ARTICLE ACCEPTED FOR PUBLICATION

discourse analyst and her primary research focus is on the relationship between discourse and identity. She is the co-author (with Elizabeth Stokoe) of *Discourse and Identity* and has published articles and chapters on discourse and reader identity, on discourses and representations of masculinity in popular culture, on tutorial discourse and student identity and on healthcare and health complaints interactions. She conducted a pilot study on interactional approaches to complaints to the NHS in Scotland with May McCreaddie, and is now co-investigator on the NIHR funded project: 'Enhancing the patient complaints journey: harnessing the power of language to transform the experience of complaining' with colleagues in the universities of Ulster, Stirling, Queen Margaret and Loughborough, which uses conversation analysis to understand the longitudinal experience of complaining and to develop training materials for healthcare professionals handling complaints.

Catrin S. Rhys is Senior Lecturer in linguistics and Head of School in the School of Communication and Media at Ulster University. She works in a Conversation Analytic framework to examine language and social interaction in a range of different institutional and mundane settings. Having come from a background in formal linguistics, with a particular interest in the syntax semantics interface, there is an emphasis in her research on the interaction between the interactional and the linguistic properties of language in use. Her current focus is the Real Complaints project: an NIHR funded project researching the language of complaints handling in the NHS with colleagues at the Universities of Stirling, Loughborough and Queen Margaret

Ruth Parry recently retired as Professor of Human Communication and Interaction at Loughborough University UK. She uses audio-visual recordings of real life interactions and an approach known as conversation analysis to capture and understand how we attempt and accomplish things with one another through our interpersonal interactions. She has largely worked on recordings of healthcare interactions. She has interests in difficult communication tasks such as telling someone else what is wrong with their motor performance, and talking about issues such as illness progression and dying. Ruth's other key area of interest is in using analysis of recordings to generate insights into what somewhat nebulous concepts (dignity, patient-centred care) look like in practice. In recent years, with her team, Ruth developed, disseminated, and evaluated a set of communication training resources called 'RealTalk' which incorporate both clips from real life recordings and learning points based upon insights and findings of conversation analytic research. The resources are designed for use by communication trainers within their work in the NHS, universities, and hospices, and aim to increase the evidence-base and authenticity of health and social care communication training. Ruth has also pioneered the adaptation and application of systematic review methods for conversation analytic studies.

	Richard Simmons is a Professor of Public and Social Policy, and Co-Director of the Mutuality Research Programme at the University of Stirling. Over the last decade he has led an extensive programme of research on the use of voice in public services. This includes four studies funded by the Economic and Social Research Council, a Single Regeneration Budget-funded study, and work for the NHS, Scottish Executive, National Consumer Council, Carnegie Trust, World Bank, Co-operatives UK, NESTA and the Care Inspectorate. He also writes widely on these issues for academic, policy and practitioner audiences. His book, 'The Consumer in Public Services' is published by the Policy Press. As well as a series of journal articles in high-quality international journals such as Social Policy and Administration, Policy and Politics, Annals of Public and Co-operative Economics, and Public Policy and Administration, Richard has written a number of policy-oriented publications and professional journal articles for a practitioner audience. His research interests are broadly in the field of user voice, the governance and delivery of public services and the role of mutuality and co-operation in public policy. The Mutuality Research Programme has acquired an international reputation as a centre of excellence for research, knowledge exchange and consultancy on these issues. **Adrian Kerrison** is a Postdoctoral Researcher/Postdoktor at Linköping University with the Non-Lexical Vocalizations project. His work uses Ethnomethodology and Conversation Analysis to examine how crowds operate as social actors within large-scale settings such as sporting events, artistic performances, and protests. Currently he is focused on the use of individual non-lexicals (yelps, grunts, etc.) to perform attention, understanding, and assessment of play in sporting contexts.
Any funding acknowledgement details:	

1   **Should we share qualitative data? Epistemological and practical insights from**

2   **Conversation Analysis**

3

4   **Abstract**

5   Over the last thirty years, there has been substantial debate about the practical, ethical and

6   epistemological issues uniquely associated with qualitative data sharing. In this paper we

7   contribute to these debates by examining established data sharing practices in Conversation

8   Analysis (CA). CA is an approach to the analysis of social interaction that relies on audio/video

9   recordings of naturally occurring human interactions and moreover works at a level of detail that

10   presents challenges for assumptions about participant anonymity. Nonetheless, data sharing

11   occupies a central position in both the methodology and the wider academic culture of CA as a

12   discipline and a community (ten Have, 2007). Despite this, CA has largely been ignored in

13   qualitative data sharing debates and discussions. We argue that the methodological traditions of

14   CA present a strong case for the value of qualitative data sharing (QDS) and offer open data

15   sharing practices that might be usefully adopted in other qualitative approaches.

16

19

20   Data sharing is an important principle of the Open Science agenda but is not yet

21   widespread within qualitative research. Debates emphasise ethical and epistemological

22   issues uniquely associated with qualitative data that are seen as a barrier to data sharing.

23   However, noticeably absent from these debates is the contribution of Conversation

24   Analysis (CA), a qualitative research approach with a long history of data sharing. We

25   address that gap through examining the formal and informal data sharing practices of CA

26   and the underlying epistemological position that supports those practices. We argue that

27   the particular conception of data and context adopted in CA makes redundant the very

28   distinction between primary and secondary data that currently frames data sharing debates

29   and has thus significantly influenced data sharing practices amongst CA researchers. Two

30   broad distinctions can be made with regards to data sharing: (1) *data sharing as a*

31   *research practice*, contributing to the rigour of analysis within a particular project; (2)

32   *corpus sharing*, where a complete data set is made available via a repository for other

33   research projects. Both types of sharing are well established in the CA community with

34   demonstrable benefits for research outcomes. Thus, we argue that the methodological

35   traditions of CA present a strong case for the value of qualitative data sharing (QDS) and

36   offer open data sharing practices that might be usefully adopted in other qualitative

37   approaches. This is particularly important in a context where funders, publishers and legal

38   requirements increasingly expect data to be made available for sharing.

39   **Qualitative Data Sharing: Nature and Context**

40   Interest in potential reuse/secondary analysis of qualitative data has grown since the

41   1990s (Heaton, 2008; Hughes et al. 2020). Arguments for sharing and reusing qualitative

42   data include checking of findings, fostering public trust in science, and enhancing

43   research training (DuBois et al. 2018). In addition, existing data can be analysed to

44   produce new findings, which is time and cost-effective for researchers and avoids

45   unnecessary burden on participants (Kuula, 2011). The first qualitative data repository,

46   Qualidata, was established over 25 years ago (and is now part of the UK Data Service[1]).

47   Since then, technological advances have increased capacity to store and facilitate access

---

[1] A resource that provides guidance on data management and includes a large archive of data
      and the details of other collections.

48   to large datasets in repositories (Chauvette et al., 2019, p. 2; Corti et al., 2016)[2].

49   Increasingly, research funders and publishers encourage and even mandate QDS (Antonio

50   et al., 2019; Chauvette et al., 2019) through policies of open access, shaped by a

51   commitment to principles of transparency and scrutiny, and to maximising the social

52   value of publicly-funded research (UKRI, 2021). In the UK, important milestones for

53   QDS were the adoption by all research councils in 2011 of the Common Principles on

54   Data Policy and in 2012 of the Policy on Access to Research Outputs which required

55   research outputs to make explicit how the research data would be made available (Bishop

56   & Kuula-Luumi 2017). With 55% of research in UK HE funded by research councils,

57   these policies "strongly influence research practices" (Bishop & Kuula- Luumi, 2017, p.

58   2). The UK Research and Innovation (UKRI) conducted a review, in the UK (published

59   February 2022), of open access policies and the data sharing landscape which reiterated

60   their concordat on Open Research Data (UKRI, 2021) emphasising that publicly funded

61   research should be openly available with as few as restrictions as possible.

62   Nonetheless, concern about QDS persists. Resistance is often expressed on ethical

63   and epistemological grounds (Chauvette et al., 2019; Mozersky et al., 2020a). Two large-

64   scale surveys of scientific researchers identified a number of recurring concerns: fears

65   relating to participant anonymity, misinterpretation in secondary analyses, invalid

66   conclusions, data errors, being scooped, and researcher burden. Overarching and unifying

67   guidelines, policies, and mandates are also still lacking internationally. Moreover, where

68   they exist, data sharing approaches, policies, and repositories are mostly established with

69   quantitative research in mind (Antonio et al., 2019; Tsai et al., 2016). For example, pre-

---

[2] Data repositories were initially developed to increase the transparency and sharing of data
    from clinical trials (Antonio et al., 2019).

registration forms required by some funders are often inadequate adaptations of forms designed for quantitative work (Humă & Joyce, frth). Recent empirical research on QDS preparedness in the US found that even repository specialists lacked experience and knowledge relevant to QDS (Mozersky et al., 2020a). Many felt unprepared to advise qualitative researchers – particularly on decisions about sensitive data. Similar limitations in knowledge and preparedness were found amongst qualitative researchers and institutional ethics committee members. This US study found little experiential knowledge of QDS; each group (researchers, ethics committees and repository staff) felt that the primary responsibilities and decisions about QDS lay elsewhere. Even if those groups were to become more experienced, there remains a lack of agreement and guidance on best practice, exacerbated by different requirements between institutions and repositories, and different countries' laws relating to cross-border data sharing. Arguably, resolving these infrastructural and practical concerns depends on also addressing the debates about QDS to which we now turn.

Debating Qualitative Data Sharing

Literature reflecting on the viability of QDS is dominated by discussion of ethical and epistemological challenges posed by the various forms of qualitative data, each with different affordances, including (inter alia) observational data (e.g. fieldnotes, audio/video recordings), participant produced data (e.g. diaries) and researcher elicited data (e.g. interviews). Across this section we review literature outlining these challenges and the impact they have on qualitative researchers' commitment to data sharing and data reuse. We discuss ethical and epistemological challenges that exist across the range of qualitative methods as well as relating to specific methods.

Ethical Debates

95   Some qualitative researchers argue that it is impossible to ensure that research participants

96   know what they are consenting to when it comes to data sharing – how data might be used

97   in future projects and by other researchers (Parry & Mauthner, 2004; see also Chauvette

98   et al., 2019). Consent for data sharing can only ever be secured in a general manner for

99   and about the process itself, rather than for specific research (Irwin, 2013, p. 297). The

100   issue, then, is whether it can be considered ethical to share data when *fully informed*

101   consent for the myriad ways data might be used can never be achieved.

102   The counterargument, however, is that it is impossible to be fully 'informed' about

103   all aspects of research (even research questions, for example, may not be formed prior to

104   data collection in some qualitative approaches) (Bishop, 2009). Hence, this is not a reason

105   to dismiss QDS. Additionally, qualitative research participants, who invest time but also

106   emotionally in the research, generally seem to support data sharing and to assume it

107   occurs to a greater extent than it does (Kuula, 2011). Indeed, participants in sensitive

108   qualitative studies, interviewed by Mozersky and colleagues (2020b), reported broad

109   support for QDS where data are anonymised, although, when pressed, expressed concern

110   about confidentiality and potential misuse/misunderstanding in future research. This

111   research also suggests that qualitative research participants trusted research institutions

112   and their researchers to be sufficiently transparent with data collection and sharing plans.

113   These findings echo similar work on the trust and value of research, such as Parry et al.

114   (2016) who surveyed research participants and reported that most regard qualitative

115   video-based research as acceptable, and Williams et al. (2010) who reported an

116   overwhelming majority of participants believing that recording was worthwhile.

117   There are, however, other potential ethical issues. Qualitative data can be highly

118   sensitive, confessional and intimate. Ensuring confidentiality and anonymity, and

119 protecting participants from unintended identification is vital, but rigorous anonymisation

120 is time and labour intensive and challenging. In some forms of qualitative research, even

121 if all possible measures of anonymisation are taken, there is still the possibility for

122 identification as total anonymisation is impossible (Hopkins, 1993). For example, in

123 longitudinal data collection or in other forms of data collection that link different sets of

124 data together, the accumulation of information and associations potentially present a

125 greater disclosive risk (Law, 2005). Similarly, qualitative research that takes place in

126 small communities or on phenomena that are rare may be particularly hard to fully

127 anonymise (Chauvette et al., 2019; Hardy et al., 2016).

128 This poses a challenge to QDS on two-fronts: the integrity and quality of certain

129 forms of data is compromised when it is digitally altered to ensure participant

130 confidentiality (such as video data) (Bishop, 2009: 262), and the reuse of original

131 (unedited) data runs the risk that researchers outside of the original project may not know

132 what should be anonymised and how it should be anonymised. These are challenges

133 whose answers may have profound consequences for the scientific analysis (Corti et al.,

134 2000).

135 It is true that there are no straightforward or 'one size fits all' answers to these

136 issues (see Humă & Joyce, frth). However, it has also been argued[3] that "too often the

137 critics of reusing qualitative data have narrowly construed the debate to focus solely on

138 participants – to the exclusion of other agents – and rights – to exclusions of duties. Such

139 arguments do not do justice to the depth of moral debate required" (Bishop, 2009, p. 258).

140 In this sense, advocates of QDS argue for a broader range of ethical considerations. From

---

[3] This is influenced by deontological ethics. See Bishop (2009, p. 257-260) for an overview.

141 this perspective, the benefits to knowledge, policy and society from data sharing (weighed

142 against, in the majority of cases, minimal risks to participants) need to be given greater

143 emphasis (Bishop, 2009). For example, data sharing can mean avoiding the unnecessary

144 intrusion and burden on participants that result from collecting data that already exist.

145 ***Epistemological Debates: The Problem of Context***

146
147 The central epistemological debate in QDS literature concerns the contextual nature of

148 qualitative data. One sense of this contextuality relates to the relationship between the

149 researcher and the overall research process. In this sense, qualitative data is argued to be

150 contextual in that it reflects the original researcher's positionality, beliefs, judgments,

151 disciplinary assumptions and boundaries, as well as their theoretical and methodological

152 inclinations and intentions within those disciplinary boundaries (Irwin, 2013). These

153 aspects are embedded in the data, uniquely shaping its constitution and analysis.

154 Reflexivity is thus a central practice of the qualitative paradigm, requiring the researcher

155 to explicitly examine how these underpinning beliefs and practices have shaped the data

156 and its analysis. However, for many qualitative approaches this practice is not available

157 in secondary data analysis (Mauthner et al., 1998). As a consequence, many qualitative

158 researchers maintain that only they (and their team) can analyse their data in a

159 contextually fitted and adequately reflexive manner.

160   A second aspect of the contextuality of qualitative data focuses more narrowly on

161 different conceptualisations of context, particularly with respect to how it relates to

162 talk/text. van den Berg (2008, p. 186-188) describes three ways of conceptualising

163 context: as (broad) *extra-discursive template,* where the relation between text and context

164 is predefined (e.g. Critical Discourse Analysis); as (narrow) *intra-discursive product*,

165 where context is only relevant when demonstrably made relevant in a participant's talk

166 (e.g. Conversation Analysis); and as (intermediate) *conditions of discursive production,*

167   where the necessary contextual information depends on the focus of the research and the

168   data being used (e.g. ethnography of communication). The broad and intermediate

169   conceptions of context view qualitative data as generated in a specific time and setting of

170   which the primary researcher necessarily has first-hand, intimate experience. In this

171   sense, ethnographic fieldnotes, for example, may be difficult or impossible to

172   meaningfully interpret by a researcher who did not participate in the original research

173   (Chuavette et al., 2019). As Hammersley (2010, p. 3) notes, "in the process of data

174   collection researchers generate not only what are written down as data but also implicit

175   understandings and memories of what they have seen, heard, and felt, during the data

176   collection process.". According to this perspective, the extent of contextual understanding

177   in secondary analysis will necessarily be more limited and interpretive (Hammersley,

178   2010). This understanding of context, however, is not in harmony with CA's perspective

179   which we will consider later.

180   Other contributors to context as an intra-discursive product posit that data are

181   constructed and not independent of the research process – that, in other words, 'context'

182   is not ontologically separate from data (Mauthner & Parry, 2009). Ethnomethodologists

183   (Garfinkel et al., 1981; Lynch, 1982) have questioned how "data" is even first granted

184   that status by researchers, and how a discipline's technical language and concepts must

185   be deployed to alert others to the presence of data (Maynard & Clayman, 1991). From

186   this perspective, researchers do not 're-use' data, because data are constituted for the first

187   time in a particular research project (Moore, 2007; see also Bishop, 2007).

188   The past two sub-sections have explored the existing literature concerned with

189   data sharing in qualitative research. This has shown gaps in researchers' agreement on

190   the value of – and commitment to the practice of – QDS; significant interpretive and

191   practical difficulties associated with this; and a series of ethical and philosophical

192   questions regarding the sharing and reuse of data. The paper now turns to understandings

193   and approaches to data sharing in CA and considers their potential to address these gaps,

194   difficulties and questions.

195   **Centrality of Data Reuse in Conversation Analysis**

196

197   Before detailing CA's contributions to data sharing debates, it is necessary to provide a

198   brief overview of CA's history and its arguably unique relationship with data sharing[4].

199   CA draws focus on practices and social actions rather than people or experiences which

200   means it is commonplace to ask very different questions of reused data. In the early days

201   of CA, data collections tended to be limited to audio only. Initially, Harvey Sacks, while

202   a researcher at The Suicide Prevention Centre, analysed recorded phone calls for his PhD

203   thesis in 1966[5].

204   From its inception, data has been central to CA's concerns – with much of the

205   early development of CA by Sacks drawing on two collections of audio recordings: calls

206   to the suicide hotline, and group therapy sessions (Sacks, 1992). Over the following

207   decades, a number of phone call corpora were created, notably recordings taken from

208   around Santa Barbara in California which the CA community refer to as the Newport

209   beach corpus or more commonly: "Classic data", and Elizabeth Holt's "Holt corpus" of

210   phone calls recorded by a British family over three years. These are data which are widely

211   available and widely reused and on which much of the ground-breaking work in CA is

---

[4] The cumulative relationship between CA and QDS is unique but the specific practices and procedures are not unique to the approach.

[5] For a fuller picture of the founding of CA see: Psathas (1994) ten Have (2007) Sidnell (2011) Silverman (1998).

212 based[6]. The (re)use of this data speaks to a core ideal in CA – that data used in research

213 should be made available to check findings (often in the form of transcripts, but also with

214 visual representations (see Walker, 2017) or in the sharing of audio/video data). Sacks

215 illustrates this point:

216   "It was not from any large interest in language or from some theoretical

217   formulation of what should be studied that I started with tape-recorded

218   conversations, but simply because I could get my hands on it and I could study it

219   again and again, and *also, consequentially, because others could look at what I*

220   *had studied and make of it what they could, if, for example, they wanted to be*

221   *able to disagree with me"*

222   (Sacks, 1984, p. 26, emphasis added).

223 The technology of the time influenced the practices of the discipline: recordings could be

224 replayed, scrutinised by others and made available for future studies by other researchers.

225 Sacks goes on to explain how he chooses the data he works with:

226   "People often ask me why I choose the particular data I choose. [..] And I am

227   insistent that I just happened to have it, it became fascinating, and I spent some

228   time at it" (Sacks, 1984, p. 27).

229 Data which CA researchers *just happen to have* has been used and reused in a number of

230 studies addressing a wide range of interactional phenomena. CA with its grounding in

231 ethnomethodology examines the observable practical common-sense reasoning as

232 revealed in the data itself to make sense of how the social world is constituted in local

233 environments. This means the distinction between primary and secondary analyses

---

[6] Most modern CA research no longer draws on classic data with that collection being normally
     reserved for teaching.

234 disappears because the source of evidence is always constituted for the first time. This

235 fits with Moore's (2007) understanding that in data reuse, analysis is always primary but

236 of a *different order of data*, Hughes et al. (2020) extend this position by articulating "the

237 range of approaches and practices involved in producing different orders of data" (p. 568).

238 One example is Gibson's (2019) primary (rhetorical) analysis of Milgram's classic

239 obedience experiment data which offers a reinterpretation of the core insights around

240 obedience and persuasion. Similarly, Hughes et al. (2020) show how interview data might

241 be reused to examine features of relational dynamics. In this way QDS can open up novel

242 avenues of research and lead to further scrutiny of prior findings[7].

243   Opening up novel avenues of research is one benefit of QDS, other benefits are

244 explored by Jepsen and colleagues who reflect on reasons for creating their primary care

245 consultations archive: a corpus of recordings of GP consultations, linked survey

246 responses and patient records named the 'One in a Million' corpus:

247   "Data sharing provides considerable added value in terms of minimising data

248   collection costs, reduced environmental impact, and patient and practice burden.

249   This will support low-cost studies including doctoral-level research, thus building

250   research capacity in primary care." (Jepson et al., 2017, p. 350).

251 Their argument expands Sacks' point – that a chief reason for sharing data is so that other

252 researchers (particularly those who are at an early stage of their careers) can just *happen*

253 *to have* that data. This is a point the paper will return to later, but first it is necessary to

254 describe the kinds of data that CA usually works with and the essential characteristics of

255 that data.

---

[7] See Humă & Joyce (frth) for a discussion on the relationship between the culture of data

sharing and the culture of continuous refinement and replication in CA.

256   CA draws on recordings of naturally occuring[8] social interaction. This can include

257   any site where interaction occurs between participants, including (but not limited to):

258   Online chat logs (e.g. Meredith & Stokoe, 2014), Institutional encounters (e.g. Drew &

259   Heritage, 1992), Phone calls (Holt, 1996), AI (e.g. Suchman, 2007; Mair et al., 2020),

260   Video recordings (e.g Mondada, 2018) and so on. In this vein, research interviews can be

261   viewed as a site of interaction (Potter & Hepburn, 2012). CA research can be broadly

262   placed within one of two camps: 'pure CA' and 'Applied CA'. Antaki (2011) explains

263   that 'pure CA' focuses on interactional practices and procedures detached from *any* type

264   of context (e.g. Jefferson, 1988) and takes an endogenous orientation to the conversation

265   itself rather than drawing on analytic insights about the institutional context. Compared

266   to 'applied CA' which focuses on interactional practices *within* a certain setting (e.g.

267   Drew & Heritage, 1992; ten Have, 2007) and provides an evidence-base for interventions

268   (e.g. Stokoe, 2014; Wilkinson, 2015). A discussion of the debate regarding the two terms

269   can be found in Antaki (2011).

270   The procedures for both camps are largely the same – recordings of social

271   interaction are gathered and analysis proceeds with 'unmotivated looking'; that is, as

272   Psathas (1990) explains, the researcher discovering what is happening in the recordings

273   and not searching for predetermined phenomenon. Data collection of interaction

274   recordings do not normatively involve the researcher which thus enhances the usefulness

275   of the data for reuse and reanalysis. To outsiders the unmotivated looking and efforts to

---

[8] "Naturally occuring" is a slogan in the CA enterprise and usually contrasts with researcher

    elicited data or scripted talk, but see the debate in Discourse Studies which problematises the

    'natural' and 'non-natural' data distinction (Speer, 2002a; 2002b; ten Have, 2002; Lynch,

    2002; Potter, 2002).

276 remain exogenous to the data collection process may seem unstructured and haphazard,

277 but the methodological technology imposes a high degree of rigour to account for and

278 evidence unmotivated discoveries (see Liddicoat, 2007, p. 9 and Schegloff, 1996a, p. 172-

279 173 on accounting for phenomena). In short, data analysed in this way aims to avoid

280 mediation by the subjective perspective of the researcher.

281 Despite the growing range of data that CA researchers draw upon, core data

282 sharing principles remain unchanged since its foundation – that others ought to have

283 access to the data, including, ideally, the original video/audio recordings, so to scrutinise

284 the analysis of the researcher and that data is usually made available for pre-publication

285 data sharing sessions (referred to as "data sessions") as an integral part of the method.

286 Recordings which are particularly sensitive may be subject to greater sharing restrictions

287 which may be mitigated by heavily anonymising the recordings (e.g. voice altering, video

288 manipulation etc.), or by asking data session participants to sign a non-disclosure

289 agreement and return all materials after a meeting. These more extreme measures are

290 often the result of ethics committee requirements, and not of the science itself, with many

291 sensitive anonymised data sets shared without such restrictions in place.

292 To summarise, over the course of CA's history the core principle that data should

293 be shared and reused has established formal and informal practices for handling data. The

294 remainder of the paper describes CA's understanding of and approach to ethics and

295 epistemology, and explores CA's established procedures and practices for sharing data

296 with the intent of widening ongoing debates and allaying some of the persistent concerns

297 in qualitative research about data sharing.

298 **Ethics, Context and Conversation Analysis**

299

300 *Ethics*

301 Conversation Analysts deal with many of the same ethical dilemmas experienced by

302 other forms of qualitative research. For CA studies, which collect recordings of social

303 interaction, ensuring anonymity for participants in shared data can be technically

304 complex. It minimally requires the deletion of names, dates and locations. However,

305 other components that make participants identifiable require more complex

306 anonymisation decisions, for instance, anonymising voices and faces, or whether to

307 remove specific details such as references to a participant's medical condition.

308 Moreover, a participant may, in the course of a recording, indicate (in the recording)

309 that some part should be anonymised through either explicit mention (e.g. Speer &

310 Hutchby, 2003), or by blocking the recording equipment (e.g. Mondada, 2014). There is

311 a profound understanding in CA that simply removing names, dates and locations might

312 not always be sufficient for anonymisation.

313   It is not possible to predict every ethical dilemma which may arise in the course

314 of research hence ethical solutions cannot be prescribed *a priori*. These ethical

315 questions will persist as (hopefully all) researchers endeavour to protect their

316 participants from harm. However, when the possibility of data sharing is built into

317 research procedures ethical safeguards become even more central to the research design

318 (see Albert & Hofstetter, frth for a discussion). This includes providing information to

319 participants about data reuse (and its associated risks) along with consent forms which

320 allow participants to decide whether, and in what forms/contexts, their data may be

321 shared for future research[9]. This has, for example, been the approach taken in the

---

[9] Participants are rarely in a position to fully understand the research process and a discussion of
this and how ethics panels are not geared to handle qualitative data sharing warrants a future
paper (but see Hammersly, 2014; ten Have, 2007: 79-81).

322   National Institute for Health Research (NIHR) funded CA project known as 'Real

323   Complaints' (Real Complaints, 2021).

324

325   *Context*

326   CA can uniquely contribute to debates about the problem of context in QDS.

327   Precisely what is meant by 'context' is arguably "fuzzy" (van Dijk, 2007, p. 285) across

328   the social sciences, as it can be a shorthand to denote a specific situation, or the

329   historical/geographical/cultural environment, of the object of investigation. However,

330   CA's particular way of dealing with 'context' does not entertain contextual explanations

331   of phenomena. Handling context in CA has been debated at length[10] (see ten Have, 2007,

332   p. 58-59 and Wooffitt, 2005, p. 168-179 for reviews). In short, CA does not assume that

333   aspects of context such as social categories (race, gender, power, class etc.) are relevant

334   *a priori*. Rather, context is dealt with analytically if, and only if, it is *procedurally*

335   *relevant* and *demonstrably attended* to by the interlocutors themselves[11] (see Schegloff,

336   1992). Hence Irwin's (2013) concerns about the contextual qualities of qualitative data

337   are not normally relevant in CA.

338   This should not be read as necessarily advocating for this way of handling context

339   in qualitative research generally. The point being made is that when sharing data, it cannot

340   be foreseen how it may be (re)used. The endogenous understanding of context espoused

---

[10] This was discussed and responded to at length between Emmanuel Schegloff (1997; 1998;

    1999b; 1999c), and Margaret Wetherell (1998) and Michael Billig (1999a; 1999b) who took

    issue with Schegloff's original 1997 paper.

[11] It is this additional step that distinguishes CA from other inductive approaches such as

    ethnography and grounded theory.

341 by CA means that it does not carry any 'burden' of externally imposed context to delimit

342 what it can be used to demonstrate. It cannot be expected that the data which researchers

343 share will only be used by those within the same discipline or even those who share

344 similar interests – rather, researchers ought to anticipate that the shared data may be used

345 beyond the scope of the original research, (see the previous discussion on the practices of

346 producing different orders of data (Hughes et al. 2020)).

347   Returning to discussion of the extent to which data may be meaningfully

348 interpreted by researchers outside of the original research, new and fruitful avenues of

349 investigation can be found in the reuse of data collected for alternate purposes (e.g.

350 Gibson, 2019; Hughes et al. 2020). For certain qualitative approaches the arguments by

351 Irwin (2013) and Chuavette et al. (2019) are salient, but different approaches with

352 different epistemologies may make use of data in ways unforeseen by the original

353 researchers. Data collected for one particular purpose may be meaningfully reinterpreted

354 with CA because of its focus on phenomena demonstrably enacted and treated as relevant

355 by participants in the discourse.

356 **Data Sharing in Conversation Analysis: Practical Aspects**
357

358 As a fundamentally collaboratively discipline CA has fostered a culture[12] and tradition of

359 data sharing out of which has emerged a community of practice:

360   "CA is a *community*, although with various degrees of intensity. As it has become

361   established as a quite solidly and specifically defined approach in the human

362   sciences, you can, by working in the CA tradition, become 'a member' of that

363   community" (ten Have, 2007, p. 11)

---

[12] We refer to "culture" in the sense of disciplinary culture rather than epistemology.

364 Typically, research which makes its data available does so following completion of the

365 project – whether defined as the publication of an (final) article, or the overall

366 conclusion of the funding period. Data collections may be described on websites such as

367 the Open Science Framework during the research process but are not commonly

368 available until after project completion. We refer to this widespread form of QDS as

369 'corpus sharing' distinguished from the practices and solutions employed by the CA

370 community to add levels of transparency and rigour to the analysis, which we refer to as

371 'data sharing as a research practice'. This section returns to the fears outlined previously

372 and discusses the practical aspects of the CA approach to data sharing both as corpus

373 sharing and as established research practice.

374 ***Data Sharing as a research practice***
375
376 CA is a community of practice with a particularly democratic impulse – that both the

377 analysis and the research process builds from the ground up with students, practitioners,

378 and experienced CA researchers able to contribute insights through data sharing sessions.

379 Data sharing is thus "baked into" the research process. During the research process, there

380 are options for qualitative researchers to share data and findings in progress at

381 conferences, seminars and research meetings but CA is distinctive in that focused data

382 sharing meetings (referred to as 'data sessions') are an integral part of the scientific

383 process and disciplinary culture – researchers regularly share their data at data sessions.

384 Data sessions are structured research meetings where direct access to the

385 recordings is made available to other researchers to scrutinise[13]. Although data is often

---

[13] Examples of groups include the Conversation Analysis Reading and Data Sessions (CARDS)

at Ulster University, and the long-standing Discourse and Rhetoric Group (DARG) at

Loughborough University. A list of groups is maintained here: https://rolsi.net/data-sessions/

386   analysed by multiple researchers during the research process across the gamut of

387   qualitative (and quantitative) research approaches, data sessions are distinctive in the

388   sense that data is subject to scrutiny and analysis by others outside of immediate research

389   teams and institution. Findings can be independently checked, and ideas collaboratively

390   explored (ten Have, 2007).

391   The procedures of a data sharing session can vary amongst research groups, but

392   usually the data presenter shares recordings and transcripts with the group. The transcripts

393   will contain more detail than is perhaps necessary for the data owner's interest "because

394   even if, say, pauses or overlaps are not germane to the current analysis, some other

395   researcher might want to use the same materials for checking findings or for novel

396   analytic purposes" (Jordan & Henderson, 1995, p. 48). Participants see/hear the recording

397   several times as transcripts are recognised to be a static and partial representation of

398   interaction. After a moment or two of quiet 'thinking' time, the group will propose

399   observations. There is usually some rule that group members are initially limited to a

400   single observation to encourage a collaborative, democratic ethos with equal access to

401   contribute.

402   Crucially, sharing data allows for a more transparent analytic process for the data

403   owner and for learners of the method. The practices and procedures of the data session

404   might be usefully adopted by other qualitative approaches to fit with the Open Science

405   movement (see Humă & Joyce, frth) by not only corroborating findings, but also making

406   explicit the discussions and analysis of data often done behind closed doors. We argue

407   that collaborative analysis of data adds another level of rigour in the analytic process

408   where banal observations may be retold as composed and refined analytic points and

409   flawed analysis, or invalid conclusions recognised and corrected. The value of the data

410   session cannot be overstated and highlights possibilities for data sharing beyond making

411 data sets available in a repository post-project. Worries about being scooped by sharing

412 one's data prior to publication are greatly outweighed by the benefits of the data session,

413 and indeed, new projects may be launched, and collaborations proposed following such

414 sessions. Sharing is thus baked into the research process through the tradition of the data

415 session through which both new and seasoned researchers and, where relevant,

416 participants involved in the data, are invited to witness and scrutinise data on their own

417 terms.

418 ***Corpus sharing***

419 'Data sharing' typically refers to post-project data sharing in a repository – where a final

420 corpus of data is made available to other researchers. This paper treats a corpus of data

421 as data (in whatever form) associated with a single project, whereas a repository of data

422 is a resource where multiple corpora are stored and made accessible to others. The gold

423 standard of data sharing is typically regarded as unrestricted access to data, which is

424 shared with accessibility in mind, and is fully described and indexed so that other

425 researchers can easily search and understand the data (e.g. Mass Observation Archive,

426 British Library Sound Archive), and importantly, check analysis. CA is not unlike other

427 qualitative disciplines in that corpora are held in various places (some, such as TalkBank,

428 are specific to communication data), and although the ideal is unrestricted sharing, in

429 practice there may be gatekeepers or restrictions on accessing data. Moreover, data is

430 often shared through informal networks rather than through a formal repository.

431 It is impossible to predict how data may be reused which carries with it benefits

432 and drawbacks. Within CA the units of analysis are discursively realised practices and

433 social actions, rather than people or experiences. The approach allows for the study of

434 phenomena whose context is endogenously constituted within the talk itself. In this way,

435 subsequent researchers can ask very different questions about the data, focusing on what

436   is *being done* in the interaction and not the original purpose of the data collection (which

437   for CA researchers is a context that is not relevant to the analysis). For example, data

438   from Heritage and colleagues' 2007 study investigating how doctors encourage patients

439   to voice concerns in consultations was used by Heritage (2012) in a study focusing on the

440   relationship between epistemic status and stance. Similarly, interview data from a study

441   by Hepburn and Brown (2001) asking how secondary school teachers use 'stress' to

442   manage their accountability and make sense of their institutional role was used by Potter

443   and Hepburn (2005) to critique the (over)use of interviewing in qualitative psychology.

444   This illustrates that *all* data irrespective of method of collection might be repurposed to

445   generate novel findings potentially in novel ways.

446   The generation of findings is, however, only one argument for post-project data

447   sharing. Conversation Analysts have long advocated for direct access to the data[14]

448   presented in empirical articles. Providing corroborative evidence, which is prepared

449   effectively (see Walker, 2017), for research claims allows others to independently check

450   those claims. To repeat Sacks's observation, "others could look at what I had studied and

451   make of it what they could" (Sacks, 1984, p. 26). Much of the foundational CA work

452   which reused recordings (e.g., from the Newport Beach corpus or the Holt corpus) had

453   such an impact in the community because fellow researchers were familiar with the data

454   and able to independently check findings.

455   **Conclusion**
456
457   This paper contributes to discussions around the viability and usefulness of QDS, adding

458   insights from the established traditions of CA to widen those discussions, and to advocate

---

[14] 'Direct access' may be confused with access to the in-the-moment encounter, but here we

refer to the original recording of the encounter.

459   for a more open and flexible approach to QDS. Ongoing debates emphasise the ethical

460   and epistemological barriers to QDS and are framed by a distinction between primary and

461   secondary data. We argue that CA's conception of data and context makes this distinction

462   redundant. Currently, the demands of funders, Open Science and legal restrictions

463   influence decisions about what gets shared and how, but inexperience and lack of

464   consensus on best practice for QDS persist. Our aim has been to reflect on the long history

465   of sharing data in CA, the impetus for sharing within the CA community, and how these

466   procedures might be drawn on by other qualitative approaches.

467   We discuss two types of data sharing that are baked into the design of CA studies:

468   sharing as research practice, and corpus sharing. We show how the 'data session', while

469   not unique to CA (see also, grounded theory), enables the research process to build from

470   the ground up and that this collaborative analysis adds rigour to the analytic process.

471   Corpus sharing is a more traditional understanding of data sharing – and while matters of

472   context and ethics present as barriers to reusing data, for conversation analysts, having

473   access to the original recordings of analysed encounters is considered gold-standard. The

474   expectations, tools and procedures of CA facilitate more transparent QDS within the

475   community but as with most other approaches they rely on the researcher(s) having

476   sufficient means to engage in QDS.

477   Beyond the practical barriers to engaging in sharing as a research practice or

478   establishing and sharing a data corpus, many authors point to significant ethical and

479   epistemological barriers for QDS. For ethics, data reuse presents challenges for informed

480   consent and the high level of anonymisation potentially required might make the data

481   difficult to work with. While we accept that attempting to solve all ethical issues relating

482   to participants consent or prescribing ethical solutions *a priori* is a fool's errand, building

483   in the possibility of data sharing into the research design, as CA does, foregrounds ethical

484 safeguards which can alleviate potential dilemmas. For epistemology, many qualitative

485 approaches consider reflexivity a central practice making secondary analyses impossible

486 and indeed, different qualitative approaches conceive of 'context' very differently which

487 again, makes the data difficult to work with. CA, as an illustration, does not face these

488 concerns. The distinct way that CA conceives of 'context' dissolves the distinction

489 between 'primary' and 'secondary' data meaning that data is always constituted for the

490 first time. Unlike a number of qualitative approaches, CA is not 'burdened' by externally

491 imposed context to delimit what it can be used to demonstrate. To be clear, we are not

492 advocating that all qualitative approaches follow CA's conception of context but instead

493 are arguing that future use of any data can never be predicted and that all data, irrespective

494 of method of collection, might be repurposed to generate novel findings in potentially

495 novel ways.

496   We have demonstrated that sharing insights from CA can allay fears and barriers

497 to QDS and that the long-established and refined tools of CA make QDS much more

498 achievable. CA is, however, not a panacea for all QDS challenges and for many

499 qualitative researchers, particularly early-career researchers, or those in marginalised

500 areas, barriers to QDS – whether ethical, epistemological, or economic – can prove

501 difficult to overcome without sufficient support and funding and thus they may be

502 reluctant to share their data (Pownall et al., 2021) which can adversely impact career

503 outcomes (Siegel & LaMarre, 2019). This is a crucially important topic which we have

504 not discussed at length and so encourage further scholarship on this issue.

505   We conclude by reiterating Sacks' (1984, p. 26) explanation of how he came to

506 study the data that he did: "simply because I could get my hands on it and I could study

507 it again and again, and also, consequentially, because others could look at what I had

508 studied and make of it what they could". CA was built on the ideal that data *should* be

509   shared for the benefit of the primary investigator and the research community. The

510   overall intention of this paper was to spark further discussion of what QDS could look

511   like across the range of qualitative approaches.

512   **References**
513
514   Albert, S. & Hofstetter, E. (frth). Data Management 1: Privacy, Security and Access.
515
516   Antaki C. (2011). Six Kinds of Applied Conversation Analysis. In C. Antaki (Ed.),
517   *Applied Conversation Analysis* (pp. 1-14). Palgrave Advances in Linguistics.
518   Palgrave Macmillan, London.
519
520   Antonio, M. G., Schick-Makaroff, K. Doron, L.S., White, L. & Molzahn, A. (2019).
521   Qualitative Data Management and Analysis within a Data Repository. *Western*
522   *Journal of Nursing Research,* 42(8), 640-648.
523   https://doi.org/10.1177/0193945919881706
524
525   Berg, H. v. d. (2008). Reanalyzing qualitative interviews from different angles: the risk
526   of decontextualization and other problems of sharing qualitative data. *Historical*
527   *Social Research,* 33(3), 179-192. https://doi.org/10.17169/fqs-6.1.499
528
529   Billig, M. (1999a). Whose terms? Whose ordinariness? Rhetoric and ideology in
530   conversation analysis. *Discourse & Society,* 10, 543–58.
531   https://www.jstor.org/stable/42888872
532
533   Billig, M. (1999b). Conversation analysis and the claims of naivety. *Discourse &*
534   *Society,* 10, 572–6. https://doi.org/10.1177/0957926599010004007

535   Bishop, L. (2007). A reflexive account of reusing qualitative data: Beyond
536   primary/secondary dualism. *Sociological Research Online, 12*(3), 43-56.
537   https://doi.org/10.5153/sro.1553

538   Bishop, L. (2009). Ethical sharing and reuse of qualitative data. *Australian Journal of*
539   *Social Issues,* 44(3), 255-272. https://doi.org/10.1002/j.1839
540   -4655.2009.tb00145.x

541   Bishop, L. & Kuula-Luumi, A. (2017). Revisiting Qualitative Data Reuse: A Decade
542   On. *SAGE Open*, 1-15. https://doi.org/10.1177/2158244016685136

543   Chauvette, A. Schick-Makaroff, K. & Molzahn, A.E. (2019). Open Data in Qualitative
544   Research. *International Journal of Qualitative Method,* 18, 1-6.
545   https://doi.org/10.1177/1609406918823863

546 Corti, L., Day, A. & Backhouse, G. (2000). Confidentiality and informed consent:
547   Issues for consideration in the preservation of and preservation of access to
548   qualitative data archives. *Forum Qualitative Sozialforschung/Forum:*
549   *Qualitative Social Researc*h http://www.qualitative-research.net/fqs-texte/3
550   -00/3-00cortietal-e.htm.

551 Corti, L. Fielding N. & Bishop, L. (2016). Editorial for Special Edition, Digital
552   Representations: Re-Using and Publishing Digital Qualitative Data. *Sage Open*,
553   1-3. DOI: 10.1177/2158244016678911

554 Drew, P. & Heritage, J. (1992). *Analyzing Talk at Work. Interaction in Institutional*
555   *Settings*. Cambridge: Cambridge University Press.

556 DuBois, J. M., Strait, M. & Walsh, H. (2018). Is It Time to Share Qualitative Research
557   Data? *Qualitative Psychology,* 5(3), 380-393. doi: 10.1037/qup0000076

558 Garfinkel, H., Lynch, M. & Livingstone, E. (1981). The work of a discovering science
559   construed with materials from the optically discovered pulsar. *Philosophy of the*
560   *Social Sciences,* 11(2), 131–158.

561 Gibson, S. (2019). *Arguing, Obeying and Defying: A Rhetorical Perspective on Stanley*
562   *Milgram's Obedience Experiments*. Cambridge University Press.

563 Hammersley, M. (2010). Can We Reuse Qualitative Data Via Secondary Analysis.
564   Notes on Terminological and Substantive Issues. *Sociological Research*
565   *Online,* 15(5), 1-7. https://doi.org/10.5153/sro.2076

566 Hardy, L. J., Hughes, A., Hulen, E. & Schwartz, A. L. (2016). Implementing qualitative
567   data management plans to ensure ethical standards in multi-partner
568   centers. *Journal of Empirical Research on Human Research Ethics,* 11(2), 191
569   198. https://doi.org/10.1177/1556264616636233

570 Harris, J., Theobald, M. A., Danby, S. J., Reynolds, E., & Rintel, S. (2012). "What's
571   going on here?" The pedagogy of a data analysis session. In A. Lee & S. J.
572   Danby (Eds.), *Reshaping doctoral education: International Approaches and*
573   *Pedagogies* (pp. 83–96)*.* London: Routledge.

574 Heaton, J. (2008). Secondary analysis of qualitative data: An overview. *Historical*
575   *Social Research/Historische Sozialforschung,* 33(3), 33-45.

576 Hepburn, A. & Brown, S. D. (2001). Teacher stress and the management of
577   accountability. *Human Relations,* 54(6), 691-715.

578 Heritage, J., Robinson, J. D., Elliott, M. N., Beckett, M. & Wilkes, M. (2007). Reducing
579   patients' unment concerns in primary care: the difference one word can make.
580   *Journal of Gen Intern med,* 22(10), 1429-33.

581   Heritage, J. (2012). Epistemics in Action: Action Formation and Territories of
582   Knowledge. *Research on Language & Social Interaction,* 45(1), 1-29.
583   https://doi.org/10.1080/08351813.2012.646684
584
585   Holt, E. (1996). Reporting on Talk: The use of Direct Reported Speech in Conversation.
586   *Research on Language and Social Interaction,* 29(3), 219-245.
587
588   Hopkins, M. (1993). Is anonymity possible? Writing about refugees in the United
589   States. In C. B. Brettell (Ed.), *When they read what we write: The politics of*
590   *ethnography* (pp. 121-129). Westport, CT: Bergin & Garvey.
591
592   Hughes, K., Hughes, J. & Tarrant, A. (2020). Re-approaching interview data through
593   qualitative secondary analysis: interviews with internet gamblers. *International*
594   *Journal of Social Research Methodology*, 23(5), 565-579.
595   https://doi.org/10.1080/13645579.2020.1766759

596   Humă, B. & Joyce, J. B. (frth). 'One Size Doesn't Fit All'. Lessons from Interaction
597   Analysis on Tailoring Open Science Practices to Qualitative Research.

598   Irwin, S. (2013). Qualitative Secondary Data Analysis: Ethics, Epistemology and
599   Context. *Progress in Development Studies,* 13(4), 295-306.
600   https://doi.org/10.1177%2F1464993413490479

601   Jefferson, G. (1988). On the sequential organization of troubles talk in ordinary
602   conversation. *Social Problems,* 35(4), 418-442.

603   Jepson, M., Salisbury, C., Ridd, M. J., Metcalfe, C., Garside, L. & Barnes, R. (2017).
604   The 'One in a Million' study: creating a database of UK primary consultations.
605   *British Journal of General Practice*, 67(658), e345-e351.
606
607   Jordan, B. & Henderson, A. (1995). Interaction Analysis: Foundations and Practice. *The*
608   *Journal of the Learning Sciences,* 4(1), 39-103.

609   Kuula, A. (2011). Methodological and Ethical Dilemmas of Archiving Qualitative Data.
610   *IASSIST Quarterly*, 34(3-4), 12-17.

611   Law, M. (2005) Reduce, reuse, recycle: Issues in the secondary use of research data.
612   *IASSIST Quarterly,* 29, 5–10.

613   Liddicoat, A. (2007). An Introduction to Conversation Analysis. London: Continuum.

614   Lynch, M. (2002). From naturally occuring data to naturally organized ordinary
615   activities: comment on Speer. *Discourse Studies*, 4(4), 531-537.

616   Lynch, M. (1993). Scientific Practice and Ordinary Action: Ethnomethodology and
617   Social Studies of Science. New York, Cambridge University Press.

Mair, M., Brooker, P., Dutton, W. & Sormani, P. (2020). Just what are we doing when we're describing AI? Harvey Sacks, the commentator machine, and the descriptive politics of the new artificial intelligence. *Qualitative Research,* 21(3), 341-359. https://doi.org/10.1177%2F1468794120975988

Mauthner, N.S., Parry, O. & Backett-Milburn, K. (1998). The data are out there, or are they? Implications for archiving and revisiting qualitative data. *Sociology, 32*(4), 733-745.

Mauthner, N.S. & Parry, O. (2009). Qualitative data preservation and sharing in the social sciences: On whose philosophical terms?. *Australian Journal of Social Issues,* 44(3), 291-307. https://doi.org/10.1002/j.1839-4655.2009.tb00147.x

Maynard, D. & Clayman, S. E. (1991). The diversity of ethnomethodology. *Annual Review of Sociology,* 17, 385–418

Meredith, J. & Stokoe, E. (2014). Repair: Compaing Facebook 'chat' with spoken interaction. *Discourse & Communication,* 8(2), 181-207. https://doi.org/10.1177%2F1750481313510815

Milgram, S. (1974). *Obedience to Authority: An Experimental View*. Harper & Row.

Mondada, L. (2014). Ethics in Action: Anonymisation as a Participant's Concern and a Participant's Practice. *Human Studies,* 37(2), 179-209.

Mondada, L. (2018). The multimodal interactional organization of tasing: Practices of tasting cheese in gourmet shops. *Discourse Studies,* 20(6), 743-769. https://doi.org/10.1177%2F1461445618793439

Moore, N. (2007). (Re) using qualitative data?. *Sociological Research Online,* 12(3), 1-13.

Mozersky, J. Walsh, H. Parsons, M. McIntosh, T. Baldwin, K. and DuBois, J.M. (2020a). Are we ready to share qualitative research data? Knowledge and preparedness among qualitative researchers, IRB Members, and data repository curators. *The International Association for Social Science Information Service and Technology Quarterly,* 8(43), 1-26. doi:10.29173/iq952.

Mozersky, J., Parsons, M., Walsh, H. Baldwin, K. McIntosh, T. & DuBois, J.M. (2020b). Research Participant Views Regarding Qualitative Data Sharing. *Ethics and Human Research,* 42(2), 13-25. https://doi.org/10.1002/eahr.500044

Parry, O. & Mauthner, N.S. (2004). Whose Data Are They Anyway? Practical Legal and Ethical Issues in Archiving Qualitative Research Data. *Sociology,* 38(1), 139-152. https://doi.org/10.1177%2F0038038504039366

Parry, R., Pino, M., Faull, C. & Feathers, L. (2016). Acceptability and design of video-

658   based research on healthcare communication: evidence and recommendations.
659   *Patient education and counselling*, 99(8), 1271-1284.
660

661 Potter, J. (1996). Discourse Analysis and Constructionist Approaches: Theoretical
662   Background. In J.T.E. Richardson (ed.) *Handbook of Qualitative Research*
663   *Methods for Psychology and the Social Sciences*. Leicester: BPS. Books: 125
664   -140.
665

666 Potter, J. (2002). Two kinds of natural. *Discourse Studies*, 4(4), 539-542.
667

668 Potter, J. & Hepburn, A. (2005). Qualitative interviews in psychology: problems and
669   possibilities. *Qualitative Research in Psychology,* 2, 281-307.
670   https://doi.org/10.1191/1478088705qp045oa
671

672 Pownall, M., Talbot, C. V., Henschel, A., Lautarescu, A., Lloyd, K. E., Hartmann, H.,
673   Darda, K. M., Tang, K. T. Y., Carmichael-Murphy, P. & Siegel, J. A. (2021).
674   Navigating Open Science as Early Career Feminist Researchers. Psychology of
675   Women Quarterly, 45(4), 526-539.
676   https://doi.org/10.1177%2F03616843211029255
677

678 Psathas, G. (1990). Introduction: Methodological issues and recent developments in the
679   study of naturally occurring interaction. In G. Psathas (Ed.), *Interaction*
680   *Competence* (pp. 1-30). Washington DC: International Institute for
681   Ethnomethodology and Conversation Analysis.
682

683 Psathas, G. (1994). *Conversation Analysis: The Study of Talk-in-Interaction*
684   (Qualitative Research Methods). SAGE Publications.
685

686 Real Complaints (2021). *Information for Participants*. Accessed 18.04.21. Available at:
687   https://www.realcomplaints.org/research.
688

689 Sacks, H. (1984). Notes on Methodology. In J. M. Atkinson & J. Heritage (Eds.),
690   *Structures of Social Action: Studies in Conversation Analysis* (pp. 21-27).
691   Cambridge: Cambridge University Press.
692

693 Sacks, H. (1992). *Lectures on Conversation*, vols. 1 and 2, edited by G. Jefferson.
694   Oxford: Basil Blackwell.
695

696 Schegloff, E. A. (1992). In Another Context. In A. Duranti & C. Goodwin (Eds.),
697   *Rethinking Context: Language as an Interactive Phenomenon* (pp. 191-227).
698   Cambridge: Cambridge University Press.
699

700 Schegloff, E.A. (1997). Whose text? Whose context? *Discourse & Society,* 8, 165–87.
701   https://doi.org/10.1177%2F0957926597008002002
702

703 Schegloff, E.A. (1998). Reply to Wetherell. *Discourse & Society,* 9, 413–16.
704   https://doi.org/10.1177%2F0957926598009003006
705

706 Schegloff, E. A. (1996a). Confirming allusions: Toward an empirical account of action.
707   *American Journal of Sociology,* 104, 161-216.
708   https://www.jstor.org/stable/2782190
709

710 Schegloff, E.A. (1999b). "Schegloff's Texts" as "Billig's Data": a critical reply.
711   *Discourse & Society,* 10, 558–72.
712   https://doi.org/10.1177%2F0957926599010004006
713

714 Schegloff, E.A. (1999c). Naivety vs. sophistication or discipline vs. self-indulgence: a
715   rejoinder to Billig. *Discourse & Society,* 10, 577–82
716

717 Sidnell, J. (2011). *Conversation Analysis: an introduction.* London: Wiley-Blackwell.
718

719 Siegel, J. A. & LaMarre, A. (2019). Navigating "publish or perish" as qualitative
720   researchers. *Nature behavioural and social sciences.*
721   https://socialsciences.nature.com/posts/54648-navigating-publish-or-perish-as-
722   qualitative-researchers
723

724 Silverman, D. (1998). *Harvey Sacks and Conversation Analysis.* Cambridge: Polity
725   Press.
726

727 Speer, S. A. (2002a). 'Natural' and 'contrived' data: a sustainable distinction?
728   *Discourse Studies*, 4(4), 511-525.
729

730 Speer, S. A. (2002b). Transcending the 'natural'/'contrived' distinction: a rejoinder to
731   ten Have, Lynch and Potter. *Discourse Studies*, 5(5), 543-548.
732

733 Speer, S. A. & Hutchby, I. (2003). From Ethics to Analytics. Aspects of Participants'
734   Orientations to the Presence and Relevance of Recording Devices. *Sociology,*
735   37(2), 315-337. https://doi.org/10.1177%2F0038038503037002006
736

737 Stevanovic, M. & Weiste, E. (2017). Conversation-analytic data session as a
738   pedagogical institution. *Learning Culture and Social Interaction,* 15, 1-17.
739

740 Stokoe, E. (2014). The Conversation Analytic Role-play Method (CARM): A method
741   for training communication skills as an alternative to simulated role-play.
742   *Research on Language and Social Interaction,* 47(3), 255-265.
743

744 Suchman, L. A. (2007). *Human-machine reconfigurations: Plans and situated actions*.
745   Cambridge: Cambridge University Press.
746

747 ten Have, P. (2002). Ontology or methodology? Comments on Speer's 'natural' and
748   'contrived' data: a sustainable distinction? *Discourse Studies*, 4(4), 527-530.
749

750 ten Have, P. (2007). *Doing Conversation Analysis*. SAGE Publications.
751

752 Tsai, A.C., Kohrt, B.A., Matthews, L.T., Betancourt, T.S., Lee, J.K., Papachristos, A.V.,
753   Weiser, S.D. & Dworkin, S.L. (2016). Promises and pitfalls of data sharing in

754   qualitative research. *Social Science & Medicine,* 169*,* 191-198.

755   https://doi.org/10.1016/j.socscimed.2016.08.004

756

757 UK Research and Innovation (UKRI) (2021). *Open research.* Available from:

758   https://www.ukri.org/about-us/policies-standards-and-data/good-research

759   -resource-hub/open-research/ [Accessed 15th April 2021]

760

761 Van Dijk, T. A. (2007). Comments on Context and Conversation. In N. Fairclough, G.

762   Cortese, & P. Ardizzone (Eds.), *Discourse and Contemporary Social Change*

763   (pp. 290-295). Bern: Peter Lang.

764

765 Walker, G. (2017). Visual Representations of Acoustic Data: A Survey and

766   Suggestions. *Research on Language and Social Interaction,* 50(4), 363-387.

767   https://doi.org/10.1080/08351813.2017.1375802

768

769 Wetherell, M. (1998). Positioning and interpretive repertoires: conversation analysis

770   and poststructuralism in dialogue. *Discourse & Society,* 9, 387-412.

771   https://doi.org/10.1177%2F0957926598009003005

772

773 Williams, B., Dowell, J., Humphris, G., Themessl-Huber, M., Rushmer, R., Ricketts, I.,

774   Boyle, P. & Sullivan, F. (2010). Developing a longitudinal databse of routinely

775   recorded primary care consultations linked to service use and outcome data.

776   *Social Science & Medicine,* 70(3), 473-478. DOI:

777   10.1016/j.socscimed.2009.10.025

778

779 Wilkinson, R. (2015). Conversation and aphasia: advances in analysis and intervention.

780   *Aphasiology,* 29(3), 257–268. https://doi.org/10.1080/02687038.2014.974138

781

782 Wooffitt, R. (2005). *Conversation analysis and discourse analysis: a comparative and*

783   *critical introduction.* London: Sage.

784

785 **Acknowledgements**

789

790 **Contributions**

791 TD conceived the initial idea for the paper. The manuscript was drafted by JJ and TD.

792 BB, CR and JJ contributed to the development of key ideas, the structure, the rewriting

793 of several sections of the manuscript, and the editing of the manuscript. RP and RS

794 helped develop the initial ideas on which the paper was written, and, along with AK

795 offered feedback and guidance on drafts of the paper.