

Routledge
Taylor & Francis Group

Research on Language and Social Interaction

ISSN: 0835-1813 (Print) 1532-7973 (Online) Journal homepage: www.tandfonline.com/journals/hrls20

Coding Social Interaction: Prospects, Pitfalls, and Procedures

Steven E. Clayman & John Heritage

To cite this article: Steven E. Clayman & John Heritage (2025) Coding Social Interaction: Prospects, Pitfalls, and Procedures, Research on Language and Social Interaction, 58:3, 221-239, DOI: 10.1080/08351813.2025.2528489

To link to this article: https://doi.org/10.1080/08351813.2025.2528489

Published online: 02 Sep 2025.

Submit your article to this journal

Article views: 136

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=hrls20

RESEARCH ON LANGUAGE AND SOCIAL INTERACTION
2025, VOL. 58, NO. 3, 221–239
https://doi.org/10.1080/08351813.2025.2528489

Routledge
Taylor & Francis Group

Check for updates

Coding Social Interaction: Prospects, Pitfalls, and Procedures

Steven E. Clayman and John Heritage

Department of Sociology, UCLA

ABSTRACT

Coding interactional phenomena for formal quantification poses substantial challenges, especially for conversation analysts committed to the emic validity of analytic categories. Recognizing that each step in the process merits more sustained attention than is possible within a single paper, our aim is to sketch in broad strokes the overall process of developing and implementing coding systems. We touch on key issues and, when possible, suggest guidelines and best practices for four phases of the process: (1) developing the coding system, including the importance of context-sensitive case analysis as a basis for coding categories; (2) determining what will count as a codable case, including the problem of incompletely realized and compound cases; (3) testing and refining the coding system; and (4) coding "for real," addressing various methods for maintaining reliability and promoting a common coding culture. Data are in English.

At its inception, conversation analysis was firmly opposed to the coding of interactional phenomena, and the quantification it is normally designed to enable (Heritage & Atkinson, 1984; Hill & Crittenden, 1968). Nonetheless, over the course of its development the field has included studies that draw upon coding and quantitative methods as an adjunct to sequential analysis (Goodwin, 1981; Schegloff, 1968) and these have increased over time.[1] Other early initiatives explored the association between gender and practices related to turn taking (West & Zimmerman, 1983; Zimmerman & West, 1975) and laughter (Jefferson, 2004), representing sporadic attempts to link conversational practices with attributes of speakers. Later studies, mainly in the field of medicine, attempted to establish relationships between actions and/or their design and medically relevant outcomes (Heritage et al., 2007; Heritage & Robinson, 2006; Mangione-Smith et al., 2015; Stivers, 2002; Stivers et al., 2003). At about the same time, studies emerged in the area of media and politics, and connected practices of question design in presidential news conferences to aspects of the sociopolitical context in which they were asked (Clayman & Heritage, 2002a, 2023; Clayman et al., 2006, 2007). Finally, quantitative analyses have been used to examine various aspects of the interaction order itself: relationships between action design and immediate interactional responses (Heritage & Greatbatch, 1986), preference organization (Kendrick & Holler, 2017; Kendrick & Torreira, 2015), the length of turns at talk (Robinson et al. 2022), the normative length of transition-space silence (Jefferson, 1989), and cross-linguistic, cross-cultural variations in interactional conduct (Enfield et al., 2019; Floyd et al., 2018; Rossi et al., 2023; Stivers et al., 2009).

We suggest therefore that the time is ripe to reflect on how the core principles of CA can be, and indeed are, embodied in the design and implementation of coding systems within our field. In the spirit of Schegloff's (1993) cautionary reflections and Stivers' (2015) guarded optimism, we offer

CONTACT Steven E. Clayman clayman@soc.ucla.edu; John Heritage heritage@ucla.edu 375 Portola Plaza, Los Angeles, CA 90095-1551.

[1]An early set of reflections on the potential role of quantification in CA can be found in the following interrelated papers: Jefferson (1993), Drummond & Hopper (1993), and Zimmerman (1993).

© 2025 Taylor & Francis Group, LLC

guidelines and suggestions about ways of working in a quantitative mode that remain grounded in sequential analysis or are maximally sensitive to the issues that arise from it.

We begin, by way of background, with a discussion of the rationale lying behind the conversation analytic stance toward coding and quantification and demonstrate that this rationale supports a cautionary and principled approach rather than one of blanket opposition. We then sketch some of the scientific objectives for which systematic coding of interactional data may be advantageous, and consider some of the ways in which coding instruments should be designed and fitted to these objectives. Finally, we touch on various phases of the coding process from system development to testing and implementation, paying particular attention to key obstacles and best practices for addressing them.

To code or not to code: Coding in CA and its discontents

The opposition to coding and quantification in conversation analysis arose from a variety of interrelated considerations. First there was a keen awareness of the reflexivity and context-sensitivity of language and action that was a primary insight of Garfinkel's ethnomethodology (Garfinkel, 1967). This awareness in turn stimulated an appreciation of the single case as the primary locus of order that emerged with the development of conversation analysis (Sacks, 1992) and would remain a fundamental methodological precept (Schegloff, 1987, 1993) in conjunction with the analysis of case collections. Combined with the comprehensive insistence that analytic claims be grounded "emically" in the orientations of the participants themselves, CA took a stand against all forms of exogenous, "top-down" coding that forced the analysis of interaction into theoretically preconceived categorical frameworks. This anti-coding position in CA, and the analytic approach pursued in its place, was the antithesis of exogenously generated coding systems then in vogue (e.g., Bales, 1950; Roter & Larson, 2002).[2] In short, the Garfinkelian arguments concerning the irremediably context-sensitive nature of language use and action were, in their fundamentals, irrefutable (Heritage, 1984, pp. 142–159), and Schegloff's (1993) supplementary arguments against coding and quantification, particularly the top-down variety, further barred the door to such work. Why, then, resurrect these issues at all?

Notwithstanding the powerful arguments put forward in Schegloff's critique, the paper was neither written nor intended as a blanket proscription against all forms of quantification. Indeed, he acknowledged the existence of so-called informal quantification "represented in the common use in some conversation analytic writing of terms such as *massively, overwhelmingly, regularly, ordinarily*, and . . . *commonly*" (Schegloff, 1993, p. 99). With these terms, authors clearly convey estimates of frequency, though without submitting them to the rigors of actual counting and the coding that would enable it. Moreover, as Stivers and Rossi (2025/this issue) note, in this and a subsequent paper on comparative work in CA, Schegloff (2009) did not shut the door on quantitative studies, but rather aimed at establishing conditions for the defensible use of quantitative techniques, for which indeed he expressed support as a resource for increasing confidence in analytic claims.

In a fundamental way, the "coding question" is not so much a matter of resurrection as it is a matter of inexorability. Consider that the naming of any individual object, for example as a "table," or an action in its context, for example as a "question," is necessarily an act of classification that adumbrates generalization (Mervis & Rosch, 1981; Rosch, 1987) and embodies a primitive form of coding (Robinson, 2024). Using the *same* name to describe *another* object or individual action in its context is a more profound act of classification that asserts some measure of comparability across the first and second cases (Mondada, 2023). At its heart, then, naming is tantamount to categorization, and

[2]The Bales coding system was designed to study interaction in small groups. Its categories were based on Parsons' theory (Parsons & Bales, 1955) that in effective groups some members were task oriented, while others were more focused on managing relationships within the group. The method is still quite widely used today. The Roter Interaction Analysis System (RIAS) adapted Bales's categories to the characteristics of medical encounters.

using a name to describe one or more actions or events does not differ in its fundamentals from an act of coding.

Just as the assimilation of single cases into categories is unavoidable in any enterprise involving the use of natural language, the same can be said of quantification when used to make claims beyond the single case. As Schegloff (1993) acknowledged in his remarks about informal quantification in CA, and as is evident from the use of cross-case comparisons and small collections in Sacks' (1992) earliest lectures and in subsequent foundational papers by Sacks and his colleagues (Schegloff, 1968; Schegloff & Sacks, 1973; Sacks, Schegloff, & Jefferson, 1974), the combination of context-sensitive case analysis with the analysis of patterns and their robustness across collections of cases has been a hallmark of the CA approach from the outset (Clayman et al., 2022; Clift & Raymond, 2018; Schegloff, 1997). It bears emphasis however that such quantification, however small-scale and informal it may be, is necessarily built upon a foundation supplied by single instances of data categorization.

While even the most qualitative of CA studies necessarily deploy forms of classification that amount to coding in its most primitive and basic form, these studies remain entirely distinctive from traditional coding-based studies in the social sciences. The distinctiveness is this: The classifications used in CA studies are grounded in the orientations of the participants themselves that are emergent in real time, and demonstrable in context by reference to both sequential and concurrent elements of conduct (Robinson et al., 2024). Conversation analytic categories are thus rooted in context-sensitive case analysis and are on that basis endogenously validated. There is a marked contrast between these categories and their "emic" manner of validation, as opposed to exogenously-generated categories ungrounded in the participants' own orientations, and their validation in terms of other exogenous constructs.[3]

In recent years, as previously noted, more explicit forms of coding and quantification in CA have become more frequent. Three main interests have contributed to this development. First were efforts to address institutionally specific and applied questions in medicine, mass media, and politics, involving linkages between conversational practices and actions and exogenous variables of various kinds (Stivers & Rossi, 2025/this issue). Subsequently, research on cross-language comparisons mandated coding because they could not be convincingly supported without quantitative evidence. Most recently, quantitative methods have been deployed to substantiate and elaborate upon classic CA topics such as turn taking, repair, and preference (e.g., Hayano, 2025/this issue; Healey et al., 2025/this issue; König & Pfeiffer, 2025/this issue). These interests notwithstanding, coding-based quantitative studies can never displace the primary investigation of interactional practices using methods established during fifty years of conversation analytic studies (Robinson et al., 2024). For the emic validity of these quantitative studies rests on the primary findings concerning language use and practical reasoning that are established in the focused and detailed analysis of singular instances of interactional conduct, and they lose their moorings when they become detached from them.

Varying aims of coding projects

Unlike traditional conversation analytic projects, which may commence from unmotivated looking (Sacks 1984) or serendipitous observation, coding projects necessarily involve predetermined objectives, and therefore the design of coding instruments and systems must begin with a consideration of the goals to be pursued. Here we review some of the various objectives motivating extant research and comment on their distinctiveness and import for coding system development.

Documenting the prevalence of conduct. Quantitative projects may be geared to a relatively straightforward objective of determining the sheer prevalence of some interactional phenomenon. These projects may be motivated by hypotheses derived from previous research or theoretical frameworks.

[3]As documented in Clayman and Heritage (2021), the "emic validation" provided by conversation analytic methods is unique and entirely distinct from the panoply of validation principles and techniques featured in discussions of coding and measurement in the social sciences (see Krippendorff, 2018: 155–168 for a representative overview).

For example, following Sacks et al. (1974), Robinson et al. (2022) documented the frequency of single-unit turns as a proportion of all turns at talk. By showing that approximately two thirds of all turns are comprised of a single turn constructional unit, they validated the Sacks et al. hypothesis that the turn taking system for conversation is biased in favor of single-unit turns.[4] Similarly, Stivers et al. (2009) were able to show that the modal length of silence between questioning and answering turns was zero seconds, thus partially validating another hypothesis derived from Sacks et al. (1974) that the turn taking system is organized to minimize gap and overlap in turn transition. Before these two studies, Jefferson (1989) established a standard maximum silence between turns of 1.0 seconds, a finding that was later indirectly confirmed by Stivers et al. These studies' conclusions would not have been convincing without explicit coding of turns and turn constructional units together with measurements of timing.

One step beyond the sheer prevalence of single phenomena are studies of the comparative prevalence of sequentially relevant alternatives. For example, the literature on preference organization has demonstrated qualitatively that different practices are engaged in the production of preferred and dispreferred second actions (Davidson, 1984; Pomerantz, 1984; Pomerantz & Heritage, 2013; Pillet-Shore, 2017), and a cluster of studies has recently marshalled quantitative evidence in support of these claims. For example, dispreferred responses have been found to be systematically delayed relative to preferred responses (Stivers et al., 2009), and turn-initial practices (e.g. *well*-prefacing, audible inbreaths, and particles associated with repair, e.g., *uh*) are, as predicted in the foundational literature, systematically associated with dispreferred responses (Kendrick & Torriera, 2015). Kendrick and Holler (2017) also showed that gaze aversion by speaker is associated with dispreferred responses.

Connecting Practices with Outcomes. It has long been suggested that the openings of conversations can be predictive of what will transpire within the conversation as well as its non-conversational outcomes (Pillet-Shore, 2012; Schegloff, 1986). This association between openings and their sequelae has been particularly relevant in studies of medicine, where interactions are frequently goal-directed from the outset and geared toward decision-making. For example, particular forms of opening question design have been shown to index the nature of the visit (Robinson, 2006), and to invite expanded or unexpanded responses from patients (Heritage & Robinson, 2006) and in related contexts to influence the content of patient responses (Robinson et al., 2016; see also Heritage et al., 2007). More generally, question design forms can also influence overall outcomes of medical visits, such as the fulfillment of patient medical objectives and patient satisfaction (Boyd, 1998; Robinson & Heritage, 2006; Robinson et al., 2016). In other instances from medicine, the expression of patients' candidate diagnoses correlates with physician beliefs that patients expect antibiotics (Stivers et al., 2003) and increased prescription of antibiotics (Mangione-Smith et al., 2015). Other associations between conduct and outcomes have been documented in contexts ranging from calls for emergency service (Kent & Antaki, 2020) to political speeches and audience reactions (Heritage & Greatbatch, 1986).

Connecting Practices with Social Contexts. A central tenet of CA concerns the phenomena of recipient design and, more broadly, context sensitivity. As Goffman (1955) was the first to propose, every turn at talk embodies a stance toward the context in which it occurs including the relevant identities of speakers and recipients and the social occasions in which they are lodged. Against this backdrop, interactional practices have been used as expressive indices of relevant contextual elements. For example, the design of greetings is commonly sensitive to social relations between the parties (Pillet-Shore, 2012; Schegloff, 1979), while as Jefferson (1980) showed, responses to *how are you* inquiries can hint at "hidden" troubles. Similarly, physicians' opening questions to patients convey their assumptions about the context of the medical visit and patients' reasons for being there (Robinson, 2006). Moving beyond the purely expressive level, action design has been used to take the measure of phenomena as diverse as the observance of professional norms such as journalistic neutralism (Heritage & Roth, 1995) and adversarialness (Clayman & Heritage, 2002b; Clayman et al.,

[4]This contrasts, for example, with the turn taking system for news interviews in which multi-unit responses by interviewees are frequent and normative (Clayman & Heritage, 2002b).

2006), the existence of implicit racial biases (e.g., in medical questioning; Stivers & Majid, 2007), and recipients' growing competence in an instructed skill (e.g., driving instructions; see Deppermann, 2018). More broadly still, the quantitative study of interactional practices can provide an important window into cultures and cultural differences. This is most apparent in cross-language studies, for example Floyd et al.'s (2018) study of the expression of gratitude in eight languages, Rossi et al.'s study (2023) of recruitment, and Enfield et al.'s (2019) study of responses to polar questions (see also Brown et al., 2021).

While the studies described above treat language practices as indicators of sociocultural phenomena, other research pursues associations between changes in interactional practices and larger socioeconomic and sociocultural developments. In analyses of U.S. presidential news conferences, questioning practices have been found to vary in relation to local conditions related to question content (domestic versus foreign affairs topics), the president's term in office (first versus second terms), and more distal conditions such as movements in the business cycle (Clayman et al., 2007) and periods of sociopolitical upheaval (Clayman et al., 2010; see also Comstock, 2023; Deppermann & Pekarek Doehler, 2021).

The coding-based studies reviewed in this section include primary studies of language practices and their prevalence, and studies of the association between practices and outcomes as well as connections to both local and more distal sociocultural contexts. These studies could not have been performed without coding and the quantification it enables. It is important to bear in mind that coding systems are tools for addressing particular empirical questions and theoretical puzzles. They are always devised for a motivating purpose and, save for their use in replications and adaptations to other linguistic and cultural contexts, they may have limited utility and may indeed lapse from use once they have served their purpose.

Project aims and coding approaches

As we have suggested, research objectives necessarily guide coding system development, and the diversity of objectives makes it impossible within a single paper to characterize all of these contingencies. In what follows we touch on a few key considerations that that are broadly relevant across projects of various kinds, and that should inform coding system development in the context of particular scholarly aims.

A primary consideration is whether the coding project targets language practices per se as opposed to investigating them as a window into something "larger": e.g., interpersonal relations, social norms, or other aspects of the sociocultural context. For the former type of project, exemplified by the studies of TCUs and between-turn silences noted earlier, the coding system can be "narrowly" focused on the particular practices of interest. When these topics have been extensively researched conversation analytically, coding systems can build upon this foundation and can be accomplished with a high level of reliability.[5]

By contrast, studies using language practices as a window into "larger" sociocultural phenomena may require coding systems that target a wider range of practices treated as indicative of the (possibly varying) sociocultural phenomenon of primary interest (Clayman & Heritage, 2021). For example, in a project focusing on changing levels of deference and adversarialness toward U.S. presidents, Clayman and colleagues (Clayman et al., 2006; Clayman & Heritage, 2002a) examined a dozen language-based indicators for five dimensions of adversarial questioning derived from earlier work on news interview interaction (Clayman & Heritage, 2002b). Because of the sociocultural interests motivating the project, the coding system had, by necessity, to encompass both formal features of question design (e.g., question formats embodying "pressure") as well as aspects of question content (e.g., topics and themes that are critical or oppositional toward the president).

[5]For example, a recent study by Robinson et al. (2023) reports levels of agreement on a range of turn-taking issues ranging from .95 to 1.0 as measured by Krippendorf's alpha (Krippendorf, 2013).

In a related vein are studies of the expression of medical authority which, in the context of diagnosis, have documented the increasing role of diagnostic explanations (Peräkylä, 1998, 2002, 2006) and the growth of expressions of uncertainty (Heritage & McArthur, 2019) suggestive of an erosion of medical authority. Similarly, a study of treatment recommendations (Stivers et al., 2018; see also Heritage, 2021) showed a marked tendency for authoritatively-expressed recommendations to be clustered around antibiotics and related prescription medications. Coding in these studies required ways of classifying diagnostic statements and treatment recommendations in a manner that would capture such issues as certainty and uncertainty, unilateral versus collaborative orientations, and the nature and type of physician explanations.

All of these projects, whether focused on singular language practices or a range of practices as a 'sociocultural window', were largely built upon extensive conversation analytic research into political and medical interactions long before more quantitative efforts were attempted. Without this basis, they would have lacked the emic validity that arises from elementary grounding in the behaviors and enacted understandings of the participants themselves.

A second consideration concerns the granularity of the project (see also Deppermann & Gubina, 2025/this issue): does the study target a singular type of practice (e.g., prefacing a turn with *well*), or an "action" and the variety of practices through which it may be achieved.[6] As an example of the latter, consider the various forms that a question can take (Bolden et al., 2023a; Enfield et al., 2010): interrogatives, declaratives constituted as questions epistemically or with rising intonation, tag questions, alternative choice questions, etc. At a still greater level of complexity are studies that concern integrated systems of practice such as turn taking (Sacks et al., 1974) or requests (Rossi, 2025). Finally, at perhaps the highest level of abstraction, analysis can combine various practices and actions to form conduct-based measures or scales of social phenomena, such as the multidimensional measures of aggressive journalistic questioning developed by Clayman and Heritage (Clayman et al., 2006).

Particular studies may engage one or more of these issues with different relative emphases or foci. Ordinarily a less-granular type of project addresses a broader array of practices implicating a wider range of actions and their contexts; but even highly granular projects targeting singular practices may need to attend to environing conduct as well in order to pinpoint their functionality, and indeed many CA studies describe granular practices of conduct by locating them in relation to particular sequential contexts. This approach has proven particularly successful when aspects of turn design and their consequences are considered in relation to particular actions and sequential positions (e.g., Clayman & Raymond, 2021; Curl 2006; Heritage, 2018; Heritage & Robinson, 2006; Kendrick and Drew 2014; Pomerantz & Rintel, 2004; Rossi 2025; Stivers, 2002; Stivers et al., 2003).

Coding social action: The relevance of position and composition

Earlier we noted that CA was from its inception opposed to the use of extant coding techniques based on externally-imposed action categories that largely ignored the contexts in which actions occurred. This stance was taken in the interest of building analyses of conduct around action categories that the participants are demonstrably oriented to, together with their context-shaped implementation. In this section we provide an initial consideration of how a coding enterprise may be developed within the "emic" constraints of the CA perspective.

CA's focus on the study of situated social action is encapsulated in its uniquely endogenous originating question: *Why that now?* From the vantage point of participants within interaction, this question is the overarching consideration that drives the ascription of intentionality to utterances and other situated behaviors, which is in turn the level of social meaning that drives the forward development of interaction. For participants, the primary resources for this action-ascriptive process include both: (1) the *composition* of the interactional contribution, as well as (2) its *position* within the developing interaction and its context (Pillet-Shore, 2024). Correspondingly, coding that incorporates

[6]This distinction parallels that between semasiological and onomasiological approaches within linguistics.

a conversation analytic sensibility must embrace analysis in similar terms, and that means attending not only to the internal design of conduct but also its social embeddedness.

Position can in principle embrace a variety of contextual dimensions: e.g., after a question, before a request, in the lead-up to a closing, or more narrowly, in overlap with another speaker, while gazing away, after a recipient's concurrent gesture or facial expression, etc. Outside of ordinary conversation, e.g., a news interview, clinical consultation, survey interview, etc., some of these sequential positions may be re-specified in relation to distinctive turn-taking systems, activity phases, or overarching interactional goals (Robinson et al., 2016; Schaefer et al., 2013). In general, many of the more successful examples of coding within CA have looked at *position* primarily in terms of location within a sequence (i.e., first, second, or third position; Heritage, 2018 on *oh* and *well*), or location within a turn or turn constructional unit (e.g., Clayman & Raymond, 2021 on *you know*), or varying responses in the same sequential slot (Enfield et al., 2019; Stivers, 2022, Stivers et al., 2009).

For example, prefacing a turn with *well* can have a different import depending on whether the turn is an initiating "first position" action, a responsive "second action," or in third position (Heritage, 2015, 2018; Kim, 2013). In (1) below, *well* appears in all three positions (arrowed) and does discriminably different work in relation to the action being implemented at each point (Heritage, 2015). At line 7 Amy's *well*-prefaced turn shifts the topic from Bea's connection with a man interested in "spiritual healing" to a well-known spiritual healer ("Kathryn Kuhlman"). At 9, Bea's *well*-prefaced response resists the terms of Amy's negatively polarized query, treating it as obvious that she would know the person in question. Finally, Amy's *well*-prefaced third turn forwards the sequence, retrojectively providing that the preceding turns were prefatory her announcement in line 10 (Kim, 2013).

```
(1) [SBL 2:2:4 8:55]
 1 Bea:     we've been working together for some time, and
 2          we have uh certain periods during the day thet
 3          uh kind of tune in to the [↑vibrations and uh
 4 Amy:                               [Yeah.
 5 Bea:     he is int'rested in anything where there is uh
 6          spiritual healing.
 7 Amy: 1-> yea- Well, ↑tell me. (.) Do you- does the name
 8          (.) Kathryn Kuhlman mean anything to you?=
 9 Bea: 2-> =Well I should say so.
10 Amy: 3-> Wul do you know that she's in Los Angeles? n'
11          giving uh: ((continues))
```

In this example, then, the practice of *well*-prefacing a turn is deployed in the service of identifiably distinct actions, all of which are facets of its systematic deployment in the three sequential contexts illustrated here.

When looking at *composition*, it is common to look at different ways of building turns at talk within some particular sequence with a view to explicating their differential import. For example, affirmative responses to polar questions can take the form of interjections such as *yes*, or repetitions of some aspect of the preceding question, normally the verb. The significance of this difference turns on the fact that whereas interjection responses tend to acquiesce to the terms of the question, repetitional responses tend to be more "assertive" or "confirmational" thus embodying the respondent's epistemic or other claims vis a vis the questioner (Bolden et al., 2023b; Enfield et al., 2019; Heritage & Raymond, 2012; Raymond, 2003; Raymond et al., 2021; Stivers, 2022; cf. König & Pfeiffer; 2025/this issue).

To illustrate this contrast, in (2) a mother and her daughter (Lesley) are talking about the daughter of a third party (Ann). The mother is explaining that Ann's daughter wants to follow a new fashion for multiple earrings in a single ear, provoking a disagreement with her mother. Commenting on the daughter's stance, Lesley observes that the fashion is 'very cheap' (line 5), producing the observation in the form of a [statement] + [tag]. Mum's response to that takes the form of a confirming repeat of Lesley's turn followed by an agreement *('Yes')* (line 6).

```
(2) (Field 1.1:8)
 1  Mum:     .hh We:ll. (0.2) Uh you know (.) there's
 2           a cra:ze with the girls now to have (.)
 3           a secon:d. (1.1) ring ih- a secon:d
 4           uh earring in on[e ear.
 5  Les:                     [Oh: it's very chea:p isn'it.
 6  Mum: ->  It's very cheap yes'n this is u- this is
 7           what Ann said. An' Ann said (0.3) she-
 8           she'll haf (.) tuh have another. (0.5)
 9           .hh (.) hole in'er ear...((continues story))
```

Here the ordering of turn components in which the repetition precedes the interjection yields a turn that is centered on the act of confirmation that repetitional responses achieve. Mum subsequently underscores her confirmation by reference to the child's mother's opinion ('this is what Ann said'). Here Mum privileges confirmation over affirmation or agreement, in the process leveraging her access to Ann (a party to the dispute) as an element of epistemic primacy in the opinion.

And in (3), by contrast, Mum responds to Lesley's initial query about garlic tablets with an acquiescent interjection followed by a brief expansion (line 2). However when Lesley pursues the point by asking whether she's taking them, Mum responds with a more agentive repetitive response confirming that she "started taking them today" (line 4).

```
(3) [Field 1:1:89-94]
 1  Les:     Uh didyuh get yer garlic tablets.
 2  Mum:  -> Yes I've got them,
 3  Les:     Have yuh t- started tak[ing th'm
 4  Mum:  ->                        [I started taking th'm t'da:y
 5  Les:     Oh well do:n[e
 6  Mum:                 [Garlic'n parseley.
```

While Mum's first interjective response simply affirms the question's embedded proposition, her second response – a repetition – adopts a more confirmatory stance. Thus these distinctive responses enact distinctive stances towards the polar questions to which they are addressed.

Related comparisons can involve more macro contexts as well. For example, Rossi (2012, 2025) demonstrates that the choice between imperatively- and interrogatively-formed requests is sensitive to whether the request involves the continuation of an ongoing bilateral project, or whether it obliges the recipient to depart from a current activity to fulfill the request (see also Stivers & Sidnell, 2016 for parallel observations). Or again, in pediatric consultations, problem presentations expressing a candidate diagnosis are associated with different physician perceptions and prescribing outcomes compared with problem presentations that simply describe the child's symptoms (Mangione-Smith et al., 2015; Stivers, 2002; Stivers et al., 2003).

In sum, the analysis of sequentially organized practices of talk has been most successful when attending to the twin relevancies of position and composition in the constitution of social action. Correspondingly, studies involving coding and quantification are most effective when they are sensitive to these same dimensions, including especially position within a sequence or activity context. What this means in practice will, of course, necessarily hinge on the aims of a given project. For studies seeking to determine the action-import of previously unexamined practices, it may be necessary to explicitly code for both positional and compositional features. For these discovery-oriented projects, the comparative logic exemplified above – i.e., holding the practice constant while varying its sequential position, or holding sequential position constant while varying the practices employed – has proven to be highly effective. On the other hand, for projects aiming to extend or apply previous findings, it may suffice merely to incorporate

a sensitivity to sequential and activity contexts in the process of rendering coding decisions. In any event, attention to sequential and activity contexts, and not merely isolated units of behavior, has been a hallmark of conversation analytically informed coding projects and is essential to their success and generativity.

Case inclusion/exclusion considerations

-Before coding can begin, it is necessary to determine which cases within a given dataset are relevant or appropriate to code (see also Stivers & Rossi, 2025/this issue). This preliminary determination is often taken for granted, but it is a significant pre-coding "coding" decision and a prerequisite for the application of any coding system to a pool of cases. The aims of the project may matter here also; sensible resolutions may be contingent on what the research is designed to achieve.

An issue that emerges in almost every coding study is how to handle *cases that are not fully realized*, such as those that are elided or abandoned prematurely by the speaker, or alternatively interdicted by the recipient. Decisions concerning these cases tend to be highly practical and, again, sensitive to the project's objectives. They may often hinge on whether "enough" of the target phenomenon has been realized for relevant coding categories to be defensibly applied.

Similarly, case-inclusion issues can arise in the context of *compound cases*. For example, in studies of question design, a questioning turn at talk may contain two or more identifiable questions. The researcher may then need to consider whether these amount to a "real" multi-part question imposing multiple response demands on the recipient, or alternatively whether they involve the replacement and re-doing of a first question by a second (Clayman & Heritage, 2002a).

Case-inclusion puzzles can also arise from *the sometimes-fuzzy boundaries of the interactional occasion*. For example, doctor-patient encounters frequently contain episodes involving "small talk" (about vacations, etc.) which can emerge prior to or in the midst of the main business at hand. Similarly, presidential news conferences can also include preliminaries that may encompass "questions" but not those understood to be "journalistic" or "on the record." Decisions about the inclusion of this kind of conduct must clearly turn on the research objectives in play.

Ideally, inclusion criteria will be clearly specified in the project codebook (discussed further below). For an illustration of what this might look like in practice, here are the case-inclusion criteria from the codebook for our project on question design in presidential news conferences. Our list of criteria provide instructions for dealing with partially realized cases, side-sequences involving repair, cases extraneous to the institutional occasion, and other problematic situations.

What Counts as a Codable Turn:

- Code all sequence-initiating turns by reporters seeking to elicit substantive information/opinion from the president.
- Exclude questions seeking to repair minor problems of hearing or understanding what the president just said (eg., "Huh?," "Did you say the House or the Senate?," etc.).
- Exclude turns responding to a Q from the president, or informing the president after "don't know" response.
- Exclude turns that seem extraneous to the main business of the conference (eg., requests to open the window, attempts to get the president's attention or to quiet others, wholly frivolous joking, etc.).
- Exclude turns that are interrupted or otherwise aborted so as to be rendered inconsequential. However, if enough of the question is produced for it to be intelligible and for the president to respond to it, then code it.
- Exclude any clusters of Qs that get no response.

- Exclude prefaces interrupted before the Q is delivered.
- If the president has trouble hearing the question and asks for it to be repeated, code only the second reissued question. However, if only a single word or phrase gets repeated, code the original full-form question.

While specified criteria are extremely useful resources for identifying codable turns, they cannot resolve all of the problems that can arise in the coding process. Perhaps the thorniest inclusion issue arises from the potential for any target phenomenon to be vulnerable to *the reflexive relations between the phenomenon and its subsequent treatment*. In his first lecture, Harvey Sacks invites us to consider the following sequence:
The call-taker's "Can I help you" is apparently designed as a conventional way of soliciting an

(4) [Sacks 1992: 10]
1 SPC: Can I help you?
2 Clr: I don't know hheh
3 ((pause)) I hope you can.
4 SPC: Uh huh Tell me about your troubles.

initial problem report from the caller (cf. Robinson, 2006), and this is confirmed by the call-taker's pursuit at line 4. However, the caller's "I don't know" response treats it as merely a query about the call-taker's ability to help, in effect exploiting the surface form of the question as a means of sidestepping the solicitation. The coding puzzle that cases like this raise is fundamental: are we to code the call-taker's turn at line 1 as the problem solicitation it was ostensibly "intended" to be? Or alternatively do we code it in terms of the "next-turn proof procedure", i.e., the treatment it received as a question about the call-taker's ability to help (Schegloff, 1996)? How are these alternatives to be reconciled? This matters for case inclusion because, depending on the researcher's judgment, the case may or may not be included as, for example, an "offer of help." Although we've introduced this problem in our discussion of case-inclusion, it is in fact a far more general problem in the application of coding systems to actual data (see Deppermann & Gubina, 2025/this issue).

There is no one-size-fits-all solution to this dilemma, but again varying approaches that would be reasonable in the context of particular projects. Some projects may be fundamentally "about" the stance taken by speakers through language practices; others may be essentially about concurrent treatment and uptake by others. Distinctive goals of this kind may mandate different approaches to the resolution of this dilemma.

Categories and their application

While other contributions to this special issue devote primary attention to the process of developing coding categories in relation to particular topics, in this and succeeding sections we provide an overview of various phases of coding development and implementation and suggest best practices toward this end. We take up, in turn, the creation and use of a codebook, testing and refining the coding system, and maintaining coding reliability while coding "for real."

Making a codebook

Coding studies are unlikely to be successful unless the system and rules for its use are carefully elaborated and preserved in a codebook. Codebooks are essential for keeping track of the coding categories and associated rules for their application. As part of the collective memory of the coding

system, they can incorporate reasons for creating codes, and they may be adjusted as necessary over time. Codebooks also function as foundational documents to which coders can return at moments of difficulty, confusion, or disagreement.

To illustrate what this can look like in practice, consider this passage from the codebook we developed for presidential news conference questioning. The passage, slightly simplified for this illustration, concerns the coding for "question complexity", a variable with 5 categorical values: single-TCU questions (code 1),prefaced questions (code 2), compound or multi-part questions (code 3), and so on. The codebook lists an exhaustive set of categories to be applied to each questioning turn at talk, and a set of instructions guiding the application of each code. It also contains additional guidelines for handling unusual or ambiguous cases: for instance, how to code statements that are included in the turn but not in the usual prefatory position; how to code statements that are clausally dependent on the question rather than separate sentential units; and so on.

Question Complexity

1 One questioning TCU
2 Statement(s) + Q
3 2 or more Qs
4 Statement(s) + 2 or more Qs
5 Declarative assertion(s) only

CODE 1 includes all single TCUs that function as Qs, whether they are formatted as standard interrogatives, single statement + tag Qs, B-event statements, or 3rd-party attributed statements. Code 1 also includes single statements followed by final questions that are grammatically discrete but substantively content-less and hence parasitic on the prior statement (eg., statement + "is that correct/right/true").

CODE 2 includes all single TCU Qs that contain preliminary statements that express propositions bearing on the topic at hand (e.g., "Last week oil prices rose sharply." + Q). This code also includes Q + statement (which may be thought of as "post-face" statements). Exclude statements that merely name a topic ("Concerning oil prices." + Q) or project an action (e.g., "Let me as two questions").

NOTE: Preliminary subordinate clauses (e.g., "Regarding interest rates," + Q) generally should not be counted as preliminary statements. However, if such a clause is elaborate and substantive, flag it for consideration in our weekly staff meetings.

NOTE: For questions in the form [Q + statements + reissued Q], where the final Q is another version of the initial Q, the initial Q should be understood as part of the preface (and not as a separate Q or as a Q cascade). Hence, CHOOSE CODE 2 FOR THIS TURN DESIGN.

This question complexity example concerns coding categories that are highly structural and relate to the "syntax" of turn design. Other coding categories and category systems may be addressed to more sequential-pragmatic aspects of turn construction, as in the example below which concerns a "pre-sequence" type of question mainly found in early presidential press conferences. Because this coding category, which we termed a *preparatory condition question*, is less structural and more content-based, the coding instructions include not only rules for the category's application but also a number of prototypical examples illustrating the phenomenon to be captured by the category, and well as examples illustrating boundary cases that are superficially similar but should be excluded from the category.

Column 13a: Preparatory Condition Qs
0 not a preparatory condition Q
1 preparatory condition Q

A preparatory condition Q asks if the president is prepared (either by his own actions or by his staff) to provide an answer. These address conditions that make the president *able to* respond, but rather than such conditions being grammatically encoded in modal verb frames (*can you/could you*), here they are the explicit focus of the question.

Examples: *Have you put your mind to the issue of X?*
Do you know whether X?
Are you aware of the fact that X?

Have you been briefed on the matter of X
Has the matter of X come across your desk?

NOTE ON BOUNDARY CASES: Exclude pre-sequence Qs addressing conditions other than the president's own preparedness to answer. For instance: "Has the bill taken final form yet"; "Have you met with Prime Minister X since his arrival in the U.S. yesterday." These may be preliminary to soliciting the president's viewpoint, but since they reflect exogenous events largely beyond the control of the president or the administration, they do not count as preparedness Qs.

Codebooks should be developed in tandem with the coding system itself. They should encompass and track the full range of coding categories, guidelines for their application (which may include rules as well as illustrative examples), and pre-coding case-inclusion guidelines. The codebook should be continuously updated in the face of encounters with anomalies, difficult cases, and other bumps in the road. These encounters will prompt revisions to the codebook and it is important to keep a record of these revisions via successive iterations of the codebook. What is central, here, is to maintain an explicit and continually updated record of the evolving grounds for coding decisions. This iterative process is one of the means by which the coding system is developed in a direction that aligns with the orientations of the participants and thus remains "emically" grounded.

Testing and refining the coding system

Coding systems, no matter how solidly grounded and carefully developed, are rarely ready for implementation straightaway. Normally they require extensive testing that will almost always result in the discovery of unanticipated problems in using the system, which may then require corrective efforts. For example, upon testing it may become apparent that specific codes are insufficiently definite or are difficult to apply to concrete cases, or too often generate coder unreliability or disagreement. These problems may mandate the addition or subtraction of coding categories, or elaboration of the category-application guidelines. This process of refinement may take considerable time and may require many iterations.

Some limitations can be rectified with the addition of coding categories or category values. For example, in the presidential news conference project, we initially attempted to code questions' policy content as a simple binary variable: foreign versus domestic policy. As it turned out, this was a recurrent source of disagreement in the face of questions about military budgeting, which could be construed as either "foreign" or "domestic" depending on whether its import for

international relations or domestic funding priorities is emphasized. After discussions with the coding team (discussed further below), we resolved this problem by adding a third category for military and national security issues. This simple addition circumvented inter-coder disagreements and coding on this issue proceeded straightforwardly.

Testing may also reveal ambiguities in the framework which are resolvable by collapsing existing categories. For instance, in a project on *voir dire* questioning of prospective jurors (Clayman & Fox, in press), there was an interest in capturing the extent to which questions about juror bias formulated "bias" indirectly. The authors initially distinguished two types of indirect reference: (1) *litotes* reference which refers to bias by problematizing its opposite (e.g., *difficulty being fair*), and (2) *allusive* reference which avoids the language of fairness/bias altogether (e.g., *difficulty serving as a juror*). The distinction seemed clear from prototypical cases, but testing revealed so many "fuzzy" cases that reliable coding proved persistently elusive despite repeated efforts to elaborate the coding guidelines. In the end the time investment and poor results motivated the abandonment of the distinction by collapsing *'litotes'* and 'allusive' into a single and reliably codable category for 'indirect references to bias.'

Finally, a cautionary note about reliability. It is possible to acquire a subjective sense that the coding system has achieved a superior threshold of reliability so that it is now ready for implementation. This sense can, however, be illusory, sustained through the tacit operations of *ad hoc* reasoning and tendencies toward confirmation bias. Accordingly, it is advisable to submit trial coding results, whether created by individual coders or by teams, to a statistical test of reliability, such as Cohen's *Kappa* or Krippendorf's *Alpha* (Lombard et al., 2002) before coding commences "for real." Unreliable codes will necessarily require further team discussion, evaluation, and correction.

Coding "for real"

Once the coding system is developed and successfully tested, further issues can emerge in the context of its implementation.

Consider, first, the different conditions that obtain for a solo researcher as opposed to coding with a collaborating colleague, or alternatively delegating the process to a team of research assistants. In all three conditions, maintaining a stable common culture of coding is of paramount importance to prevent coder "drift" and more generally promote reliability.

In our experience, coding with a colleague or more generally in pairs has been more reliable than solo coding, largely because each coder acts as a check on the other. This can expose differences in interpretation (of cases or coding categories or both) resolvable through discussion and thereby enhancing inter-coder reliability. For this reason, if coding is delegated to a team of research assistants, it is desirable, funding permitting, to have the coders work in pairs to build in a continuous check on coding decisions. Coding in pairs is not, of course, a magic bullet; coding partners can 'drift' together in a *folie à deux* process if wayward judgments become persuasive and mutually-reinforcing among partners. In a context where teams of coders are working in this way, coder reliability can be measured at the level of the pair, having all pairs code a common data subset to expose possible drift of this kind.

Whether coding is performed in pairs, or by team members working independently, the maintenance of a common coding culture may require additional steps. In the presidential news conference project, we were assisted by a sizable team of fourteen coders. Our research team members were, as much as possible, trained in concert with one another. Training involved a series of test coding runs involving all coders working on a common data subset. We sustained the common culture across the coding crew throughout the duration of the coding process through weekly meetings of the entire team to discuss problem cases in a public forum. These

meetings have a dual significance: they allow for the shared resolution of problematic cases, while also renewing and maintaining the common culture of coding. Each week coders would be reminded of the complex guidelines and their import for specific cases. In addition, we periodically rotated coders into different pairs, inhibiting any tendencies toward *folie à deux* by requiring that each coder be exposed to the independent judgments of different partners. All of these measures were designed to stabilize and universalize the culture of the coding project.

Perhaps the thorniest aspect of coding interaction concerns the importance of contextual knowledge. This is a general and unremitting problem for the coding of any kind of discourse or interaction. One aspect of this context concerns the knowledge that interactional participants have of one another and of the world at large. Ideally this knowledge should be accessible to the coder as a condition of realistic coding decisions. For example, in understanding the import of arrangements to get together, it may be necessary to know who can or cannot drive, who has or does not have a car, as well as the locations of the participants in relation to the rendezvous point (Heritage & Clayman, 2024). Access to this kind of taken-for-granted common knowledge is essential for appropriate coding decisions and lack of this access can seriously hamper the process.

This general problem is exacerbated in contexts where the participants' knowledge is remote from common experience. For example, in the study of interactions between physicians to review whether particular patients were eligible for surgery (Boyd, 1998; Heritage et al., 2001), coding depended on considerable knowledge of upper respiratory infections and middle and inner ear conditions, together with a grasp of the review guidelines for treatment. By the same token, in the presidential news conference project, which reached back over a half-century, coders could not analyze the early questions in the database without acquiring at least some basic sociopolitical knowledge such as the identity of historical figures such as Senator Estes Kefauver (a key proponent of anti-trust legislation in the 1950s), processes involving congressional committees and floor votes, and particular legislation being targeted for inquiry. Most often the coders needed only enough background knowledge to determine whether a given question was innocuous or whether its content was adversarial toward the president, but here again access to this kind of background knowledge is essential for coding decisions.[7]

In general, coder reliability will, other things being equal, deteriorate when larger coding groups are involved. One way of addressing this problem, beyond the various techniques already identified above, is to make coding categories turn on formal criteria rather than requiring a great deal of interpretation or context-sensitive judgment. A consequence is that with larger coding teams there can be a tendency to shift towards formalized criteria and explicit decision rules with a concomitant sacrifice of validity for reliability.

Coding, as Stivers (2015) has observed, involves an unavoidable "flattening" of complex historically and temporally specific actions into general categories (see also Schegloff, 1993). Nowhere is this more obvious than when the coder encounters momentous sociopolitical events and scandals. There is of course no metric whereby the Watergate scandal, the Iran-Contra scandal, and the Monica Lewinsky scandal can be arrayed against one another in any meaningful way. The sacrifice of specificity that coding can involve can only be justified by the prospect of generalizability.

Discussion

Early in this paper we suggested that every act of naming or description carries an explicit or implicit claim of comparison and resemblance, and this in turn entails categorization. It's through this process

[7]The pervasive relevance of common sense and contextual background knowledge can also create significant difficulties in "exporting" coding schemes within and across cultures.

that reference and description escape from what might be thought of as the solipsism of the single case. Systematic coding is an explicit, disciplined, and purposeful extension of this elementary process.

While the study of discourse has been associated with a great variety of coding systems, many of them are rooted in *etic* categories; that is, categories which reflect the theoretical interests or preoccupations of the investigator. Conversation analytic coding, by contrast, is disciplined by a commitment to categories that are *emic* in character; i.e., grounded in social conduct and thus demonstrably oriented to by the participants themselves. Such grounding provides a form of validation that is distinctively based in the warp and weft of actual social life. To the best of our knowledge, CA coding is the only form of coding in the social sciences that is validated in this way (Clayman & Heritage, 2021).

We have described a variety of important ingredients for the successful management of conversation analytic coding studies. Firstly, at the phase of coding system development, we have underscored the centrality of both the position and composition of the conduct targeted for coding, while also identifying a number of further considerations that are broadly relevant across coding projects of various kinds. These include the focus and granularity of the phenomenon of interest; whether the study is designed to connect interactional practices or actions to one another, or alternatively to some aspect of their socio-interactional context, or to the outcomes of interaction.

At the system testing phase, we have stressed the necessity of extensive reviewing of coding categories using test runs with actual data to identify ambiguities and other inadequacies of the category system, as well as testing for inter-coder reliability. We also stressed the vital importance of a detailed codebook to anchor the thinking of coders in relation to the material under investigation. These measures, if successful, will establish an active and shared "culture of coding."

Stabilizing and maintaining this coding culture over time is a primary concern across the post-testing phase of research. In this context active and continuous updating of the codebook is essential to sustain the collective memory of the coding system and its rationales, and thus strengthen the cohesion of coders' decision-making. Beyond this, we also advocated for coders to work in rotating pairs (resources permitting) as a prophylactic against coder drift. Weekly meetings of the entire coding team are a further valuable resource for resolving problem cases collectively, and in the process reinforcing and updating a common culture of coding.

Finally, because successful *emic* coding is heavily reliant on knowing what the participants know to grasp the meaning of social actions, the coders' possession of relevant contextual knowledge is always and everywhere an important consideration. This can become more problematic, for example, in technical professional domains and historically remote time periods, as the commonsense knowledge of the interactional participants becomes remote from that of the coders. There is, of course, no comprehensive "solution" to this problem. Nonetheless, sustained sensitivity to and vigilance about this problem is a desirable stance which can lead to progress in all but the most recondite areas.

The principles and practices discussed in this paper are designed to maximize the extent to which quantitative investigations of interaction are built upon coding categories that are, in the tradition of conversation analysis, emically grounded. The approach we have sketched is an ideal toward which researchers in the field may gear their efforts, but there are unavoidable limits to its applicability. Many of the studies we have touched on investigate linkages between interactionally-situated conduct and either contexts or outcomes, but the coding of contexts and outcomes may not lend itself to emic validation. For instance, social categories such as gender or race or socioeconomic status, socio-economic conditions such as the business cycle or the boundaries of an historical era, or the medical appropriateness of a case for surgery – these dimensions of context are not necessarily amenable to emic grounding and may require "etic" classification that is exogenous to the orientations of the participants. Accordingly, studies working with categories of this kind require a hybrid of conversation analytic coding together with other bases of classification.

Although studies involving the coding of interactional data have progressed relatively slowly over the past twenty years, their results have been striking and at times counterintuitive. Coding studies have suggested that thanking others for assistance is infrequent (Floyd et al., 2018); that requests accompanied by *please* are similarly sparse (Chalfoun et al., 2024); that the design of questions has massive impacts in a wide variety of interpersonal and institutional contexts (Bolden et al., 2023a); and that fundamentals of conversational interaction such as turn taking and its impacts are widespread, and possibly universal, across languages and cultures (Stivers et al., 2009). It is likely that coding will remain a part of the CA toolkit for the foreseeable future.

Acknowledgments

We thank Arnulf Deppermann and Tanya Stivers, and the two ROLSI reviewers, for their valuable comments on an earlier draft of this paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Ethics statement

The data reproduced in this article were collected prior to the institution of informed consent regulations. The preponderance of data displayed here were collected through the voluntary collaboration of participating families, but the rights of some participants to agree to be recorded may not have been respected at the time of recording. Participating families were invited to erase any conversations of their choosing, and we have reason to believe that at least some of them exercised that right.

References

Bales, R. F. (1950). *Interaction process analysis: A method for the study of small groups*. Reading, Mass. Addison-Wesley.

Bolden, G., Heritage, J. & Sorjonen, M.-L. (Eds.) (2023a). *Responding to polar questions across languages and contexts*. Amsterdam, Benjamins.

Bolden, G., Heritage, J. & Sorjonen, M.-L. (2023b). Introduction: Polar questions and their responses. In G. Bolden, J., Heritage & M.-L. Sorjonen (Eds.), *Responding to polar questions across languages and contexts* (pp. 1–39). Amsterdam, Benjamins.

Boyd, E. A. (1998). Bureaucratic authority in the "company of equals:" The interactional management of medical peer review. *American Sociological Review*, *63*(2), 200–224.

Brown, P., Sicoli, M. A. & Le Guen, L. (2021). Cross-speaker repetition and epistemic stance in Tzeltal, Yucatec, and Zapotec conversations. *Journal of Pragmatics*, *183*, 256–272.

Chalfoun, A., Rossi, G. & Stivers, T. (2024). The magic word? Face-work and the functions of please in everyday requests. *Social Psychology Quarterly*: DOI: 10.1177/01902725241245141

Clayman, S. E., Elliott, M. N., Heritage, J. & Beckett, M. (2010). A watershed in white house journalism: Explaining the post-1968 rise of aggressive presidential news. *Political Communication*, *27*, 229–247.

Clayman, S. E., Elliott, M. N., Heritage, J. & McDonald, L. (2006). Historical trends in questioning presidents 1953-2000. *Presidential Studies Quarterly*, *36*, 561–583.

Clayman, S. E. & Fox, M. P. (in press). Asking about bias: Managing the sensitivities of voir dire questioning. *Language in Society*.

Clayman, S. E. & Heritage, J. (2002a). Questioning presidents: Journalistic deference and adversarialness in the press conferences of Eisenhower and Reagan. *Journal of Communication*, *52*(4), 749–775.

Clayman, S. E. & Heritage, J. (2002b) *The news interview: Journalists and public figures on the air*. Cambridge: Cambridge University Press.

Clayman, S. E. & Heritage, J. (2021). Conversation analysis and the study of sociohistorical change. *Research on Language and Social Interaction*, *54*(2), 225–240.

Clayman, S. E. & Heritage, J. (2023). Pressuring the president: Changing language practices and the growth of political accountability. *Journal of Pragmatics*, *207*, 62–74.

Clayman, S. E., Heritage, J., Elliott, M. N. & McDonald, L. (2007). When does the watchdog bark?: Conditions of aggressive questioning in presidential news conferences. *American Sociological Review*, *72*, 23–41.
Clayman, S. E., Heritage, J. & Maynard, D. (2022). The ethnomethodological lineage of conversation analysis. In D. Maynard. & J. Heritage (Eds.), *The ethnomethodological program: Legacies and prospects* (pp. 252–286). New York, Oxford University Press.
Clayman, S. E. & Raymond, C. W. (2021). An adjunct to repair: You know in speech production and understanding difficulties. *Research on Language and Social Interaction*, *54*(1), 1–21.
Clift, R. & Raymond, C. W. (2018). Actions in practice: On details in collections. *Discourse Studies*, *20*(1), 90–119.
Comstock, L. (2023). Journalistic practice in the international press corps: Adversarial questioning of the Russian President. *Journal of Language Aggression and Conflict*, *11*(2), 145–175.
Curl, T. S. (2006). Offers of assistance: Constraints on syntactic design. *Journal of Pragmatics*, *38*, 1257–1280.
Davidson, J. A. (1984). Subsequent versions of invitations, offers, requests, and proposals dealing with potential or actual rejection. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 102–128). Cambridge, Cambridge University Press.
Deppermann, A. (2018). Changes in turn-design over interactional histories - the case of instructions in driving school lessons. In A. Deppermann & J. Streeck (Eds.), *Time in embodied interaction. Synchronicity and sequentiality of multimodal resources* (pp. 293–324). Amsterdam: Benjamins.
Deppermann, A. & Pekarek Doehler, S. (Eds.) (2021). Longitudinal CA: How interactional practices change over time. *Research on Language and Social Interaction* 54 (2).
Drummund, K. & Hopper, R: (1993). Backchannels revisited: Acknowledgment tokens and speakership incipiency. *Research on Language and Social Interaction*, *26*, 157–177.
Enfield, N. J., Stivers, T., Brown, P., Englert, C., Harjunpää, K., Hayashi, M., Heinemann, T., Hoymann, G., Keisanen, T., Rauniomaa, M., Raymond, C. W., Rossano, F., Yoon, K.-E., Zwitserlood, I. & Levinson, S. C. (2019). Polar answers. *Journal of Linguistics*, *55*, 277–304.
Enfield, N. J., Stivers, T. & Levinson, S. C. (2010). Question-response sequences in conversation across ten languages. *Journal of Pragmatics*, *42*, 2615–2860.
Floyd, S., Rossi, G., Baranova, J., Blythe, J. Dingemanse, M., Kendrick, K. H., Zinken, J. & Enfield, N. J. (2018). Universals and cultural diversity in the expression of gratitude. *Royal Society Open Science*, *5*(5), 180391.
Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ, Prentice Hall.
Goffman, E. (1955). On face work. *Psychiatry* 18: 213–231.
Goodwin, C, (1981). *Conversational organization: Interaction between speakers and hearers*. New York, Academic Press.
Heritage, J. (1984). *Garfinkel and ethnomethodology*. Cambridge, Polity Press.
Heritage, J. (2015). *Well*-prefaced turns in english conversation: A conversation analytic perspective. *Journal of Pragmatics*, *88*, 88–104.
Heritage, J. (2018). Turn-initial particles in English: The cases of *Oh* and *Well*. In J. Heritage & M.-L. Sorjonen (Eds.), *Turn-initial particles across languages* (pp. 149–183). Amsterdam: Benjamins.
Heritage, J. (2021). The expression of authority in US primary care: Offering diagnoses and recommending treatment. In C. Gordon (Ed.), *Approaches to discourse analysis* (pp. 104–122). Georgetown, Georgetown University Press.
Heritage, J. & Atkinson, J. M. (1984). Introduction. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 1–16). Cambridge, Cambridge University Press.
Heritage, J., Boyd, E. A. & Kleinman, L. (2001). Subverting criteria: The role of precedent in decisions to finance surgery. *Sociology of Health and Illness*, *23*(5), 701–728.
Heritage, J. & Clayman, S. E. (2024). Making arrangements: A sketch of a 'big package'. *Research on Language and Social Interaction*, *57*(3), 279–300.
Heritage, J. & Greatbatch D. (1986). Generating applause: A study of rhetoric and response at party political conferences. *American Journal of Sociology*, *92*(1), 110–157.
Heritage, J. & McArthur, A. (2019). The diagnostic moment: A study in US primary care. *Social Science and Medicine*, *228*, 262–271.
Heritage, J. & Raymond, G. (2012). Navigating epistemic landscapes: Acquiescence, agency and resistance in responses to polar questions. In J. P. de Ruiter (Ed.), *Questions: formal, functional and interactional perspectives* (pp. 179–192). Cambridge, Cambridge University Press.
Heritage, J. & Robinson, J. D: (2006). The structure of patients' presenting concerns: Physicians' opening questions. *Health Communication*, *19*(2), 89–102.
Heritage, J, Robinson, J. D., Elliott, M., Beckett, M. & Wilkes, M. (2007). Reducing patients' unmet concerns: The difference one word can make. *Journal of General Internal Medicine*, *22*, 1429–1433.
Heritage, J. & Roth, A. (1995). Grammar and institution: Questions and questioning in the broadcast news interview. *Research on Language and Social Interaction*, *28*(1), 1–60.
Hill, R. J. & Stones Crittenden, K. (1968). *Proceedings of the purdue symposium on ethnomethodology*. Purdue IN, Purdue Research Foundation.

Jefferson, G. (1989). Preliminary notes on a possible metric which provides for a 'standard maximum' silences of approximately one second in conversation. In D. Roger & P. Bull, P. (Eds.), *Conversation: An interdisciplinary perspective* (pp. 166–196). Clevedon, Multilingual Matters.
Jefferson, G. (1993). Caveat speaker: Preliminary notes on recipient topic-shift implicature. *Research on Language and Social Interaction*, *26*, 1–30.
Jefferson, G. (2004). A note on laughter in 'male-female' interaction. *Discourse Studies*, *6*(1), 117–133.
Kendrick, K. H., & Drew, P. (2014). The putative preference for offers over requests. In P. Drew & E. Couper-Kuhlen (Eds.), *Requesting in social interaction* (pp. 87–114). John Benjamins.
Kendrick, K. H. & Holler, J. (2017). Gaze direction signals response preference in conversation. *Research on Language and Social Interaction*, *50*(1), 12–32.
Kendrick, K. H. & Torreira, F. (2015). The timing and construction of preference: A quantitative study. *Discourse Processes*, *52*, 255–289.
Kent, A. & Antaki, C. (2020). Police call-takers' first substantive question projects the outcome of the call. *Applied Linguistics*, *41*, 640–661.
Kim, H. R. S. (2013). Retroactive indexing of relevance: The use of well in third position. *Research on Language and Social Interaction*, *46*(2), 125–143.
Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*, Sage.
Lombard, M., Snyder-Duch, J. & Campanella Bracken, C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, *28*(4), 587–604.
Mangione-Smith, R, Zhou, C., Robinson, J. D., Taylor, J. A., Elliott, M. & Heritage, J. (2015). Communication Practices and antibiotic use for acute respiratory tract infections in children. *Annals of Family Medicine*, *13*, 221–227.
Mervis, C. B. & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, *32*(1), 89–115.
Mondada, L. (2023). The situated and methodic production of accountable action: The challenges of multimodality. In D. Maynard & J. Heritage (Eds.), *The ethnomethodology program: Legacies and prospects* (pp. 289–321). New York, Oxford University Press.
Parsons, T. & Bales, R. F. (1955). *Family, socialization and interaction process*. Glencoe: IL, Free Press.
Peräkylä, A. (1998). Authority and accountability: The delivery of diagnosis in primary health care. *Social Psychology Quarterly*, *61*(4), 301–320.
Peräkylä, A. (2002). Agency and authority: Extended responses to diagnostic statements in primary care encounters. *Research on Language and Social Interaction*, *35*(2), 219–247.
Peräkylä, A. (2006). Communicating and responding to diagnosis. In J. Heritage & D. Maynard (Eds.), *Communication in Medical Care: Interactions between primary care physicians and patients* (pp. 214–247). Cambridge, Cambridge University Press.
Pillet-Shore, D. (2012). Greeting: Displaying stance through prosodic recipient design. *Research on Language and Social Interaction*, *45*(4), 375–398.
Pillet-Shore, D. (2017). Preference organization. In J. Nussbaum (Ed.), *Oxford research encyclopedia of communication*. New York: Oxford University Press. DOI: 10.1093/acrefore/9780190228613.013.132.
Pillet-Shore, D. (2024). Where the action is: Positioning matters in interaction. In J. D. Robinson, R. Clift, K. H. Kendrick & C. W. Raymond (Eds.). *The Cambridge handbook of methods in conversation analysis* (pp. 577–611). Cambridge University Press.
Pomerantz, A. M. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 57–101). Cambridge, Cambridge University Press.
Pomerantz, A. M. & Heritage, J. (2013). Preference. In J. Sidnell & T. Stivers (Eds.), *Handbook of conversation analysis* (pp. 210–228). New York, Wiley-Blackwell.
Pomerantz, A. M. & Rintel, S. (2004). Practices for reporting and responding to test results during medical consultations: Enacting the roles of paternalism and independent expertise. *Discourse Studies*, *6*(1), 9–26.
Raymond, C. W., Clift, R. & Heritage, J. (2021). Reference without anaphora: On agency through grammar. *Linguistics*, *59*(3), 715–755.
Raymond, G. (2003). Grammar and social organization: Yes/no interrogatives and the structure of responding. *American Sociological Review*, *68*, 939–967.
Robinson, J. D. (2006). Soliciting Patients' presenting concerns. In J. Heritage & D. Maynard, D. (Eds.), *Communication in Medical Care: Interactions between primary care physicians and patients* (pp. 22–47). Cambridge, England, Cambridge University Press.
Robinson, J. D. (2024). Coding and statistically associating inter-action to advance conversation-analytic findings. In J. D. Robinson, R. Clift, K. H. Kendrick. & C. W. Raymond, C. W. (Eds.). *The Cambridge handbook of methods in conversation analysis* (pp. 452–484). Cambridge University Press.
Robinson, J. D., Clift, R., Kendrick, K. H. & Raymond, C. W (Eds.) (2024). *Handbook of research methods in conversation analysis*. Cambridge, Cambridge University Press.
Robinson, J. D. & Heritage, J. (2006). Physicians' opening questions and patients' satisfaction. *Patient Education and Counseling*, *60*, 279–285.

Robinson, J. D., Rühlemann, C. & Taylor Rodriguez, D. (2022). The bias toward single-unit turns in conversation. *Research on Language and Social Interaction*, *55*(2), 165–183.
Robinson, J. D., Tate, A. & Heritage, J. (2016). Agenda-setting revisited: When and how do primary-care physicians solicit patients' additional concerns? *Patient Education and Counseling*, *99*(5), 718–723.
Rosch, E. (1987). Wittgenstein and categorization research in cognitive psychology. In: M. Chapman & R. A. Dixon (Eds.), *Meaning and the growth of understanding* (pp. 151–166). Berlin, Springer.
Rossi, G. (2012). Bilateral and unilateral requests: The use of imperatives and *Mi X?* In Italian. *Discourse Processes*, 49, 426–458.
Rossi, G. (2025). *Systems of social action: The case of requesting in italian*. New York: Oxford University Press.
Rossi, G., Dingemanse, M., Floyd, S., Baranova, J., Blythe, J. Kendrick, K. H., Zinken, J. & Enfield, N. J. (2023). Shared cross-cultural principles underlie human prosocial behavior at the smallest scale. *Scientific Reports*, *13*(1), 6057.
Roter, D. & Larson, S. (2002). The Roter Interaction Analysis System (RIAS): Utility and flexibility for analysis of medical Interactions. *Patient Education and Counseling*, *42*, 243–251.
Sacks, H. (1984). Notes on Methodology. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 21–27). Cambridge: Cambridge University Press.
Sacks, H. (1992 [1964-72]). *Lectures on conversation (2 Vols.)*. Oxford, Basil Blackwell.
Sacks, H., Schegloff, E. A. & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*, 696–735.
Schaeffer, N. C., Garbarski, N., Freese, J. & Maynard, D. W. (2013). An interactional model of the call for survey participation: Actions and reactions in the survey recruitment call. *Public Opinion Quarterly*, *77*(1), 323–351.
Schegloff, E. A. (1968). Sequencing in Conversational Openings. *American Anthropologist*, *70*, 1075–1095.
Schegloff, E. A. (1986). The routine as achievement. *Human Studies*, *9*, 111–151.
Schegloff, E. A. (1987). Analyzing single episodes of interaction: An exercise in conversation analysis. *Social Psychology Quarterly*, *50*(2), 101–114.
Schegloff, E. A. (1993). Reflections on quantification in the study of conversation. *Research on Language and Social Interaction*, *26*, 99–128.
Schegloff, E. A. (1996). Confirming allusions: Toward an empirical account of action. *American Journal of Sociology*, *102* (1), 161–216.
Schegloff, E. A. (1997). Practices and actions: Boundary cases of other-initiated repair. *Discourse Processes*, *23*(3), 499–545.
Schegloff, E. A. (2009). One perspective on conversation analysis: Comparative perspectives. In J. Sidnell (Ed.), *Conversation Analysis: Comparative Perspectives* (pp. 357–406). Cambridge, Cambridge University Press.
Schegloff, E. A. & Sacks, H. (1973). Opening up closings. *Semiotica*, *8*, 289–327.
Stivers, T. (2002). Presenting the problem in pediatric encounters: "Symptoms only" versus "candidate diagnosis" presentations. *Health Communication*, *14*(3), 299–338.
Stivers, T. (2015). Coding social interaction: A heretical approach in conversation analysis? *Research on Language and Social Interaction*, *48*(1), 1–19.
Stivers, T. (2022). *The book of answers: Alignment, autonomy, and affiliation in social interaction*. Oxford, Oxford University Press.
Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoon, K.-E. & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, *106*(26), 10587–10592.
Stivers, T., Heritage, J., Barnes, R. K., McCabe, R., Thompson, L. & Toerien, M. (2018). Treatment recommendations as actions. *Health Communication*, *33*(11), 1335–1344.
Stivers, T. & Majid, A. (2007). Questioning children: Interactional evidence of implicit bias in medical interviews. *Social Psychology Quarterly*, *70*(4), 424–441.
Stivers, T., Mangione-Smith, R., Elliott, M. N., McDonald, L. & Heritage, J. (2003). Why do physicians think parents expect antibiotics? what parents report vs what physicians perceive. *The Journal of Family Practice*, *52*(2), 140–148.
Stivers, T. & Sidnell, J. (2016). Proposals for Activity Collaboration. *Research on Language and Social Interaction*, *49*(2), 148–166.
West, C. & Zimmerman, D. (1983). Small insults: A study of interruptions in cross-sex conversations with unacquainted persons. In B. Thorne, C. Kramarae & N. Henley (Eds.) *Language, Gender and Society* (pp. 102–117). Rowley MA, Newbury House.
Zimmerman, D. H. (1993). Acknowledgment tokens and speakership incipiency revisited. *Research on Language and Social Interaction*, *26*(2), 179–194.
Zimmerman, D. H. & West, C. (1975). Sex roles, interruptions and silences in conversation. In B. Thorne & N. Henley (Eds.), *Language and Sex: Difference and Dominance* (pp. 105–129). Rowley, Mass. Newbury House.