The Role of Numbers and Statistics within Conversation Analysis

Jeffrey D. Robinson

To cite this article: Jeffrey D. Robinson (2007) The Role of Numbers and Statistics within Conversation Analysis, Communication Methods and Measures, 1:1, 65-75, DOI: 10.1080/19312450709336663

To link to this article: http://dx.doi.org/10.1080/19312450709336663

Published online: 05 Dec 2007.
The Role of Numbers and Statistics within Conversation Analysis

Jeffrey D. Robinson
Rutgers University

In this article, I discuss the role of numbers and statistics within Conversation Analysis (hereafter CA). Due to space limitations, I do not review CA (see Heritage, 1984). Furthermore, I do not focus on CA’s concerns regarding the use of quantification (e.g., ratios) to derive or operationalize statistically manipulable variables (see Schegloff, 1993). Finally, I do not focus on CA’s concerns regarding the validity of statistically associating variables, especially when they are not demonstrably relevant to, or procedurally consequential for, participants (see Schegloff, 1992). The latter enterprise is sometimes called applied CA when one of the variables is a product of CA. However, applied CA is a misnomer because it implies that such work represents a type of CA, which it does not. Rather, it represents an application of CA (which takes nothing away from its value; see Heritage, 2004).

Instead, I address numbers and statistics in terms of proving certain types of CA claims, with the hope of clarifying ambiguity and increasing analytic rigor. I begin by reviewing CA’s conception of order and then outline two different, yet interrelated, CA projects: analyzing single cases and documenting practices of action. The latter project involves claims about communication rules that generate regular patterns of understanding and interactional organization. As such, achieving this goal involves evidentiary requirements that implicate numbers and statistics. Conversation analysts should (perhaps reluctantly) find these requirements to be methodologically acceptable, although they are neither easily nor typically satisfied, at least in single articles/chapters.
ORDER AT ALL POINTS

Harvey Sacks, the cofounder of CA (with Emanuel Schegloff and Gail Jefferson), lamented that “[t]he important theories in the social sciences have tended to view a society as a piece of machinery with relatively few orderly products, where, then, much of what else takes place is more or less random” (1984, p. 21). Sacks’s observation applies to more recent views of interaction by communication scholars: “There are an infinity of examples of interactional sequences, the vast majority of which must be random pairings rather than regular pairings” (Cappella, 1990, p. 239). Contrary to this view, Sacks argued that, in every bit of conduct-in-interaction,¹ analysts will find detailed order in terms of interactants’ production, recognition, and interpretation of their own and others’ actions. People are irremediably making sense all the time, or as Sacks put it, there is “order at all points” (p. 22). The methodological consequence is that researchers cannot, a priori, assume that interaction is idiosyncratic or that any order of interactional detail is disorderly, accidental, or irrelevant. The argument for order at all points is supported by 40 years of research showing that the ability of communities of humans (and, in some cases, all humankind) to “make sense” in and of inter-“action” is largely predicated on the existence of socially conventionalized (i.e., recognizable and predictable) orchestrations of conduct-in-interaction that are understood, by all members, to make particular communication norms/rules relevant and accountable. I strongly encourage readers to review Heritage (1984) for CA’s conception of rules, which differs from others.

SINGLE CASES VERSUS PRACTICES OF ACTION

CA primarily involves two different, but necessarily interrelated, projects: (a) the analysis of (multiple) single cases, and (b) the systematic documentation of practices of action (Schegloff, 1987).

Analyzing Single Cases

The integrity of single-case analysis is grounded in the assumption of order-at-all-points. Here, numbers of cases is not important. In single cases, analysts attempt to demonstrate, from participants’ conduct-in-interaction, that partici-

¹Conduct-in-interaction is short for talk-and-other-conduct-in-interaction. CA’s definition of conduct includes vocal and nonvocal behavior, the use of artifacts, and any other feature of context that participants can be shown to orient to (vocally or nonvocally), including identities, roles, and relationships.
pants understand particular features of interaction in particular ways. These data-internal, or emic (vs. etic), understandings are assumed to reflect orderly processes, and thus are used to make claims about rules-based structures of interaction (Goodwin, 1984; Jacobs, 1986; Schegloff, 1987). For example, see Extract 1, which was recorded in a private British household and involves an interaction between a husband, his wife who recently gave birth, and a federally sponsored health visitor (HV) hired to ask particular questions and give post-natal advice. At lines 2–3, the HV addresses the wife’s employment status.

Extract 1: NURSE [HV:1A1:20:50]
01 (5.3)
02 HV: Were you a uh:m (1.5) what- (1.0) uh: (.)
03 you were a nurse at thuh churchill.
04 (6.3)
05 HUS: Yea[h
06 WIF: [Yea:h
07 (2.6)
08 HV: Are you going to go ba:ck,
09 HUS: Mm hm,
10 WIF: Yea:h
11 (6.3)
12 HV: Uhm (.) an ‘S’ ‘E’ ‘N’.
13 WIF: Yeh(p)
14 (2.6)
15 HUS: [.mhh
16 WIF: I’ve taken maternity lea:ve,

I make two observations. First, in response to the HV’s “you were a nurse at thuh churchill.” (line 3), both the husband and wife respond sufficiently with the same action (i.e., confirmation) achieved by the same token (i.e., Yeah; lines 4 & 5). Second, in response to the HV’s question “Are you going to go ba:ck,” (line 10), which is designed for the wife, the husband not only responds (affirmatively; see “Mm hm” at line 11), but responds first. Both of these observations can be used to make claims about the (normative) structure of interaction. Regarding the first observation, the spouses’ responsive actions display their understandings that the HV’s turn (at lines 2–3) performs a particular action: a simple request for confirmation. Readers familiar with studies of discourse will recognize the HV’s turn as a B-event statement (Labov & Fanshel, 1977) and that the spouses’ responses conform to its rules. Regarding the second observation, the husband treats himself as having primary rights to answer the HV’s question. As is discussed later, readers familiar with discourse studies will recognize that this behavior violates rules. In both cases, analysts can begin to examine the HV’s conduct for how it was orchestrated so as to make relevant such understandings, and for the normative underpinnings of such orchestrations.
Documenting Practices of Action

CA is primarily interested in identifying practices of social action, meaning that they are intersubjectively understood and normatively binding across at least a delimited range of contexts and participants. A practice of action is a structured orchestration of multiple aspects of conduct-in-interaction that is regularly produced and understood as implementing a particular action (or set of actions)\(^2\) that regularly instantiates a particular set of norms/rules, and thus that is regularly associated with patterns of conduct that evidence participants’ understandings of the operation of such rules. CA is only secondarily concerned with idiosyncratic rules (Sigman, 1980), or those shared uniquely by a single dyad, such as two relational partners.

Analysts carry the burden of proving the regularities of practices, and this cannot be achieved with a single case. As Schegloff (1988) noted, although single cases can serve “to launch a proposal” (p. 442) about a practice of action, this proposal is just “a conjecture” (p. 442) until “a substantial number of occurrences” (p. 451) can be assembled. This point is highlighted by the second observation made about Extract 1. From prior research (Lerner, 2003), we know that the HV’s “Are you going to go ba:\-ck,” (line 10) implements the practice of selecting a particular next speaker to respond next; it allocates the right and obligation to respond next to the wife (and only the wife). That the husband responds first, and immediately, violates these rules. Analyzed in isolation, Extract 1 stands as (possible) data-internal evidence against this turn-allocation practice; it leads analysts astray from validly characterizing the practice.\(^3\) Of importance, this does not mean that the husband’s response was idiosyncratic or disorderly (more on this is later). In sum, “proving” a practice requires numbers of cases. Here, conversation analysis and distributional analysis are interdependent (vs. incommensurate) enterprises.

DISTRIBUTIONAL EVIDENCE FOR PRACTICES OF ACTION

Completely documenting an empirically grounded account of a practice of action is an arduous analytic task, involving sorting through large numbers (e.g., hun-

---

\(^2\)There is a distinction between a practice and a practice of action, the former being used to build the latter. For example, multiple practices of turn design, lexical choice, intonation, and sequential position frequently get orchestrated, in context-sensitive ways, to achieve single practices of action. Participants are concerned with the product of this orchestration in terms of action, whereas analysts are concerned in addition with the practices used to build actions.

\(^3\)Lest we become overly pessimistic about the reliability of single cases, re-consider the first observation made about Extract 1. From prior research (Labov & Fanshel, 1977), we know that the HV’s “you were a nurse at thuh churchill.” (line 3) implements a practice called a B-event statement. Although space limitations prevent a full analysis, the husband’s and wife’s responses (lines 4–5) stand as data-internal evidence for this practice.
dreds) of cases, some of which are discarded as noninstances of the practice, others of which are set aside as boundary instances (Schegloff, 1997), and others of which become core instances, or ones that contain all of the practice’s constitutive features (i.e., the core collection). Readers are directed to Schegloff (1996) for an example of this process that includes methodological requirements from a CA perspective. It is extremely important that readers understand that the upcoming recommendations regarding numbers and statistics only apply to the analysis of practices of action (vs. single cases), and only apply after researchers have completely documented a practice from a CA perspective. CA is not, in principle, opposed to coding and counting, but rather to doing so at the expense of detailed single-case analyses, and thus at the expense of interactional phenomena as oriented to by participants. As Schegloff (1993) noted, “in examining large amounts of data, we are studying multiples or aggregates of single instances,” and “[q]uantitative analysis is, in this sense, not an alternative to single case analysis, but rather is built on its back” (p. 102; emphases original).

Within core collections, a vast array of different types of cases can support analysts’ claims about practices. For example, take the practice of allocating turns to specific next speakers. One very frequent pattern is for selected speakers to respond immediately next, as in Extract 2, which is drawn from an American dinner conversation between two couples (Beth & John; Ann & Don).

**Extract 2: HOUSE [Chinese Dinner]**

01 BET: Did they do a lot with it though or did they
02 buy it kinda intact. ((referring to a house))
03 DON: They bought it intact. (. ) re:al [ly.]
04 ANN: [Mm: ] [An’]
05 ANN: you pick out all thuh stuff ahead. an’ they do it all
06 BET: [Oh y’=mean it’s li]ke all the o- (0.2)
07 ANN: [Each room. ]
08 Q-> BET: Is it a bran’ new hou:se* (((=Gazing at Ann))
09 A1> ANN: Mm hm?
10 (0.7)
11 A2> DON: It’s a bran’ new house. It’s- z- it’s on a development...

At line 8, Beth implements the practice of selecting Ann to speak next by addressing a question to Ann, gazing directly at Ann at the end of the turn, and using the pronoun you (Lerner, 2003). Note that Ann answers (confirms) first, and immediately (A1>), even though Don is able to, and eventually does, provide his own response/answer (A2>).

Although cases that accord with the claimed rules of a practice tend to be the most frequent, showing that participants conform to rules is not sufficient to demonstrate the normality of rules. In addition, analysts must show that participants
orient to rules in the face of their being flouted, violated, and so on. For example, see Extract 3, which is drawn from the same data as Extract 2. When non-selected speakers respond next, they tend to do so only after selected speakers demonstrate some trouble with responding and thus after some response delay (Stivers & Robinson, 2006). At line 1, Beth is discussing the purchase of her new dishes, which are on the table.

Extract 3: ORANGE [Chinese Dinner]

01 BET: They didn’t have all the colors, .hh the orange
02 is really nice but they only had it in, .hh
03 these bowls, and uhm, (0.5) (the coffee mugs).
04 DON: (Mmkh) ((possibly an ‘eating’ noise))
05 Q-> ANN: Which is orange.* ((*=Gazing at Beth))
06 *-> (0.5)
07 A1> DON: The reddy orange.
08 A2> BET: This one.
09 ANN: Oh:.

At line 5, Ann implements the practice of selecting Beth to respond next by addressing a (repair-related) question to Beth and gazing directly at Beth at the end of the turn (Lerner, 2003). Don has neither the right nor obligation to respond next. At the end of Ann’s turn (i.e., at “or…” of “orange.”), Don is gazing directly at Beth and can see that she is gazing down at her plate (i.e., not at Ann), where Beth continues to gaze for at least several tenths of a second into the silence at line 6. Although Don responds first (A1>), he does so only after waiting through a significant silence (*>-) in which Beth accountably does not respond (i.e., after Beth displays generic trouble responding). Cases like Extract 3 not only “prove the rules” of the practice, but highlight their normativity, because nonselected recipients (e.g., Don) orient to the norm that selected recipients (e.g., Beth) respond first.

In addition to cases represented by Extracts 2 and 3, others can “prove the rules,” such as cases in which rule violations are sanctioned by participants (i.e., deviant cases). In sum, there will be different types or groups of cases—and, in all likelihood, contextually idiosyncratic single cases—that confirm analysts’ claims about practices. Although not the focus of this article, knowing the relative frequency of different types of confirmatory cases within a single practice can be important. In publications, largely due to space constraints, conversation analysts tend to present prototypical cases as evidence. However, as Jacobs (1988) noted, “Neither the fact that an example is prototypical in appearance nor the fact that it was recorded from naturally occurring conversation warrants the inference that [it is] typical in occurrence” (p. 437; emphasis added). Knowledge of different types of confirmatory cases and their relative frequency aids analysts’ understandings of
practices themselves, and of the (proto)typicality of their supporting cases (Jacobs, 1986).

Having discussed confirmatory cases, my main point is that, in virtually every core collection, there will be cases that have two characteristics: (a) Participants’ conduct disconfirms the claimed rules of the practice; and (b) There is no data-internal evidence (that analysts can find) that explains or accounts for the rule deviation. Stated differently, there will almost always be unexplainable disconfirmatory cases that do not support, and sometimes actively run against, analysts’ claims about the operation of practices. Of importance, these cases do not necessarily invalidate analysts’ characterizations of practices, nor the assumption of order-at-all-points. That is, participants do not have to behaviorally orient to rules (and thus provide data-internal evidence) in order to understand rules and their accountable implications, and this is just “tough methodological luck” for conversation analysts. For example, regarding the practice of selecting particular next speakers, Extract 1 is (currently) an unexplainable disconfirmatory case. At line 10, the HV’s question contextually selects the wife, and rules out the husband, as a respondent. Nonetheless, the husband responds first, and immediately (line 11). There is no data-internal evidence (that I have yet found) that participants orient to the husband’s answer as nonnormative or accountable in any way. However, this does not necessarily mean that the husband’s answer is not orderly or meaningful for participants. Regarding order-at-all-points, the husband’s answer may be orderly precisely by reference to its violation of the rules instantiated by the practice of speaker selection. For instance, when the husband answers for his wife without allowing her the rightful chance to answer for herself, he may be understood as constructing a particular identity for himself, such as ‘dominant.’

Personally, I have not encountered a core collection that does not have at least one unexplainable disconfirmatory case. Frankly, their absence is analytically suspect. A key feature of analytic induction is the search for such cases, which force analysts to re-assess the claimed essential constitutive features and/or rules of practices. Thus, unexplainable disconfirmatory cases are not bad in any sense of the word (e.g., they do not represent “bad variance,” akin to “bad cholesterol”). Rather, for CA, they are reality.

One major consequence of unexplainable disconfirmatory cases is that, although they do not necessarily invalidate analysts’ characterizations of practices, they do not definitively (i.e., empirically) support their claimed operation. Because

---

4There is only audio data for Extract 1. It is possible to “dream up” (plausible) reasons why the husband responds first, such as that his wife is occupied with their baby (note, though, that this reason is contraindicated by the fact that the wife responds first at line 8 and, given the observations made about Extract 3, does not account for why the husband responds immediately at line 11). However, such reasons are not empirically demonstrable, and if analysts choose to include audio-only data in their core collection, Extract 1 would be an unexplainable disconfirmatory case.
practices inherently involve claims of “regularity,” analysts must demonstrate: *For all cases in a core collection, those that empirically support the claims of the practice (Group A) should be more frequent than unexplainable disconfirmatory cases (Group B)*. Note that coding cases into these groups is motivated purely by CA concerns. The italicized statement represents a hypothesis (\( \text{Ha} \)), and testing it requires some type of statistical method for dealing with quantitative evidence (Cappella, 1990; Jacobs, 1986). The method of comparing percentages is not persuasive according to extant social-scientific standards because it does not prove that the two group frequencies are significantly different from each other. A more persuasive statistical method is a test of group differences that accommodates categorical data, such as the *chi-square goodness-of-fit test*. Failure to demonstrate that Group-A cases are more frequent than Group-B cases—that is, failure to reject the null hypothesis (\( \text{H}_0 \))—is (one major type of) evidence against analysts’ claims, and should send them “back to the CA drawing board.” Admittedly, statistically rejecting \( \text{H}_0 \) does not sufficiently prove \( \text{Ha} \); doing so is only one part of a larger package of evidence, most of which is conversation analytic that must (from a CA perspective) precede statistical testing.

**Sampling**

Most statistical tests, such as *chi-square*, assume a *random sample*. As is the case with much social-scientific research utilizing college-student populations, most CA data are nonrandom, convenience samples. In such cases, statistical results are crude rules of thumb, albeit ones that are commonly accepted (and sometimes demanded) by publication outlets. Conversation analysts may be tempted to use the difficulty of random sampling as an excuse to abandon statistical testing. However, doing so does not relieve analysts of the burden of rigorously testing \( \text{Ha} \) in some other way. An alternative is to address the hurdle of random sampling in more traditional ways by striving to minimize sampling error. This alternative, advocated by conversation analysts more than 20 years ago (Heritage & Atkinson, 1984), is achievable for CA purposes. The availability of national and international corpora of natural, conversational data (e.g., *The Linguistic Data Consortium*; www.ldc.upenn.edu), as well as of innovative field-sampling techniques, has made random sampling more attainable. *When making claims about practices, conversation analysts should provide information about their core-collection samples.*

**Number of Cases**

The rigorous implementation of statistical tests requires, and thus suggests to analysts, certain sample sizes. For instance, concerning a *one-way chi-square goodness-of-fit test* (i.e., to test \( \text{Ha} \), previously noted), to detect medium effect sizes at power = .80 for alpha = .05, Cohen (1992) suggested that (adequately sampled)
core collections should include at least 87 cases. *Note that many published CA studies achieve or exceed this*, including those involving ordinary (vs. institutional) data. Admittedly, Cohen’s sample-size requirements are especially challenging when dealing with “rare” practices. That is, CA is *not* concerned with the frequency with which practices of action occur in the social world. Some practices are more common than others, and their frequency does not necessarily indicate their importance. As Jacobs (1986) noted, “What people do a lot may be less important than what people can do, but don’t do very often” (p. 152). However, not having enough cases does not exempt analysts from rigorous distributional evidentiary requirements. Unless a case can be made that practices are sufficiently “rare,” analysts should collect more cases. *When making claims about practices, conversation analysts should report the numbers of cases in their core collections, including the numbers of unexplainable disconfirmatory cases.*

“Obviousness” and the Need for Distributional Evidence

Without prejudice to their empirical and theoretical importance, claims about some practices are relatively “vernacularly obvious,” at least once exposed by analysts. For example, Raymond (2003) demonstrated that grammatically formatted *Yes/No* questions prefer type-conforming responses (i.e., some version of *Yes* or *No*, respectively). In contrast, the claims of other practices are less obvious, such as that the acknowledgment token *Yeah* projects a shift in participatory status from recipient to speaker (Jefferson, 1984). In my experience, the more that claims about practices are vernacularly obvious, the less that analysts seek (Raymond excluded)—and the less that critics demand from a few (proto)typical cases—distributional evidence. However, CA has always taught that sufficient proof lies beyond plausible, pleasing, and/or parsimonious intuitions.

**CONCLUSION**

Conversation analysts tend to engage in two different, but interrelated, types of projects: analyzing (multiple) single cases, and documenting practices of action. There are standards of quality for both, and both are equally worthy of publication. I argue that the *latter* project (and not the former) necessitates statistical evidence, while recognizing that such evidence is only one component of proof, and that the validity of statistical evidence is only as good as that of the practice being investigated.

It should be mentioned that a complete account of a practice of action includes not only *showing that* it works, but also *explaining how* it works (Schegloff, 1996). The job of explanation is only partially assisted by statistical evidence. Additional help comes from examining noncore, *boundary* cases (Schegloff, 1997), such as
those in which participants demonstrably avoid using a practice (e.g., they start to use it, but then abandon it for another practice). Help also comes from examining alternative practices that accomplish similar classes of action. For example, Schegloff’s efforts to explain the practice of confirming an allusion (Schegloff, 1996) were assisted by examining different practices of “agreeing.” Note, though, that comparisons of the operation of different CA practices do require statistical evidence. These types of studies tend to be criticized not for comparing, but for doing so without first thoroughly analyzing the practices themselves.

In my experience, when documenting practices of action, the greatest challenges do not come from statistics, but from CA itself. Prior to statistical testing, analysts need to be able to answer at least the following questions in specific terms: What is the claimed practice (i.e., what are its constitutive features as an orchestra-

ACKNOWLEDGMENTS

Earlier incarnations of this article benefited from comments from John Heritage, Celia Kitzinger, and Geoff Raymond; all weaknesses are my own.

REFERENCES