use. Take, as an example, the unit to which traditionally defined units are most often compared: turns at talk. While many turns may come to be completed as sentences, such post-hoc analyses overlook the contingent, real-time interaction between speaker and hearer from which the form emerged. A different view of such units is suggested by analyses that focus on the methods speakers use to compose (spoken) contributions to unfolding

interactions in light of the contingencies associated with the local distribution of opportunities to speak next. In this view “turn-constructual units”, or TCUs, are basic components of a turn taking system for conversation (Sacks et al. 1974). As with other interactional units, turns, and the TCUs out of which they are composed, have an “overall structural organization” that parties can use to (reflexively) track their “local realization” in occasions of interaction (Schegloff 2011, 367). And just as TCUs can be used to compose turns at talk, turns can be used to compose other units, such as adjacency pairs, “a basic unit of sequence construction” (Schegloff 2007, 9). As a result, analysts interested in language as a resource for conversation “need to hold loosely (their) conceptions of structure, rule, and unit” (Ford 2004, 48), and keep in mind the flexibility of language as a resource for dealing with systematic contingencies associated with local, real time organization of action in interaction (as illustrated early on by Goodwin’s (1979) analysis of “the interactive construction of a sentence”). This focus helps guard against the analytical danger identified by Ford (2004), and elaborated on by Ford, Fox and Thompson (this volume) and Linell (this volume), that “the drive to define units may cause us to miss systematic practices that make conversation work for participants in real contexts of use” (Ford 2004, 38).

Adopting such a perspective of participant orientation to units of talk and their relation to actions, the chapters in this volume explore, amongst other things, what types of stretches of talk are treated by conversationalists as holistic entities, and whether there are previously un-described units that are relevant for talk-in-interaction. Some chapters ask how traditionally accepted linguistic units and their boundaries are realized and oriented to in conversation. Many chapters in this volume take into consideration linguistic and non-linguistic modes of interaction, and some show specifically how the analysis of different modes influences how units are perceived and constructed. In the remainder of this chapter the conceptual framework for this volume will be introduced, before the individual contributions are briefly summarised.

2. Conceptual framework

Before the issues outlined above can be addressed in the following chapters, three fundamental differences need to be acknowledged between the linguistic approaches that postulated the basic language units we use today for language study and language teaching; and the empirical approach of conversation analysis and interactional linguistics adopted by the contributors to this volume. These differences are not problematic in themselves; however, as traditional linguistic concepts and terminology dominate most people’s exposure to language study, for example, when they are learning a second language any interactional treatment of ‘units’ must start with an explicit discussion of how such a treatment differs from existing approaches.

2.1 The 'natural habitat' of language

The first difference between traditional Linguistics and Conversation Analysis concerns what is conceived of as the 'natural habitat' of language. Most linguistic approaches would consider the home environment of language to be either the brain or the mind. For example, Psycholinguistics, Neurolinguistics and Cognitive Linguistics evidently focus on the brain as the primary source of organization for language use, while Generative Grammar, with its interest in mental representations of grammar, treats the mind as the natural home of language structure. In the case of cognitive approaches, language production, perception and learning are studied empirically, primarily via experiments. While the main focus is on psychological, and therefore individual aspects of language processing (see next section), the brain activity of real-life experimental subjects is treated as the basis for investigations of language. By contrast, for generative approaches the natural habitat of language is instead the mind of an imagined, faultless native speaker. This notion has been criticised extensively in World Englishes research, where the 'native speaker' has long been considered an out-dated concept with no basis in reality: "The 'native speaker' of linguists and language teachers is in fact an abstraction based on arbitrarily selected features of pronunciation, grammar and lexicon" (Kramsch 1998, 79). Bhatt (2002), in his discussion of standard and non-standard uses of English, appropriates Anderson's (1991) concept of the "sacred imagined community" to the concept of the native speaker, showing that notions of what is or is not 'correct' language use are socially, and even ideologically accomplished. Without explicitly contributing to this area of research, the conversation analytic endeavour and its findings underpin this argument. As many contributions to this volume show, the language human beings use when they are engaged in what language evolved for – naturally occurring interaction between conspecifics – is often inadequately captured by traditional linguistic concepts, many of which can be shown to be "imagined", however "sacred" they may have become (see Ford, Fox & Thompson this volume).

In contrast to the linguistic approaches mentioned above, Conversation Analysis has argued that conversational interaction constitutes the primary ecological niche within which grammar and other resources for producing spoken language, such as prosody, word selection, and the like, have emerged (see Schegloff 2005, 2006). As a consequence, what we think of as grammar and other aspects of language can be understood as collections of deeply routinized practices (i.e. taken for granted solutions or methods) for managing the systematic contingencies associated with turn taking in conversation, and the use of talk in producing *action-in-interaction* more generally. In this view grammar constitutes a sequentially sensitive resource in that the deployment of units will project, among other things, a turn's course and duration in light of the specific sequential context in which it is used. Thus, although speakers compose utterances out of units that vary in length from a single word to a complete

‘sentence’ – and regularly build utterances that include more than one such unit, or only parts of them – participants (speakers and recipients) can nevertheless independently coordinate various forms of participation by reference to such units because their in-progress production projects the imminent possible completions they will arrive at (cf. Schegloff 2007; Lerner 1991, 1996) As Raymond and Lerner (forthcoming: 27–28; emphasis in original) observe:

When one initiates a turn at talk, the unfolding turn-so-far will project roughly what it will take to complete it. Moreover, the continuing moment-by-moment unfolding of a turn will be inspected for the progressive realization (suspension, deflection, or abandonment) of what has been projected so far. The hallmark of this realization is found in such material elements as the pace of the talk, the adjacent placement of syntactically next words and the intonation contour that carries the talk. Moreover there are circumstances in which the forward progress of a speaking turn can be delayed or sped up or even abandoned, and a set of practices by which such disturbances to the normal or normative progress of a speaking turn toward possible completion are implemented; and these practices can furnish the resources for recipient action as well. In this sense the projectability of a speaker’s turn at talk constitutes a proximate normative structure within which a range of other organizational contingencies are coordinated and managed – including the timing and design of action by others; *it is precisely this progressively realized structure that makes any deflections in its locally projected course a site of action, a recognizable form of action, and a site of action and interpretation by others.*

The appreciation of such ‘units’ as socially organized, participant-administered structures has a range of consequences for our understanding of them. These include, most centrally, the questions we ask about them, such as: How do parties to a conversation manage how such units are distributed: who gets to produce units, when, and how many? And what are the basic constituents of such units? This includes both the material elements out of which turns are built (i.e. including words and their arrangement in grammatical forms, and the prosodic packaging used to carry them in talk), as well as the methods by which these resources socially organize their temporally unfolding character as vehicles for the production of action-in-interaction.

We can further ask, how do speakers rely on the overall structural organization of units of talk (e.g. with beginnings, middles and endings) to compose recognizable actions, and what can be accomplished via the reflexive exploitation of such units that these basic structures underwrite (e.g. beginning without a beginning, or ending without an ending, etc. Schegloff 1996). Or how can participants exploit local realizations of these structures to manage action *within* them? For example, analysts have emphasized the internal infrastructure such units can provide for the coordination of a range of actions within its boundaries, as in the case of choral co-production, the collaborative completion of in-progress turns, and other forms of “conditional entry” into a turn space (see Lerner 1991, 1996, 2002, 2004; Szczeppek Reed 2006; Iwasaki 2009).

Finally we may ask by reference to what orders of organization (e.g. turn organization, sequence organization, overall structural organization, etc.) are such units produced in conversation, and how are transitions to next speakers (or next units) locally managed by reference to those orders of organization? In prior work, analysts have specified a range of practices that have emerged to handle the various context sensitive contingencies that utterance composition and completion poses for the production and recognition of social action. As these practices suggest, any specific deployment of these units is both *socially* organized – insofar as the places at which an utterance can be recognized as possibly complete are shaped by resources that are partly independent of the participants or occasion, and interactively managed – insofar as a current speaker and a next speaker jointly coordinate when a unit currently-in-progress will be complete and transition to the next speaker commences. By virtue of these same contingencies the internal organization of such units can be understood to be shaped by the relevance of a next speaker – or the absence of one – in light of their sequential context. Analysts considering these matters have demonstrated that such locally managed, participant-administered transitions between one speaker and a next are a key site for action and interaction (cf. Goodwin 1980; Schegloff 1987; Lerner 2013), and it is in the light of this evidence that interaction is considered the ‘natural habitat’ of language.

2.2 Abstract monologue vs. real-life interaction

The fundamental conceptualisation of interaction as the natural ecological niche for language results in two further differences, which concern what is considered to be language at all. While some linguistic approaches are interested in, firstly, *monologue* and, secondly, *abstract representations*, others analyse instead *conversational* and *empirically observable instances of language use*. If language is conceived of as internalised within either the brain or the mind, the phenomenon of interest is likely to be monologic language, as produced by an (imagined) individual’s mind (see above); and it is necessarily abstracted from naturally occurring instances of language-in-interaction. Interestingly, both conversation analysts and linguists have referred to each other’s object of study as ‘epiphenomena’: Chomsky (1986, 25) famously called any form of externalised (as opposed to internalised) language an ‘epiphenomenon at best’. In this volume, Ford, Fox and Thompson refer to “the social life that gives rise to the epiphenomena that linguists call grammatical units” (p. 40).

Many students of linguistics are used to conceiving of language as a theoretical system with a life of its own, where constituents ‘move’ from here to there, and so-called ‘surface’ structures can be traced back to their true, ‘underlying’ form. Traditional approaches to linguistic units have also tended to view sentence structure as a property of language, rather than of social action, and thus as a matter best defined by the scholars who study it. Indeed, it is precisely this approach that

makes individual deployments of such units by speakers and writers answerable to the structures these scholars view as ‘underlying’ reality – rather than vice versa. While this approach has proven useful in specifying many basic elements of grammar (i.e. sentences and their constituents), it has also been criticized for reifying the sentence as a unit in those cases where analysts have treated it as a given, virtually platonic unit type (and thereby adopted a thoroughgoing ‘structuralist’ orientation to its explication). By treating the sentence as a unit of language per se, analysts adopting this approach have tended to ignore as anomalous defects the various ways in which the actual production of sentence-like-units in speech are often characterized by phenomena such as restarts, repetition, bits of silence, uhs, uhms and other disfluencies (cf. Schegloff 1979; Goodwin 1980). As Schegloff notes (1979, 1996), by disattending what are thereby treated as “mere disfluencies” in speech production, these analysts miss out on many of the ways that language users orient to and exploit such “regular” units of conduct as resources for the production and coordination of action (see especially Schegloff 1996).

By contrast, conversation analysts, interactional linguists and linguistic anthropologists, and other like-minded scholars, have viewed such units as resources that *members* use – and thus as forms to be *understood* by analysts, rather than defined by them. In pursuing these matters, then, conversation analytic research has revealed that some units that have been taken for granted in traditional studies of language (in both linguistics and sociology), may have to be re-specified, and/or re-defined in light of findings from studies of interaction. For example, the sentence has been shown to be ‘semi-permeable’ (Lerner 1991, 1996) from a participant perspective; that is, it is a matter of interactional co-construction, and sentence structures are under continuous negotiation. This different approach does not argue that sentences are not, to some degree, pre-structured syntactic units; rather, it places the structuring such units provide in the hands of the participants who use them. More generally, the claim made by Schegloff (1996) for TCUs holds for many, if not all units of interaction:

What sorts of entities (described in grammatical or other terms) will be used and treated as turn-constructive units is determined by those who *use* the language (broadly understood – that is, to include gesture, facial expression, when/where relevant), not those who study it academically. Calls for formal definitions of a TCU – beyond their status as units which can constitute possibly complete turns as above – are therefore bound to be disappointed, but empirical inquiries to explore such issues should be expected to yield interesting results. (Schegloff 1996, 115, emphasis in the original)

The chapters in this volume pursue empirical inquiries of this kind, treating real-life social interactions and the dialogic behavior of those who participate in them as

their primary object of interest. All discussions of ‘units’ that occur in the following chapters, whether their starting points are units of language or of social action, reflect this approach.

3. The chapters

The subsequent ten chapters are divided into two parts. **Part One** contains four chapters each of which makes radical suggestions regarding the relevance, form and conceptualisation of linguistic units for interaction and its analysis.

In Chapter 2, **Cecilia Ford, Barbara Fox and Sandra Thompson** put forward the argument that the terminology and concepts developed for the theoretical study of language are neither adequate nor appropriate for the study of naturally occurring interaction. The authors first show that early conversation analytic work was based on traditional linguistic concepts, and suggest that the CA approach itself demands that concepts be grounded in action, rather than theory. They go on to analyse sequences of actions without reference to linguistic units and categories, focussing instead on the particular actions as they emerge, advancing a “descriptive meta-language” for the study of social interaction. The chapter presents a radical “experiment” in basing an understanding of social meaning-making entirely on the observable behaviour of those who construct their own and others’ conduct as locally meaningful.

Per Linell applies a similarly radical re-orientation to the linguistic study of talk-in-interaction in Chapter 3. Linell suggests that naturally occurring language is “internally dialogical”, via a continuous process of “incrementation”. He argues that spontaneous talk can only be captured by a theory of “linguaging”, which is able to handle the processes and resources that constitute talk, rather than the units and rules that constitute theoretical linguistic concepts. With specific reference to a number of grammatical phenomena, such as pivot constructions, non-agreement with noun phrases and slips of the tongue Linell, too, suggests “a partly new meta-language” (p. 72) and a framework that can incorporate utterance building as a succession of “decision points and continuation types” (p. 72).

In Chapter 4, **Dagmar Barth-Weingarten** presents a newly-developed, original approach to the analysis of the phonological structure of naturally occurring language. While previous linguistic studies of the ‘intonation unit’ have primarily been concerned with defining the de-contextualised characteristics of the unit and its boundaries, Barth-Weingarten suggests that talk-in-interaction makes necessary the recognition of boundaries as gradient and “fuzzy”. She presents an analysis of naturally occurring phonetic-prosodic boundaries of chunks of talk, and shows that they vary in strength, which explains prior researchers’ difficulties in identifying intonation units

in spontaneous speech. She puts forward a theory of “cesuring”, which allows analysts to take seriously the complexity of the phonetic-prosodic structure of talk.

Brendan Barnwell’s Chapter 5 presents an experimental study of naïve listeners’ perceptions of prosodic boundaries. Like Barth-Weingarten, he argues that boundaries rather than internal unit structure should be the focus of intonation analysis. Barnwell’s overarching argument is that experimental research can complement a conversation analytic pursuit of participant orientation, by showing how ordinary listeners categorise specified phenomena when explicitly asked to do so. His findings show that in the case of prosodic boundaries, there is above-chance agreement on many boundaries, but listeners are far from being in total agreement. The results suggest a gradual transition from more to less clear boundaries, with the parameters and boundaries by which intonation units are identified showing varying degrees of distinctness, a finding that is consistent with Barth-Weingarten’s analysis.

Part Two contains six chapters, each of which addresses units drawn from the analysis of interaction. While some of the chapters make reference to linguistic units, their focus is the composition of actions and sequences.

In Chapter 6, **Geoffrey Raymond** draws on previous research on Yes/No Type Interrogatives (YNIs) to establish the relevance of “slots” as an analytic concept that captures the intersecting relevance of two orders of organization: turn organization and sequence organization. Specifically, Raymond shows that type-conforming responses to polar interrogatives can be internally structured into two “slots”, which satisfy – at times separately – the different constraints imposed by sequence organization on the one hand, and turn construction on the other.

Chapter 7 is also concerned with yes/no questions. In their analysis of Danish talk-in-interaction **Jakob Steensig** and **Trine Heinemann** show that after three specific interactionally-defined question types, yes or no are not satisfactory answers, but more interactional work is required (‘yes’/‘no’+). In cases where expansions are not provided, questioners elicit them via other means. Thus, a unit (‘yes’, ‘no’) that in some contexts might constitute a full TCU does not do so in these instances. The expansion slot is a clearly identifiable intra-turn location; however, the authors do not find specific actions or linguistic unit types that correspond to expansions of the three question types.

In her analysis of Japanese interactions in Chapter 8, **Shimako Iwasaki** shows that participants may halt their production of an ongoing turn in order to create “interactive turn spaces”, i.e. locations for others to co-participate. Iwasaki’s analysis presents a deliberate move away from a focus on turn/TCU completion and transition relevance places, and towards an understanding of TCUs as, firstly, collaboratively constructed spaces for action, and, secondly, constructed of sub-components. Sub-components are locally projected, which allows interactants to negotiate participation on a moment-by-moment basis.

In Chapter 9, **Richard Ogden and Traci Walker** present an analysis of offers and their phonetic exponents. Drawing on a previous analysis of three offer types in different sequential environments, the authors ask whether actions such as offers are systematically designed with recurring phonetic features. Their findings suggest that there are no offer-specific phonetic properties. Instead the phonetic features are employed to handle turn management and sequence organization issues, such as continuing talk, designing a turn as transition relevant, or showing affiliation with prior talk.

Chapter 10 by **Darren Reed and Beatrice Szczepek Reed** presents an analysis of larger interactional projects, specifically instruction sequences in music masterclasses. By detailing the action structure of masterclass instructions and their opening and closing boundaries, the authors argue that local actions and interactions are employed by participants to construct such larger projects. It is suggested that a primarily action-based analysis is more appropriate for an investigation into naturally occurring social conduct than a linguistically grounded one.

In Chapter 11, **Li Xiaoting** reports on her findings concerning multi-modal turn construction in Mandarin talk-in-interaction. Her analyses show how turns and TCUs are achieved through participants' orientation to body movements. In particular, Li shows how the "home-away-home" movement of the torso is employed for the construction of larger interactional projects, such as story telling, non-acceptance of a previous claim and subsequent account, and counter-argument in an argumentation sequence. The chapter presents Li's discovery of a systematic interrelation between the organization of body movements and the organization of the turn-at-talk as an interactional unit.

References

- Anderson, Benedict. 1991. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.
- Bhatt, Rakesh M. 2002. "Experts, dialects, and discourse." *International Journal of Applied Linguistics* 12 (1):74–109.
- Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin and Use*. New York: Westport.
- Ford, Cecilia E. 2004. "Contingency and units in interaction." *Discourse Studies* 6: 27–52.
- Goodwin, Charles. 1979. "The interactive construction of a sentence in natural conversation." In *Everyday Language: Studies in Ethnomethodology*, ed. by George Psathas, 97–121. New York: Irvington Publishers.
- Goodwin, Charles. 1980. "Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning." *Sociological Inquiry* 50 (3/4): 272–302.
- Iwasaki, Shimako. 2009. "Initiating interactive turn spaces in Japanese conversation: Local projection and collaborative action." *Discourse Processes* 46 (2/3):226–246.
- Kramsch, Claire. 1998. *Language and Culture*. Oxford: Oxford University Press.
- Lerner, Gene H. 1991. "On the syntax of sentences in progress." *Language in Society* 20: 441–458.

- Lerner, Gene H. 1995. "Turn design and the organization of participation in instructional activities." *Discourse Processes* 19: 111–131.
- Lerner, Gene H. 1996. "On the "semi-permeable" character of grammatical units in conversation: Conditional entry into the turn space of another speaker." In *Interaction and Grammar*, ed. by Elinor Ochs, Emanuel A. Schegloff and Sandra A. Thompson, 238–276. Cambridge: Cambridge University Press.
- Lerner, Gene H. 2002. "Turn-sharing: the choral co-production of talk in interaction." In *The Language of Turn and Sequence*, ed. by Cecilia Ford, Barbara Fox and Sandra A. Thompson, 225–256. Oxford: Oxford University Press.
- Lerner, Gene H. 2004. "Collaborative turn sequences." In *Conversation Analysis: Studies from the First Generation*, ed. by Gene H. Lerner, 225–256. Amsterdam: John Benjamins.
- Lerner, Gene H. 2013. "On the place of hesitating in delicate formulations: A turn-constructive infrastructure for collaborative indiscretion." In *Conversational Repair and Human Understanding*, ed. by Jack Sidnell, Makoto Hayashi and Geoffrey Raymond, 95–134. Cambridge: Cambridge University Press.
- Raymond, Geoffrey and Gene H. Lerner. forthcoming. Towards a Sociology of the body-in-action: The body and its multiple commitments.
- Sacks, Harvey, Emanuel A. Schegloff and Gail Jefferson. 1974. "A simplest systematics for the organization of turn-taking for conversation." *Language* 50: 696–735.
- Schegloff, Emanuel A. 1979. "The relevance of repair to syntax-for-conversation." In *Syntax and Semantics. Vol. 12: Discourse and Syntax*, ed. by Talmy Givon, 261–286. New York: Academic Press.
- Schegloff, Emanuel A. 1987. "Analyzing single episodes of interaction: An exercise in Conversation Analysis." *Social Psychology Quarterly* 50 (2):101–114.
- Schegloff, Emanuel A. 1996. "Turn organization: One intersection of grammar and interaction." In *Interaction and Grammar*, ed. by Elinor Ochs, Emanuel A. Schegloff and Sandra A. Thompson, 52–133. Cambridge: Cambridge University Press.
- Schegloff, Emanuel A. 2005. "On integrity in inquiry...of the investigated, not the investigator." *Discourse Studies* 7 (4–5): 455–480.
- Schegloff, Emanuel A. 2006. "Interaction: The infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is enacted." In *Roots of Human Sociality: Culture, Cognition and Interaction*, ed. by Nicholas J. Enfield and Stephen C. Levinson, 70–96. Oxford and New York: Berg Publishers.
- Schegloff, Emanuel A. 2007. *Sequence Organization in Interaction. A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.
- Szczepek Reed, Beatrice. 2006. *Prosodic Orientation in English Conversation*. Basingstoke: Palgrave MacMillan.

PART I

Units of language revisited

Units and/or Action Trajectories?

The language of grammatical categories and the language of social action*

Cecilia E. Ford, Barbara A. Fox & Sandra A. Thompson

University of Wisconsin-Madison / University of Colorado at Boulder /
University of California at Santa Barbara

First, the embarrassing question of units.

Erving Goffman (1981)[†]

Responding to Sacks et al.'s 1974 call for linguists to join in the study of resources for turn construction, the authors of this chapter long ago took on turn formulation as an issue which linguists must account for. In this chapter, we return to this aspect of CA's charge to linguists, noting that CA continues to borrow the meta-language of linguistic unit types which are based in a tradition that does not address the practices of humans in real-time and contingent social

* We are grateful to Brendan Barnwell, Dagmar Barth-Weingarten, Joseph Brooks, Joan Bybee, Irene Checa-Garcia, Elizabeth Couper-Kuhlen, Veronika Drake, Virginia Gill, Charles Goodwin, Marjorie Goodwin, Makoto Hayashi, Ritva Laury, Douglas Maynard, Lorenza Mondada, Bracha Nir, Simona Pekarek Doehler, Felicia Roberts, Richard Sandoval, Suzanne Stevenson, Beatrice Szczepek Reed, and Alex Wahl for valuable discussion of the issue of 'units' in interaction. We are particularly grateful to Geoffrey Raymond for his insightful input on the shape of this chapter. None of them is responsible, however, for the approach to 'units' taken here or for the way we may have interpreted their input in writing this chapter. Authorship is shared equally among the three of us.

[†] We include this quote lightheartedly, just as we believe Goffman intended it in the opening of a section of "Replies and responses". The section of that essay that he opens with this sentence is, however, nicely related to our concerns in the present chapter. Goffman discusses what term might be best to designate the bounded units that interactants use and recognize as resources in conversation. He rejects grammatical terms as being "responsive to linguistic, not interactional, analysis" (23), and he settles instead on the term "move", a term he prefers "not to fix very closely" (23). He characterizes a move in relation to the activity in which it is built to play a part: a move is a "a stretch of talk or its substitutes which has a distinctive unitary bearing on some set or other of the circumstances in which participants find themselves (some 'game' or other in the peculiar sense employed by Wittgenstein)." (24)

action. We experiment in grounding accounts of turn construction in action rather than linguistic-category types, offering two detailed analyses of utterances that emerge in ordinary interaction, avoiding dependence on linguistic categories. In line with longstanding trends in CA, we experiment in moving further toward a descriptive meta-language for turn construction based in the particulars of moments of naturally occurring interaction, with attention to vocal and embodied conduct of the multiple copresent participants.

1. Introduction

An abstract notion of projectable unit-types is central to Sacks, Schegloff and Jefferson's classic 1974 account for turn taking in interaction (hereafter 'Sacks et al.'). The authors are unequivocal in presenting the turn-constructive unit (TCU) as a fundamental component to account for interactants' ability to achieve the normatively smooth turn transitions that they so evidently do. Units, or "unit-types" (702), are thus at the very foundation of the turn-taking system as Sacks et al. describe it:

- Sacks et al. describe speakers as building their turns from among projectable unit-types whose trajectories are revealed bit by bit from their beginnings: "There are various unit-types with which the speaker may set out to construct a turn." (702).
- They point to grammatical resources as providing unit-types usable for supporting the projection of where a turn unit could come to possible completion before such completion is reached, defining unit-types for English in list fashion as "sentential, clausal, phrasal, and lexical constructions". (702)
- They emphasize the interactional nature of turn construction: observing that a turn is not unilaterally managed but rather an "interactional production". (726)
- Of particular relevance for linguists, they propose that "[h]ow projection of unit-types is accomplished, so as to allow ... 'no gap' starts by next speakers, is an important question on which linguists can make major contributions." (703, n. 12)

Each of the current authors entered enthusiastically into the CA dialogue precisely in response to Sacks et al.'s radical programmatic mandate to linguistic theory. By using the term "mandate", we index both the power that the CA method offers and the responsibility that the CA method placed upon us, a responsibility to bring it to, and integrate it with, linguistics. More specifically, in the late '70s and early '80s, we were among those linguists (see, e.g. Givón 1979, 1983, 1984; Hopper & Thompson 1980; Li 1976) struggling to bring balance and 'functional' explanatory grounding to a highly abstract and formally oriented linguistics. At that time, the intellectual context was one in which many linguists were deeply committed to modeling the abstract formal resources that could economically 'generate' the infinity of 'sentences' that native speakers of a given language would judge as grammatical. Our intention in this

paper is to probe the notion of syntactic categories used in CA, using the spirit of CA's empirical standards, its agnosticism with respect to abstract categories, and its methodological practice of grounding categories in the particulars of social actions as captured in audio and video recordings of naturally occurring interaction.

In earlier work, building upon and responding to research on turn construction, subsequent to the publication of Sacks et al., we jointly explored the possibility of rethinking the 'unitness' of TCUs in terms of practices (Ford et al. 1996). In that study we suggested that even the minimal unit which Sacks et al. so compellingly propose, i.e. the minimal first TCU a speaker gets when allocated a turn, is perhaps better understood not in terms of an inventory of structural or formal unit-types, but as a provisional and contingently unfolding projection. Thus a new speaker may claim an interactional space in which to produce a turn, but how participants shape this unfolding action is determined in an incremental manner and is susceptible to contingencies involving not just the speaker but other participants as well, as recipient actions affect turn trajectories in progress (as has been demonstrated by many researchers; see, e.g. C. Goodwin 1981, 1984; M. H. Goodwin 1980; Goodwin & Goodwin 1986; Schegloff 1987). We observed that interactional data do not unequivocally demand the postulation of an abstract inventory of *a priori* grammatical structures for turn construction. What studies in turn construction in interactional data do support, however, is the need for practices for formulating trackable trajectories.¹ These trajectories are provisional and malleable claims to interactional space in which speakers contingently produce spans of talk with beginnings, trajectories, and always-negotiable points of possible completion.

In referring to *a priori* grammatical structures, we invoke the work of Harris (2003) and Hopper (1988), and the distinction within linguistic theory between '*a priori*' and 'emergent' grammar. Hopper (2011) notes that linguists (and, we would add, Sacks et al. and virtually all students of talk-in-interaction, including ourselves) have generally adopted

the standard view that speakers of a language communicate by virtue of a uniform common grammatical system. Disagreement only occurred over the source of this grammar – discourse pragmatics or mental structures. The validity of this assumption, which by some has come to be called the *fixed code* (Harris 2003) or *a priori grammar* (Hopper 1988) theory, was rarely questioned; yet when examined, it was found to be full of paradoxes. (303)

In the current chapter, then, we draw attention once again to the importance of rethinking grammatical units in interaction. Given that CA is a radically bottom-up analytic approach to interaction, and given the compelling way that CA has, from

1. See Clayman (2013) and Drew (2013) for recent overviews of turn design and turn construction.

its origins in ethnomethodology, consistently questioned the importation of *a priori* (rather than participant-constructed and emergent) social categories to explain contingent, dynamic, and locally managed interaction, we treat linguistic categories (top-down notions derived primarily through introspective methods) with the same spirit. We interrogate the reliance on such concepts for arriving at accounts for the interactional construction of turns in the course of joint activities, be they sequences or courses of action. Aiming toward an action-based metalanguage to account for turn construction in sequential context, we present a data-based inquiry into the degree to which *a priori* linguistic categories and linguistic units are relevant and necessary to account for turn construction, and into what might be gained through mindful use of action-based descriptive language. Ideally, the CA and ethnomethodological charge that categories meet the evidentiary requirement that they be treated as real by the participants is one we believe linguists, and CA practitioners who use linguistic terminology, should aim to fulfill.²

Sacks et al. allude to, and explicitly state, the need for caution in importing linguistic categories (Sacks et al. 702–703, 720–722). However, it is also the case that, in presenting their empirical evidence of participants' orientations to unit-types (702–3, n. 12), Sacks et al. draw upon linguistic categories as unexamined givens. Most specifically, Sacks et al. offer examples of participants' treatment of grammatical units as possibly complete (702–703), but those unit-types are drawn directly and without reflection from an already existing taxonomy of traditional grammatical-unit categories: "sentential, clausal, phrasal, and lexical constructions" (702). That is, they use grammatical category names as theoretical primitives. By "primitives" we mean theoretical concepts that are to be understood as self-evident and derived from intuitively obvious and shared knowledge. As Sacks et al. state it (with our italics added on the linguistic-category terms):

Our discussion in §3.1 of the turn-constructural component of the turn-taking system identifies the types of turn-constructural units as *sentential*, *clausal*, *phrasal*, and *lexical*, i.e. syntactically. The discussion of *appositionals* and *tag questions* – and, most importantly, the way in which the prospect of turn-transfer at the first possible transition-relevance place conditions decisions as between *left-embedded and conjoined sentence structures* – should indicate the deep ways in which syntax matters to turn-taking, albeit a syntax conceived in terms of its relevance to turn-taking. (720–721)

2. At the end of their linguistics-based critique of the notion of unit, Taylor and Cameron (1987, 156) make a similar point: "What research into conversation shows is that many of the basic assumptions underlying the study of verbal interaction, including those supporting such traditional linguistic domains as syntax, have to be reconsidered; and this reassessment must be performed in the light of conversation analytic discoveries about what speakers and hearers really do with words, and not just what grammarians, following an in-built scriptist bias, have for centuries been telling us that they do."

As linguists and CA practitioners concerned to ground linguistic categories in the particulars of interaction, we respectfully submit that linguistic unit-types, useful as they may seem in providing initial and provisional leverage on what may underlie turn projection, are not exempt from the fundamental commitment to understanding the moment-to-moment, locally emerging trajectories that participants build and orient to as they collaboratively do action in interaction. Thus, in alignment with the empirical program of CA, we understand the linguistic unit-types and categories such as “apposition”, “tag question”, and “left-embedded structure”, as inherited from methods that are either introspective or aimed at cognitive explanations (or both), rather than derived from a commitment to understanding forms of social action. The “syntax for conversation” (Schegloff 1979) that we are pursuing can only be arrived at by subjecting such linguistic primitives to the same standard of scrutiny that ethnomethodologists and conversation analysts bring to bear on categories dear to quantitative sociology (e.g. race, gender, class, and the like), inherited as they are from traditions and methods distinct from ethnomethodology.

With respect to the continued relevance of grounding grammatical categories in social terms, we note that throughout the several decades of CAs development, grammatical forms have been regularly cited and used in CA research (including our own). When used, the categories and forms themselves have seldom been subject to question or revision from a social and interactional perspective.³ Indeed, the goal of discovering the nature of TCUs or the practices upon which they are based, beyond the listing of grammatical categories, seems to have been set aside as a problem for linguistics rather than a challenge for conversation analysts to legitimately address. As a recent example, in a useful introductory CA textbook, Ian Hutchby and Robin Wooffitt (2008) affirm that CA is fundamentally concerned with

the nature of turn taking: how is it organized, how do participants accomplish orderly (or even apparently disorderly) turn taking, and what are the systematic resources which are used in this accomplishment? (2008, 41)

They go on to explain that TCUs “broadly correspond to linguistic categories”, listing the same grammatical structures as do Sacks et al. (1974). However, Hutchby and

3. But see Schegloff's (1996a, 450) distinction between “locally initial and locally subsequent reference” formulations (in line with Fox 1987), a distinction that could be described in linguistic terms alone (e.g. ‘Full NP’ vs. ‘Pronoun’). To us this represents precisely the sort of move toward a more action-based metalanguage that we advocate in this chapter, particularly when taken along with Schegloff's consideration of mismatches between the general norm for reference and the interactional work such distinct formulations may be doing in addition to mere reference (1996a, 451–58).

Wooffitt are unequivocal in treating the challenge of grounding TCUs as beyond the aims of CA:

[I]t is not part of the conversation analyst's aim to define in some abstract way, what a turn-construction unit is, as a linguist for instance may want to define what a sentence is. Conversation analysts cannot take a prescriptive stance on this question, because what a turn construction unit consists of in any situated stretch of talk is a members' problem. That is, such a unit is essentially anything out of which a legitimate turn has recognizably – for the participants – been built. (2008, 49–50)

While we thoroughly agree that the rules of prescriptive or even descriptive grammar do not provide appropriate grounding for a CA account of the social action of turn construction, this does not lead us to abandon the need to work toward socially grounded accounts of turn constructional practices and resources. Our interest, then, is in holding ourselves and others more accountable to a social action-based grounding for turn construction.

Since the original publication of Sacks et al.'s account for turn taking, a great deal of CA research has continued to augment our understanding of turn construction. In our analyses for this chapter, we include reference to such work where relevant. In particular, we attend to embodied conduct, an area of findings intricately related to the systematic coordination of actions in interaction.⁴ The current study takes prior CA research on turn taking as foundational for, and informing of, what we attempt here. We push a step further by explicitly calling into question the usefulness of uncritically adopting *a priori* grammatical units. Instead we model an attempt to replace grammatical terms with terms of social action, concepts grounded in the local particulars of unfolding interactions. In our experience, such care in the metalanguage we use in describing interaction serves as a heuristic, drawing us away from reliance on the terminology of grammatical units and drawing us toward actions and practices. We see this as analogous to CA's language of 'doing being' as a heuristic to help the analyst avoid mind-reading in accounts of interaction, thereby drawing attention to participants' displays and orientations (e.g. Sacks 1984).

4. See, among others, Auer (1992, 2005), Fox (2002), C. Goodwin (1979, 1984, 2000, 2007a, b), M. Goodwin (1980, 2006), Goodwin and Goodwin (1986, 1987, 1992, 2004), Hayashi (2003, 2005), Mondada (2006, 2007, 2011), Streeck (2002, 2009), Streeck and Hartge (1992), and Streeck and Knapp (1992).

2. An initial illustration

Let us begin by considering two alternative ways in which we could articulate an analysis, taking a case from Makoto Hayashi's research on Japanese face-to-face interaction. We first offer an analysis that uses traditional grammatical unit terms for the structures in the talk, and we then experiment with using language that supports a more action-oriented understanding of the unfolding talk, doing so in an effort to avoid abstract and static structural unit-types. If this can be successfully done, we will move toward more alignment with the CA perspective on what people *do* when they talk: after all, people are, in the first place, doing social actions, with grammatical patterns as epiphenomenal emergent structures⁵ from this behavior.⁶

Hayashi (2001, 2003) draws on a traditional linguistic category to describe an aspect of one speaker's talk, referring to it as a "postposition-initiated utterance": one speaker uses a "postposition" to launch a turn, thereby tying what she will continue to say to something said by an immediately prior speaker. Thanks to Hayashi's generous sharing of his data and analysis, we are able to give the context surrounding those lines as Extract (1). Hayashi (2001) attends to lines 23–24 of this extract to illustrate how the "postposition-initiated utterance" functions. Three young women, Aiko, Mami, and Sana, are talking about Mami's recent trip to Nepal and India. In lines 2–11, Aiko and Sana comment on the cost of the trip, and how *okanemochi* 'rich' Mami must be to make such a trip. In lines 13 and 17, Mami informs them that she doesn't spend her money on anything but travel.

(1) from Hayashi (2001, 319)

- 1 Mami: [() .]
- 2 Aiko: [okane (.) ikura] gurai kakatta?:
 money how.much about cost
 "How much did it cost?"
- 3 Mami: eetto ne:: sa::nju- yonjuu man::
 well FP 30 40 10,000
 "Let's see:, a little less than 300-400,000"

5. See also Hopper (2004, 153): "Grammar' is an epiphenomenon of frequent combinations of constructions. Because grammar is a result of interactions rather than a prerequisite to them, it is not a fixed code but is caught up in a continual process of local adaptation (emergence)".

6. We note that Szczepiek Reed (2012) makes a very similar point regarding phonological terminology, showing the value of talking of 'intonation phrases' as "action components".

- 4 jaku ka.
 a.little.less Q
 (yen), I guess."
- 5 Aiko: ⁰u:::n⁰ kekkoo
 quite
 "⁰Hmmmm⁰ (You're) quite..."
- 6 Sana: u:::[::n ()]
 "Yeah::::: ()"
- 7 Mami: [()]
- 8 Aiko: [okanemochi] ya na:.
 rich CP FP
 "...rich."
- 9 (.)
- 10 Mami: e:?
 "Huh?"
- 11 Aiko: [okanemo]chi ya na:.
 rich CP FP
 "(You're) rich."
- 12 Mami: [hhhhhhhh]
- 13 Mami: .hhh e datte tsukawahen mo::n. sore
 RC because don't.use FP that
 ".hhh Well, cuz I don't use money. For other
- 14 gura[i (shika).]
 about only
 than that."
- 15 Sana: [hhahh hhe][hh hh hh
- 16 Aiko: [⁰⁰u:::n⁰⁰
- 17 Sana: .hh
- 18 (0.3)
- 19 Mami: ryokoo no tame ni: tottearu kara:.
 travel LK sake PT save because
 "I save for travel, so..."
- 20 (0.3)
- 21 Aiko: ⁰u:::n⁰
 "⁰Mmhm⁰"
- 22 (0.5)

- > 23 Aiko: de! nan'nichikan gurai °sore tte.°
 and how.many.days about t_hat QT
 "And about how many days| was that
 (trip)?" |
 _____ |
 |
- > 24 Mami: ga tookakan.
 SP ten.days
 "For ten days."
- 25 (0.5)
- 26 Mami: ()
- 27 (0.5)
- 28 ?: ⁰u::[n⁰
 "°Hmmm°"
- 29 Aiko: [ja kekkoo suru n ya.
 then quite cost N CP
 "It's quite expensive then."

In response to Aiko's question at line 23, Mami begins her answer with *ga*, a form which linguists would unequivocally call a 'postposition', meaning that it normatively occurs immediately *after* a noun phrase within a single speaker's utterance. However, interestingly, what Mami does here is to *begin* her turn with this 'postposition', creating an utterance that, in traditional grammatical terms, would be an anomaly. To make sense of Mami's 'postposition-initiated' response in line 24, her recipients must have shared knowledge of *ga* as grammatically 'belonging' to, or being part of, something previously uttered.⁷ Because there is nothing previously uttered in Mami's own talk, her recipients must understand Mami's *ga* to be tied to something in Aiko's previous talk, and conclude that her turn-initial postposition and what follows it is tied to the item *sore* 'that one' in Aiko's utterance.⁸

Based on this analysis, relative to the norms for Japanese in use, we can represent the following abstract schema as relevant for, and drawn upon by, Japanese interactants:

- (2) [[Noun Phrase (*sore*)]_{NP} + postposition (*ga*)]_{Postpositional Phrase}

7. We know Aiko does make sense of Mami's utterance from her 'upshot' in line 29.

8. As Hayashi notes, the element *ga* would normally not occur following the quotative particle *tte*, so Aiko will not understand Mami's utterance as a continuation of her own utterance in line 23.

Using a 'shorthand', then, demonstrably useful in communicating with other members of our scholarly community, we can employ such terms as 'postposition' and 'nominal element', that is, linguistic unit-types. However, we note that this shorthand not only fails to capture the actions in which Aiko and Mami are engaged, but it fails to even capture the real-time tying and projecting work that Aiko and Mami are doing. And this is because these terms were arrived at in a grammatical tradition not aimed at understanding grammar in its social interactional context.

Let us, then, reformulate our description in a way that foregrounds practices rather than abstract unit types.⁹ For example, in place of the abstract unit 'nominal element', we can understand Mami's response at line 24 in terms of how she uses a familiar form, in local practice, to orient to the action of reference formulation.¹⁰ By beginning her turn with *ga*, Mami establishes a syntagmatic relation, tying her responsive action to a previous reference formulation. In this way, Mami suggests a link back to Aiko's previous action. To make sense of Mami's response, Aiko may also draw on the practice of understanding *ga* as tied to a previous reference formulation.¹¹ Thus, both Aiko and Mami seem to be oriented to *ga* as doing linking work, linking the current turn's action to some prior action (though, as we will note below, whether that prior action is done discretely with a 'nominal' or whether it is done with a more diffuse and malleable span of talk is not self-evident). In action terms, then, the work Mami does through the use of turn initial *ga* may not be well-captured by the traditional denotation of 'postposition' nor by the association with the discrete abstract unit-type

9. We note that Hayashi, both in the paper we are drawing from and throughout his career, has shared our interest in interrogating the notion of 'unit', but that, as we have done here, he made use in this paper of the 'shorthand' linguistic labels that we refer to in the conclusion to this chapter.

10. Levinson (2013:Section 3) noting that there are other 'doings' in turns besides their 'main job', suggests that these other 'doings' may often be done as 'off record' and are generally not explicitly responded to by recipients. In this chapter, we are not making a strong distinction between these two types of actions a turn may be doing. So when we say that Maureen's turn *the café de yin yang?*, is doing "reference formulation", we are not insisting that this is its 'main action'. This follows analyses such as those of Goodwin and Goodwin (1987), who demonstrate that recipients can be responsive to assessments within turns, even when those turns are primarily doing other actions and are responded to on other terms at their completion.

11. This is of course similar to what the word 'postposition' means to linguists. Our point is that without further explication, the term 'postposition' tends to evoke for most linguists the structural properties of a given single clause independently of the temporal and interactional exigencies of everyday talk, and would fail to account for the kind of tying to the previous speaker's turn that Mami accomplishes here. We do not object to the term per se, but to the unquestioning use of such terms, which has typically not taken account of the nature of talk-in-interaction.

‘postpositional phrase’, consisting of a noun phrase and a postposition. Aiko is thus able to interpret Mami’s *ga* as building upon *sore* ‘that one’ in Aiko’s own utterance, projecting that what Mami will say next will draw on that connection in building a responsive action.

In an analysis of the broader stream of participation, then, we note that the participants are involved in a number of sequentially and simultaneously unfolding actions:

- Aiko’s turn in line 23 is doing questioning (Ford 2010), requesting a specific temporal duration as a response (see Thompson et al. (frth.)), making it relevant and expectable that Mami will provide her with a response indicating a period of time.
- In doing questioning here, Aiko adds to an ongoing sequence in which Mami is telling of her recent trip to Nepal and India.
- Accordingly, Aiko’s questioning in line 23 begins with *de*, roughly translatable as ‘and’, which is routinely used by Japanese interactants to continue an agenda or activity made up of subsequent and related items (Sadler 2001, 2006); in the interactional context of (1), this *de* can be understood as sharing some features with the *and*-prefaced questions in interview sequences discussed by Heritage and Sorjonen (1994), which they analyze as implementing ‘agenda-based’ actions.¹² Thus, by starting out with *de*, Aiko is projecting that it will implement an addition to the larger activity she is carrying out of displaying interest in Mami’s account of her travels.
- Aiko brings her questioning action to a point of possible turn transition after the final particle *tte*, and Mami treats that action as complete by providing a no-delay, no-problem response to Aiko’s time-period question, the target turn in line 24.
- As Hayashi notes, to understand how Mami’s response is fitted to its position in the developing activity, i.e. how it forms a relevant responsive action, Aiko must also draw on a practice that she has often encountered in Japanese interaction, namely *ga* links back to some prior reference formulation. Mami’s action smoothly unfolds as responsive to Aiko’s turn. It begins with a ‘non-beginning’ item (Schegloff 1996a), so Aiko is alerted, from the outset of Mami’s turn, that it is built on something prior, perhaps in Aiko’s very own turn. As we noted just above, the ‘something prior’ is not entirely determined. Such uses of what are traditionally called ‘postpositions’, but at turn beginnings, are, as Hayashi puts it, “built off of, or on to, the preceding utterance, and draw on it as a resource for their construction and comprehension.” (338)

12. In Heritage and Sorjonen (1994), the ‘agenda’ derives from the institutional identity of the questioner (a home health worker), whereas we are suggesting that Aiko’s *de*-prefacing indexes an ‘agenda’ of hearing about Mami’s extensive travels. We thank Geoffrey Raymond for helpful discussion of this point.

- That is, in constructing a response to Aiko, Mami most immediately draws into use a token which is predominantly used immediately following a reference formulation within the same speaker's utterance. In using *ga* turn initially, Mami links back to, and builds from, a reference formulation, *sore* in Aiko's turn in line 23. Simultaneously, Mami appropriates Aiko's entire prior action and its formation, putting it to use in constructing her (Mami's) own current action. Through this tie and this continuation, Mami symbiotically transforms the *sore*, and the entire turn it was initially part of, to make it work as a component of her own response.¹³

In other words, we are suggesting that Mami's turn-initial *ga* guides Mami's recipient Aiko to reinterpret her own prior talk for what *ga* is building on, and for how *ga*, together with what follows, forms a responsive action. In terms of social action, *ga* is an interactional link, using *sore* as a fulcrum for relating the current turn's action to that of a previous turn. To account for Mami's use of *ga*, we evidently don't need a notion of the abstract unit 'postpositional phrase', itself internally structured through the combination of the abstract units 'noun phrase' and a 'postposition'. We have instead analyzed what Mami does in building her response to Aiko's question in terms of the unfolding social actions that form the functional foundation for what we define, in a post-hoc fashion, as abstract units. It is this kind of formulation, in terms of actions and trajectories rather than abstract grammatical units and categories, that we pursue in this chapter.

In this initial example, we have thus seen that we can account for one kind of recurrent social action without recourse to abstract grammatical categories derived from traditional linguistics. But what about the value of such categories for projecting possible turn completion? It is our position that characterizing a turn-constructional unit as having an independently projectable possible completion point based upon any fixed, decontextualized, and autonomous set of linguistic unit-types is not compatible with the highly localized and contextual nature of interaction, particularly as it is understood from a CA perspective. Sacks et al.'s (1974) account for the turn-constructional component of the turn-taking system allows the interpretation that TCUs exist independent of action context (702–703). Describing TCUs in terms of grammar, and later noting the importance of sound production (721), Sacks et al. do not foreground action context as part of turn projection or of the turn-constructional component. Thus, their account implicitly proposes projection to be based on a shared, abstract, and acontextual understanding of grammar and intonation, such that at the end of one thereby projectable TCU, a transition relevance place (TRP) occurs:

13. We can say that she is drawing into use a "practiced solution" to the management of this response at this moment (Schegloff 2006, 2007: Chapter 13). We thank Geoffrey Raymond for drawing our attention to Schegloff's articulation of this notion.

As for the unit-types which a speaker employs in starting the construction of a turn's talk, the speaker is initially entitled, in having a turn, to one such unit. The first possible completion of one such unit constitutes an initial transition relevance place. (703)

We aim to encourage a move toward describing action, rather than drawing from an inventory of previously defined abstract unit-types. For example, a particular string of words, such as English '*the editor*', when articulated in a particular context, may accomplish the action of referring to a non-co-present party rather than being categorized as a 'noun', or a 'noun phrase'. The word *the* may do the action of indexing that a reference formulation is in progress and that the reference will be one already shared among the participants (from the current interaction or from more generally shared social experience).¹⁴ Possible completion of that projectable trajectory of action would be limited by locally relevant potential objects of reference, the work that the reference formulation might need to accomplish to be taken as complete in the sequential context (e.g. is it a response? is it a part of a topic initiation? etc.), and the sound qualities and bodily movements with which the unfolding turn is produced.¹⁵

Returning to excerpt (1), Mami and Aiko's orientation to Mami's *ga*-initiated utterance can be understood from this perspective as well. Aiko's question makes relevant a temporal-duration response from Mami. As Mami's utterance is produced, Aiko is carefully monitoring it to determine how it will be that response. As we noted just above, socially, Mami's *ga* alerts Aiko that in order to project completion of this action, Aiko must 'back up' to reinterpret her own prior talk for what *ga* might be building on, and for how Mami's entire *ga*-initiated turn will unfold to constitute a possible response.

In this chapter, then, we hope to bring to the attention of linguists and conversation analysts our recognition that *a priori* linguistic units are not exempt from the fundamental commitment to understanding, in action terms, the practices that participants draw upon as they collaboratively construct action in interaction. On the contrary, the full power of the CA method *demand*s this move, both for the continued development of a socially grounded linguistic theory and for the grounding of CA in its own right.

14. The observant reader will note that, while arguing against the unquestioning use of linguistic-unit terminology, we will be using such terms as 'word' and 'velar obstruent'. We remain committed to subjecting such terminology to close scrutiny, but we also acknowledge that doing so may not always be feasible. Here, for instance, we judge that grounding the vernacular 'word' or the phonetic term 'velar obstruent' in action terms would take us far afield of our goal for this chapter.

15. We are fully in line with the linguistic and conversation analytic investigation of how action formation may become, as some linguists term it, "fixed" over time. Our argument is simply that those fixed forms are epiphenomenal and emergent at their core, based precisely on the tasks that humans do in social interaction.

Our interest, then, is in encouraging a move away from abstract formal unit-types imported from linguistics, convenient as they may be as provisional resources for analysts, as we try to account for the no-delay, no-gap nature of most turn transitions.

The body of this chapter is taken up with pushing the experiment begun in Extract (1) further by analyzing two extracts from a single videotaped interaction, and doing so without dependence on abstract grammatical units. In our final discussion, we evaluate this CA-inspired experiment in using a metalanguage of actions to account for turn construction.

3. Two cases

3.1 Case 1: The Café de Yin Yang

In our first case, we examine the very local emergence of a specific turn and how it fits within the larger flow of the interaction. Our analysis incorporates the fact that forms of visible and hearable conduct are mutually contextualizing, simultaneous aspects of turn construction (Bolden 2003; M. H. Goodwin 1980; Schegloff 1987; *inter alia*). That is, it has been well-established that during what could be characterized as verbal turns, the bodily actions of both speakers and recipients are calibrated with, and affecting of, one another. Thus, we attend as closely as we can to bodily actions, by the speaker and her recipients, in terms of both what precedes and what accompanies her talk. In our analysis of the turn's construction, as it develops in the context of already-in-progress bodily conduct and orientation to that conduct, we explore to what degree we can usefully use action terms, based in the particulars of the unfolding social interaction, in accounting for turn construction, rather than depending on abstract, autonomous grammatical unit-types without reference to social action.

We can transcribe in a single line the turn that we will argue is deeply embedded in its sequential context, including embodied actions:

(3) Maureen: the café de yin ya:ng? when he was tw- te:n?

This turn can be analyzed as composed of at least two “communicative acts” (Goodwin & Goodwin 1987, 18–19, n. 10), the acts themselves being formed up through specific vocal practices. There is clearly rising pitch and a sound stretch on *ya:ng*, and such prosody is often, though not always, associated with possible turn completion and turn transition.¹⁶ There is a very brief silence of 6/100ths of a second between the end

16. See, among others, Ford and Thompson (1996), Szczepek Reed (2004, 2010), and Ford et al. (2004).

of *ya:ng* and the start of *when*, and there is no release of the velar closure at the end of *ya:ng* until Maureen begins articulating the beginning of *when*.^{17,18} Even after what some might argue constitutes the completion of a projectable, abstract grammatical TCU, at the point where rising intonation and sound stretch accompany the delivery of *ya:ng* there is no attempt at turn transition and no visible or hearable orientation to its absence as accountable or problematic. On the contrary, it is not until the possible completion of *te:n* that a recipient initiates a spoken response, with no gap or overlap. Maureen is looking at a drawing on the wall labeled ‘The Café de Yin Yang’, drawn by the young nephew of one of the other participants:

- (4) Maureen: the café de yin ya:ng? when he was
 tw- te:n?
 Terry: yeah:.

As we will show, given the action context (viewed both locally and broadly), and with access to the embodied conduct before and during Maureen’s production of this talk, it is far from clear that Maureen’s turn is nearing possible completion as she produces the word *ya:ng*.¹⁹ On the other hand, if we were to consider the whole utterance as a complete turn, as does one recipient (Terry), then the second “communicative act” would be understood as an increment to the first, the two TCUs combining to form a complete TCU ending in a TRP.²⁰ In either analysis, the TRP at the end of *te:n* ends a stream of speech that does not constitute a grammatical unit in any linguistic sense.

Using grammatical unit-types, we could describe Maureen’s utterance (ending with *te:n*) as made up of a noun phrase (NP) followed by an adverbial clause (AdvCL):

- (5) [*the café de yin ya:ng?*]_{NP} + [*when he was tw-te:n?*]_{AdvClause}

This is a combination of grammatical units yielding an anomalous whole, a ‘non-unit’ within the inventory of such units as proposed and accepted in linguistics. That is, the

17. To our knowledge, no one working with these data has timed and represented this silence in a published transcript, though we ourselves have vacillated between putting both acts on a single line versus each on a separate line. Such transcription decisions are also analytic and theoretical decisions (Ochs 1979), representing whether or not the transcriber interprets the rules of the turn-taking system to have allowed or not allowed for transition to a new speaker (Sacks et al. 1974, 704, rule 2; Selting et al. 1998; Wilson & Zimmerman 1986).

18. See Local and Kelly (1986) on the significance of holding of glottal closure during what is transcribed as silence.

19. As we have argued elsewhere, in line with the highly contextual nature of interaction, an understanding of action in sequential context is essential to projection of turn completion and to the likelihood for speaker change to occur (Ford & Thompson 1996).

20. On ‘increments’, see, e.g. Ford et al. (2002), Walker (2004), and papers in Couper-Kuhlen and Ono (2007).

combination of [NP + Adverbial Clause] does not constitute a 'canonical' grammatical turn structure in English. Thus, one warrant for our attention to the stretch of Maureen's talk treated as complete by a recipient (Terry) is the fact that the grammatical make-up of this turn does not fit any *a priori* linguistic category that would qualify it as a coherent syntactic unit-type.

A further warrant for attending to this turn is a more fundamental methodological and theoretical one. As we have noted, unit descriptors such as Noun Phrase and Adverbial Clause were not arrived at through analytic commitment to understanding social interaction; that is, the structural terminology, both for the two parts and for the whole, is not based in action. Drawn as they are from traditional linguistic methods and commitments, these unit terms have implications within linguistics which may be largely irrelevant to our concerns with language in its natural habitat. Looking at Maureen's utterance, our interest, then, is in experimenting with a more socially grounded account for its construction, one not carrying the baggage of the very different empirical methods of linguistics. What happens if we avoid the structural language represented in (5)? What might such a shift do for us as we work to understand how this utterance came to be produced? How does this turn function, for the participants, as a recognizable trajectory of action in its sequential context? What are their visible and hearable orientations to its real-time production?

By attending to the sequential context and the multiple embodied orientations produced before and with it, we find that this unfolding turn not only reintroduces a prior assessment activity and indexes the age of the young artist when he produced the picture (*when he was two-teen?*), but it also does so with a particular kind of intonation and within a particular stream of local and broader action. Locally, the stream of bodily and vocal actions by the speaker forms an interactive matrix for a number of simultaneous and mutually elaborating trajectories of action in the stream of activity shared by all the participants at this moment.

The formulation of the turn and the precision timing (Jefferson 1973) of one recipient's vocal response are far from aberrations. The shaping of the vocal turn is fitted to the particulars of the unfolding dynamics of the broader and more local interactional environment leading up to the turn, and to the action that the turn itself enacts. An analysis that engages with the larger action context in which this turn emerges, as well as the actions and coparticipation accomplished within it, helps us understand the interactional logic at play, and should point us in the direction of a more action-based terminology for how turn trajectories are co-constructed.²¹

21. See Ford and Thompson (1996), Houtkoop and Mazeland (1985), and Lerner (1991, 1996, 2004) on collaborative turns and on prompting, as well as Jefferson (1978) and Sacks (1974) for attention to different forms of projection and to the role of sequential context and turn transition in the projection of a turn's trajectory.

In what follows, we examine the practices through which this stream of activity is organized, including the smooth transition of speakership accomplished at its completion. We offer an account for the unproblematic engagement of the participants in a course of joint action.

Three women (Maureen, Abbie, and Terry) are sitting around a table waiting for a fourth (now in the kitchen) to finish a phone call and return to the table (see Figure 1).

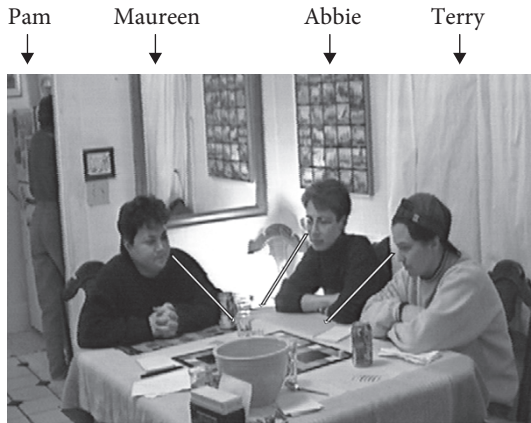


Figure 1. Maureen, Abbie, Terry at the table

As the Extract in (6) begins, a previous sequence and topic appears to be closing. Resumption of a prior turn, or opening a new sequence, are relevant next actions, but so far no one has done either. Through their gaze behavior at the very start of the extract, all three women are enacting non-engagement with one another (Goodwin 1981); they are performing minor self-grooms, and they are not orienting their bodies, faces, or gazes towards each other.

(6) Café de Yin Yang turn in context (GN 2:55)

- | | |
|----|--|
| 1 | (3.5) |
| 2 | Abbie: mmmmm. |
| 3 | (1.5) |
| 4 | Terry: gosh. |
| 5 | (1.0) |
| 6 | Abbie: [⁰ () ⁰] |
| 7⇒ | Maureen: [the <u>café</u>] de yin ya:ng? when he |
| 8 | was tw- <u>te:n</u> ? |
| 9 | Terry: yeah:. |
| 10 | (1.2) |
| 11 | Maureen: [⁰ that is really <u>something</u> . ⁰] |
| 12 | Terry: [an- an- no:te, (.)] the uhm |

- 13 (.)
 14 Maureen: is that a [realf- (.) fe]ather=
 15 Terry: [y'see on the dress?]
 16 Maureen: =on there?
 17 Terry: the yin yang?
 18 Terry: symbols?

What visible and hearable actions lead up to the target turn in line 7? How do the participants shape both the unfolding bodily movements and the talk into possible trajectories of action in lines 1–5?

In lines 1–5, both Abbie and Terry produce vocalizations, but neither produces more than turn-passing tokens; neither turn is formulated in a way that makes any specific kind of responsive action by a co-participant relevant. Furthermore, by gazing at no one during the silences at lines 1, 3 and 5,²² Abbie and Terry are also embodying disengagement. Thus, through their minimal tokens (lines 2 and 4), by not elaborating those vocalizations, and by gazing vaguely forward toward the middle of the table, they construct the moment as a disengaged one.

However, Maureen's embodied actions are different. During the 3.5-second silence at line 1, while Terry and Abbie are enacting disengagement, Maureen begins an embodied and visible trajectory of conduct. She starts to fix her gaze on something on the wall across from her, and she separates her hands from a clasped position on the table, moving her right hand toward her cheek (Figure 2).

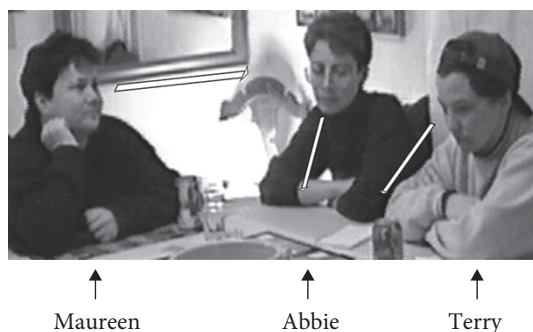


Figure 2. Just before Abbie says, *Mmmm* (line 2), Maureen moves hand to face

22. By “gazing at no one” we refer to the fact that neither participant has fixated her gaze on Maureen. Research reviewed and reported in Gullberg and Kita (2009) makes clear the fact that recipients do not need to gaze directly at the movements of other participants in order to take up information so communicated, particularly socially relevant movements. Many gestures and body movements are in a recipient’s peripheral vision, with uptake of information unaffected by this fact. We thank Charles Goodwin and Richard Sandoval for drawing our attention to this research.

At this point in our analysis, we must look further back in the previous sequential context if we are to understand how this speaker and these recipients are to interpret the trajectory of action Maureen may be forming up; relevance of action in relation to specific sequential contexts is an essential part of projecting what it may take for a turn to develop and for it to reach possible completion, making speaker change relevant and making its absence accountable. The object of Maureen's attention is a drawing that had been briefly mentioned and pointed to five minutes earlier by Pam, who is currently on the phone in another room. At that earlier time, Pam explicitly assessed the drawing as one she was proud of her nephew for creating (Figures 3A–B and Extract (7)):

(7) 5 minutes earlier, Pam had mentioned the drawing

- 1 Pam: did you notice the lovely a::rt?
 2 (0.2)
 3 Pam: my nephew did this when he was te[n].
 4 Maureen: [grea::t,

A



↑ Pam ↑ Maureen ↑ Abby ↑ Terry
 (Stacy behind Maureen)

B



The drawing they are looking at

Figure 3A–B. Did you notice the lovely art?

With this previous context noted, we might propose that Maureen's later turn in line 7 of Extract (6), *the cafe de yin ya:ng? when he was tw- te:n?*, functions as a proffer of topic resumption and a redoing of the stance enacted in the prior sequence as well as the grounds for that positive stance: the age of the artist.

Returning to Extract (6), we can see that Maureen's visible movements make available to the other participants that her attention is directed to some object, with such embodied conduct available to, and consequential for, recipient perception and processing (Gullberg & Kita 2009). We might think of Maureen's embodied action as a pre-beginning to a verbal turn (Schegloff 1996b), but the fact that she is going to speak

is only available in retrospect to us as analysts. Her embodied actions offer a possible trajectory for co-engagement, a trajectory that Abbie and Terry could join in on by moving their own gazes and bodies into responsive alignment with that of Maureen (e.g. by gazing toward the object on the wall).

As Extract (7) shows, earlier in the interaction, Maureen had delivered a clearly positive assessment of the drawing: *grea:t* (line 4). Thus, the object of Maureen's attention display at the beginning of Extract (6) has already been at the center of an interactional sequence, and has already been positively assessed by both Pam and Maureen. Viewed in this larger temporal and sequential context, we can therefore note that Maureen's embodied orientations just before she speaks in Extract (6) propose a possible return to a previous object of joint attention, a potentially already-shared stance toward that object, and a sequence of action organized around doing admiration of the art and the artist. All of this is available to Terry and Maureen, and that previous action context limits the work Maureen needs to do to resume that assessment activity.

During the 1.5-second silence at line 3 in Extract (6), by firmly settling her head into a resting position on her hand, Maureen constructs her new gaze position as an orientation that will not be fleeting (see Figure 4). Maureen projects that she will continue her gaze toward the drawing for some duration.



Figure 4. Maureen brings head to rest on hand while gazing toward drawing

This head-resting position makes Maureen's attention display further available to the others for responsive co-engagement; it can be seen as an upgrading of her offering a trajectory for joint participation in relation to the drawing on the wall. In other words, though Maureen is not gazing at either of the others and is not selecting anyone to speak, by moving her gaze toward a region on the wall which had served as a locus of joint attention five minutes earlier, and by resting her head on her hand while holding that gaze direction, she is initiating a concerted action that is peripherally visible to the others and that may invite them to join in on.

How do the others display (or not) that they are taking in Maureen's visible shift in attention and possible proffer of an object of joint attention? At first, neither Abbie nor

Terry looks at Maureen. However, as Maureen settles her head onto her hand, Terry looks toward her (movement visible in Figures 5 and 6):



Figure 5. Maureen's hand moves toward her face



Figure 6. Terry shifts her gaze toward Maureen

What we have seen so far is the fine-tuned coordination of visible orientations, reactions, and body adjustments by all three participants as Maureen shifts her arm, hand, head, and gaze. Thus, in the local context of a lull in the conversation, and in the broader context of how the picture on the wall has figured in the interaction five minutes earlier, relevant next verbal actions are already limited. That is, the context of the currently-in-progress coordination of bodies and gaze directions provides an unfolding framework in which any verbalization will be interpreted for its relevance.²³ The conduct of Maureen, Abbie and Terry, as described up to this point, is all prior to Maureen's launching of a verbal trajectory.

Just after moving her gaze toward Maureen, Terry says *gosh*, looking downward but keeping her head position at an angle that is more toward Maureen than it was

23. The ensemble of body positions in this moment of interaction involves the consequences of a "postural configuration" (Schegloff 1998).