UNDERSTANDING CONVERSATION AROUND TECHNOLOGY USE IN CASUAL-SOCIAL SETTINGS

MARTIN PORCHERON

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy.

August 2019

Martin Porcheron: *Understanding Conversation Around Technology Use in Casual-Social Settings*. © August 2019

SUPERVISORS:
Joel E. Fischer
Sarah Sharples

For all those I love,
friends and family alike.

ACKNOWLEDGEMENTS

Firstly, I would like to sincerely thank my two supervisors at the University of Nottingham, Joel E. Fischer and Sarah Sharples. Joel, from the Mixed Reality Lab, and Sarah, from the Human Factors Research Group. Both have provided continual encouragement, critique, support, care, and dedication to my work and personal development. My sincere gratitude is also due to the countless friends and colleagues in the lab, especially Stuart Reeves, who have engaged (and some might say endured) numerous provocative discussions with me throughout my PhD. Their critique, although difficult to receive at times, was ultimately always welcomed, appreciated, and most supportive.

The Horizon Centre for Doctoral Training, which supported me financially during this PhD, and the Mixed Reality Lab, where I have been based for the majority of the time, have always been welcoming and warm environments in which one can explore ideas, push boundaries, and develop a career. I am grateful to both centres.

During this journey I met—and then married!—Juan Diego. He has given me so much care, kindness, love, and support, and is a wonderful companion with whom I can explore the world and learn more about myself. He makes the world seem less scary, always encourages me to stay focused when I need it, and continually empowers me to achieve my personal goals. For this and so much more, I am unspeakably grateful. I love you.

My mother, Nicole, and father, Andrew, were both there for me whenever I needed them and they often realised I needed their help long before I did (although I am sure they feel I never need them anymore, but I assure them it is not true!). As far I can remember they both have encouraged me to be bold and courageous because they believed in me and believed I could do what I set my sights on.

I am so dearly grateful for all the support and love they have given to me. Thank you, both.

I would also like to thank (in alphabetical order, as there is no way I could order this list): Susan Ali, Kyle Arch, Pepita Barnard, Felicia Black, José Gustavo Berumen Salazar, Liz Dowthwaite, Carolina Fuentes, Panagiotis Koutsouras, Shalaka Kurup, Horia Maior, Juan Pablo Martínez Ávilla, Annie Quandt, Richard Ramchurn, Andreas Reiter, Neelima Sailaja, Gareth Shaw, Jocelyn Spence, Sam Stapleford-Allen, Hanne Wagner, and Matt Wilkinson. I am indebted to Andrés Lucero for offering and supporting me through my internship. He provided wonderful insights into his work practices and always shared his valuable experiences and knowledge with me. He has become a valuable colleague as well as a dear friend.

All of these people—and so many more—have shared so much of this journey with me and I have developed some truly cherished friendships. Furthermore, to all of my friends and family, of which there too many to name here, who have been there for me throughout, and have celebrated my highs and helped me through the lows, your support and love made this all possible.

Finally, my thanks to the School of Computer Science and Mixed Reality Lab for supporting travel to CHI 2015 and 2016, MobileHCI 2015, and CSCW 2016 and 2017; the Horizon CDT, the Human Factors Research Group, and the Mixed Reality Lab for supporting travel and accommodation during my internship; and ACM SIGCHI and the university's Graduate School for supporting travel to CSCW 2017. Ultimately, none of this would be possible without the Horizon Centre for Doctoral Training at the University of Nottingham (Grant No. EP/G037574/1) and by the Horizon Digital Economy Research Institute (Grant No. EP/G065802/1), who supported this project.

ABSTRACT

Technologies such as smartphones have had a profound impact on everyday activities, providing us with instant access to a boundless archive of information and real-time communications with others. Additionally, new devices such as standalone 'smartspeaker' devices, which are controlled entirely through voice, are making an appearance in the home. This instantaneity and pervasiveness of these technologies has not unfolded without critique, however, as many complain of situations where we are surrounded by friends but become distracted, lose track of the conversation, or have trouble balancing demands both from our devices and the conversational involvement.

Through the completion of three observational studies, this thesis reveals how people bring their device use into a multi-party conversation as part of socialising together. Furthermore, these interactions are shown to unfold collaboratively, with co-present others supporting device users to complete their interactional projects occasioned through the conversation.

This thesis studies three such technologies and examines how interaction with and around the device situationally unfolds:

- device interaction where a touchscreen is used as the primary input and output mechanism,
- device interaction using touchscreen-based devices that also feature voice-controlled interfaces that can be spoken to, and can respond by synthesising speech, and
- device interaction using voice only, where the device is controlled using spoken 'natural language' as the input mechanism, and the device synthesises speech as its response.

Each study takes place in what is categorically called a 'casual-social' setting, with these settings demarcated as places where people gather to socialise and relax. The first two of these studies adopt a participant-observer approach, and all three employ an analytic lens based on ethnomethodology. The first study observes groups of friends socialising in a pub together and reveals how natural device use becomes occasioned in and through interaction and how the device use is interleaved amongst talk in the setting. Study two adopts a similar premise but asks participants to preferentially use their device's voice-controlled 'personal assistant' in situations where they would typically type into a device. This reveals how people accountably and interactionally accomplish this practice. In both of these studies, participants were not asked to use technology or perform any activity other than to socialise together. The final study explores the use of smartspeakers in the home through a longitudinal study. These devices are designed with 'far-field' microphones to allow users to speak to them at a distance, and with speakers to allow the device to respond using a synthesised voice. Through capturing their use over a one-month period, the study reveals how devices are used as part of the multi-activity home, alongside other ongoing activities.

The thesis makes a number of contributions, such as identifying for what purposes and how people use devices in and through socialising together in a casual-social setting, including the collaborative nature of these interactions. Dealing with technical troubles with devices such as smartspeakers was identified as being a potentially collaborative activity amongst the groups, and synthesised responses from the Voice User Interfaces (VUIs) were ostensibly 'resources' for dealing with these troubles and progressing interaction with the device. Through its presentation of thick description and analysis, this thesis establishes the case for further research to examine and design for gatherings in such settings.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CSCW Computer-Supported Cooperative Work

CVR Conditional Voice Recorder

GUI Graphical User Interface

HCI Human-Computer Interaction

IoT Internet of Things

IPA Intelligent Personal Assistant

VUI Voice User Interface

LIST OF PUBLICATIONS

Parts of this thesis have been accepted by peer-review for publication in conference proceedings in Human-Computer Interaction (HCI) or Computer-Supported Cooperative Work (CSCW):

- The empirical work presented in Chapter 4 is published as:
  - Porcheron, M., Fischer, J. E. and Sharples, S. C. (2016a). 'Using Mobile Phones in Pub Talk'. In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '16. New York, NY, USA: ACM, pp. 1647–1659. ISBN: 9781450335928. DOI: 10.1145/2818048.2820014
- The empirical work presented in Chapter 5 is published as:
  - Porcheron, M., Fischer, J. E. and Sharples, S. C. (2017b). '"Do Animals Have Accents?": Talking with Agents in Multi-Party Conversation'. In: *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '17. New York, NY, USA: ACM, pp. 207–219. DOI: 10.1145/2998181.2998298
- The empirical work presented in Chapter 6 is published as:
  - Porcheron, M., Fischer, J. E., Reeves, S. and Sharples, S. (2018). 'Voice Interfaces in Everyday Life'. In: *Proceedings of the 2018 ACM Conference on Human Factors in Computing Systems*. CHI '18. ACM, New York, NY, USA: ACM. DOI: 10.1145/3173574.3174214

    This paper received a Best Paper award, reserved for the top 1% of submissions to CHI '18.

Other work in this thesis has been published in non-archival or curated venues:

- Ideas around how people interact with technology while in close proximity to each other were discussed and shaped in the workshop:
  - Porcheron, M., Lucero, A., Quigley, A., Marquardt, N., Clawson, J. and O'Hara, K. (2016). 'Proxemic Mobile Collocated Interactions'. In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '16. New York, NY, USA: ACM, pp. 3309–3316. ISBN: 9781450340823. DOI: 10.1145/2851581.2856471

  and tangentially explored in design:

  - Porcheron, M., Lucero, A. and Fischer, J. E. (2016b). 'Co-Curator: Designing for Mobile Ideation in Groups'. In: *Proceedings of the 20th International Academic Mindtrek Conference*. AcademicMindtrek '16. New York, NY, USA: ACM, pp. 226–234. ISBN: 9781450343671. DOI: 10.1145/2994310.2994350
- The findings in Chapters 5 and 6 were shaped by discussions at the workshop:
  - Porcheron, M., Fischer, J. E., McGregor, M., Brown, B., Luger, E., Candello, H. and O'Hara, K. (2017a). 'Talking with Conversational Agents in Collaborative Action'. In: *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. CSCW '17 Companion, pp. 431–436. ISBN: 9781450346887. DOI: 10.1145/3022198.3022666

I am grateful to all the reviewers of this work, the questions I received following presentations, and feedback and invaluable discussions, suggestions, ideas, recommended reading, and critique I received from countless attendees at workshops.

1

INTRODUCTION

As new technologies are introduced and increase in ubiquity they run the chance of being criticised for their impact on social order, and indeed few technologies can escape such a critique of their impact on society. Devices such as smartphones and smartspeakers are no exception, with examples ranging from popular press (e.g. Turkle (2011)) through to academia (e.g. Su and Wang (2015)) identifying troubles with the very act of 'using' of devices in the presence of others, identifying myriad negative aspects to such interactions. These devices are, however, designed such that they *can* be used with ease as part of everyday life, and are pitched as allowing users complete interactions in seconds (Brown et al., 2014). It is this notion of the problematisation of device use in society, and an academic desire to understand how users *practically* get the devices to work in and through conversation, that motivates the work within this PhD thesis.

This thesis will critically examine casual-social interactions amongst groups of people, and how people engaged in everyday talk draw upon three such technologies: smartphones, smartphones with voice-based personal assistants, and smartspeakers with voice-based personal assistants, as resources within the conversation. Precisely, the empirical chapters of this thesis unpack the gloss of *what it is to use a device* and *how this is done* while people are collocated with others[1] in a casual-social setting.

This thesis studies the social interaction around the use of devices in casual-social multiactivity settings, where multiactivity broadly refers to "the social, interactional and temporal features of situations

1 Within related literature the terms *co-located* and *collocated* are often treated as synonymous and interchangeable. Here, and throughout the thesis, the term *collocated* was arbitrarily used for consistency with publications on which this thesis is based.

and conduct in which people organise multiple activities together, concurrently or serially" (Haddington et al., 2014, p. 5) and in which the members of the setting are involved in multiple activities (Goffman, 1968). This thesis is not the first piece of literature to use the term casual-social setting, with it applied to places for smokers (Schane et al., 2009), hotel suites (Pigram, 1996), college classrooms (Yamada, 1981, pp. 30–31), and as places where designers should design mobile interactions for (Reis et al., 2012). St. Lawrence et al. (1983) define such a setting as a place where "people can openly meet and interact with one another" (ibid., p. 42). Additionally, the definition used in this thesis is also similar although not completely congruent to the notion of "third places" (Oldenburg, 1989). Third places are spaces that are outside the home or workplace where people can gather, socialise, and relax. Casual-social settings expand this definition to include any space, public or private, that is in or outside the home where socialising in a relaxed manner is a primary activity (i.e. the setting has an innate lack of formality one might expect from a work environment). In the context of this thesis, settings were selected for the fieldwork where the device use would not be considered 'out of place', i.e. the device use would be perspicuous to the setting (Garfinkel, 2002a, p. 181) and would be expected to unfold.

This thesis will critically explore conversations and how the use of technology is interleaved with them *in vivo*, and show that using a device in a social setting with others, although problematic at times, is routinely accomplished in and through everyday conversation. In order to do this, this thesis adopts an analytic perspective in line with ethnomethodology as defined through the work of Garfinkel (1967), as well as others within Human-Computer Interaction (HCI) and Computer-Supported Cooperative Work (CSCW) such as Heath et al. (2010) and Crabtree et al. (2012). Adopting this perspective will allow this thesis to show that people can successfully occasion technology use in and through conversation, can interleave interactions with the device while interacting with others around them, and can

account for and co-manage device interaction collaboratively with co-present others (see Chapter 4). While such a perspective precludes a stance on the *morality* of using devices (see Chapter 3), it does allow for an analytic orientation that reveals that, although the use of devices does engender problematic interactional sequences, these problems are quickly and methodically attended to in and through interaction collaboratively.

Research in related domains such as Mobile HCI and CSCW has long explored how to design better interactive collocated experiences based around cooperative technology use (e.g. Lundgren et al. (2015)). Work has explored many different facets of technology use such as augmenting social settings with large screens (Lucero et al., 2012), making use of mobile applications for souvenir generation (Durrant et al., 2011), cultural visiting (Fosh et al., 2013), crowdsourcing video of spectator events (Flintham et al., 2015), and connecting public displays to benefit public life and communities (Memarovic et al., 2016). The work in this thesis, however, is crucially concerned with existing practices of using technology in everyday social interactions; that is, this thesis does not intend to introduce new or augment existing technologies but will study how technology that is already 'in-the-wild' is used. In particular, this concern is with three types of device interaction: touchscreen-based portable device use, VUI use with touchscreen-based portable devices, and VUI use with non-portable smartspeakers. Therefore this work, although embedded in the HCI domain, pivots more towards studies of work as found in CSCW, with a preference for *informing design* through the examination of how people who own or have these devices in their lives make use of them in and through interaction.

The remainder of this chapter will introduce the specific problem-space under examination in this thesis and will identify how the following chapters will attempt to answer the research questions posed.

1.1 PROBLEM DEFINITION AND SPACE

It is an increasingly commonplace practice for people to use a device while around others (Brown et al., 2014), and this is often characterised as leading to problematic or undesirable situations. Conversely, however, there has been a multitude of positive reasons for technology use identified, such as dealing with anxiety (Wei and Lo, 2006), remaining in touch with family members and friends (Harmon and Mazmanian, 2013), and information retrieval that is of benefit to the user (Sohn et al., 2008). Furthermore, design work has successfully attempted to capitalise on the availability of technology to create collaborative and creative experiences in workplaces, homes, and in public spaces (e.g. Fatah gen. Schieck et al. (2014)). There remains a problem such that, although device use is often characterised as negative, there also is a wide range of positive aspects attributed to it.

Additionally, in spite of considerable progress in mobile technology, critical voices (e.g. Su and Wang (2015) and Turkle (2011)) have pointed out the ways in which the use of devices may isolate people from one another in social situations, or change the perception of the setting itself. On the other hand, socio-technical studies have shown people are skilled at interleaving and *embedding* mobile device use and social interaction in, for example, a living room (Rooksby et al., 2015), or, in the completion of specific tasks such as collaborative photo-taking setting (Durrant et al., 2011) or mobile search (Brown et al., 2015).

The work in this thesis follows on in the traditions of ethnomethodology (Garfinkel, 1967), and indeed adopts "ethnomethodological indifference" (Garfinkel and Sacks, 1970), as discussed in Chapter 3. In introductory terms, this work is not occupied in understanding models of interaction based on theoretical reasoning and considers interaction not only the site of study but the commodity with which the findings are established. This precipitates a stance to disregard the adoption of *a priori* theories of interaction in context, instead en-

gendering an approach that involves *going and seeing* what is done to construct meaningful findings from the explication of how individuals use a device as a resource in casual-social interactions.

With this theory-agnostic approach in mind, this thesis focuses explicitly on *how* people accountably use devices as resources in conversation, and collaborate in and through their device use in casual-social settings while collocated with others. This work is opposed to making moral judgements and will instead select conversation with device use interleaved within it as a site of study, to reveal *what* is accomplished in and through the device use, and *how* it is done. Although the rhetoric of device interactions as negatively impacting upon a collocated interaction has been established, it loses sight of the individual interactional achievements of members in the settings, which is, as yet, unstudied.

Therefore, in summary, the primary objective of this thesis is to explicate how and what for purpose devices are brought into ongoing everyday conversations in casual-social settings. The need to do this is motivated by the current gap in the literature that exists. There is nascent work that details the use of devices in such settings as an interactional accomplishment although there remains multiple gloss-like accounts of detrimental device use, which on an interactional level diminishes the work done by members in the setting to make device use accountable and embedded within conversation. Through the presentation of empirical data which unpacks the work of a relaxed multi-party conversation, this thesis will show how people successfully bring device use into the conversation to accomplish a problem at hand.

1.2 DEVICES UNDER STUDY

Given the problem defined above, and in line with recent developments with technology, this thesis will examine how interaction unfolds with respect to portable electronic devices such as touchscreen-based smartphones using (1) the touchscreen, and (2) with the voice-based interfaces found on the device. With regard to the second form of interaction, these Voice User Interfaces (VUIs) are interacted with by talking to the device in a "conversational" manner, with the device typically responding back in a 'conversational' manner on the screen or through simulated speech[2], with the device also making use of the Graphical User Interface (GUI) to display details about the computation and response of the user's request.

Furthermore, in line with the development of strictly voice-based interfaces (in the form of smartspeakers), this thesis will further unpack the interactional accomplishment of interleaving interaction with a VUI smartspeaker in conversation in the home. This final form under study in this thesis is where the modality of interaction is restricted solely to voice, i.e. situations where the only way to interact with the device is to talk to it, and for the device to respond with the requested action audibly and/or with synthesised talk.

Each of these devices is in different stages of mass adoption, but are widely considered ubiquitous technologies. Tangentially, theories surrounding the adoption of new technology, such as the 'Diffusion of innovations', provide an abstract understanding of technological adoption across society, which provides us with a point of reference to make sense of the how technologies are adopted by consumers at a macro-societal level (E. M. Rogers, 1995). This theory demarcates adopters of technology into arbitrary labels of 'innovators', 'early adopters', 'early' and 'late majority' and 'laggards', with the latter two

2 There is a veritable smorgasbord of different terminologies for these interfaces, such as *Conversational User Interface*, *Conversation(al) Agent*, *Intelligent Personal Assistant*, *Virtual Personal Assistants*, and so on...

categories deemed to be groups who have adopted after the 'majority of society'.

Smartphones, at the time writing this thesis (i.e. 2017), have existed for some 20 years, initially as industrial research prototypes before mass adoption. In the year prior to the first study of this thesis, Ofcom, the United Kingdom's communications regulator, remarked that the UK is now a "smartphone society", with 66% of households having at least one smartphone (Ofcom, 2015, p. 6), which would classify the smartphone adoption as already in a state of mass adoption. Conversely, at the time of work being undertaken for the second study, one survey identified personal assistants on portable devices as being used by used by 32% of respondents in the last year (Ask Your Target Market, 2016), situating it as in the stage of being adopted by the 'early majority'. Even more so, smartspeakers were included in the UK's household measure of consumer inflation (Office for National Statistics, 2019) a year or so after the underlying research in the third empirical chapter was completed, underscoring their rapid growth and pervasiveness despite only been recently released.

While these labels provide little insight in the context of this thesis' aims to understand the interactional accomplishments of bringing the device use into conversation, they proffer an understanding of the broader *context* in which the device use is brought about. In other words, they establish the backdrop against which these devices are being studied—each technology studied in this thesis is a pervasive technology that is widely used (and continuing to grow in use). The use of portable devices is so pervasive that their use regularly features in casual-social settings, as highlighted above by the critiques in literature, and that the use of smartspeakers in the home has rapidly grown to the point where it is included in national measures of inflation just a short time after this thesis was produced. This thesis does not offer to examine interaction that interleaves the use of all devices in all situations as a definitive study, but merely selects three technologies that are pervasive, have been documented as widely owned

and used, and seeks to explicate the ways in which the technology is drawn upon in conversation.

1.3 RESEARCH QUESTIONS

To answer the gap discussed above, this thesis will identify what is accomplished through the use of devices in conversation in a casual-social setting and crucially, how this device use is done. The overarching research question that creates the foundation for this thesis' contributions is:

RQ How are devices used within a multi-party conversation in a casual-social setting?

This will be achieved by performing an ethnographic study of conversations amongst groups of friends socialising together, and orienting to instances where devices are used in and through the conversation. By adopting an ethnomethodological lens to fieldwork and analysis (see Chapter 3), this thesis will show the methods through which members within a setting occasion and embed a device interaction within the wider social context. This question can be segmented into components that will be answered through the delivery of this ethnographic study.

RQ_A What is accomplished in and through the use of devices in casual-social settings?

RQ_B How is this device use interactionally and accountably organised?

Both of these questions are very much 'two sides of the same coin', revealing the nature of how and for what purpose people use a device in conversation. Therefore, the overall goal of this thesis is to develop an understanding of the efforts of individuals as they interact with each other and bring the use of everyday devices into conversation as a resource to address matters as they arise. Through this orientation

to *what* and *how* this device use unfolds, members' technical troubles with devices may be identified in interaction. Through this analytic stance, this ethnography will, in turn, identify how individuals and groups attend to these as matters in and through the conversation to accountably organise and accomplish the occasioned activity.

1.4 RESEARCH AREAS

This thesis adopts an interdisciplinary approach, drawing upon literature, grounding, and practice from several different areas, and in turn, makes a number of contributions (see 1.5) to the different fields:

- *Ethnomethodology* - Ethnomethodology is the perspective adopted for fieldwork and analysis in this thesis. Each of the three studies in this thesis adopt an applied approach to ethnomethodological analysis and explicate the sequential situated action of members in the setting.
- *Computer-Supported Cooperative Work (CSCW)* - Ethnographic studies in CSCW provide the groundwork and practical foundation to guide the study of device use as an everyday interactional accomplishment. Work in CSCW also uncovers different facets of mobile device use including qualitative and quantitative studies of mobile device use in everyday life[3].
- *Mobile Human-Computer Interaction (HCI)* - There is a plethora of interdisciplinary design work in Mobile HCI to create collaborative collocated experiences with everyday portable technologies.

1.5 CONTRIBUTIONS

The Venn diagram in Figure 1.1 shows an approximation of the influence of each of the research areas discussed previously in 1.4, based

3 Although none—as-yet—serve to reveal the interactional accomplishment of using a device in a casual-social setting, as per the objectives of this thesis.

on size. Furthermore, this diagram shows this thesis' three main contributions and how they are positioned in relation to the areas in which they contribute new knowledge:

A Details of the methodical practice and conduct of people in casual-social settings, detailing how they bring devices into an everyday multi-party conversation, and offering an insight into the differences of how device use is used to address the members' problems that arise in such settings.

B Experience and development of the methodological approach in this thesis, both in terms of the application of ethnomethodology and of the nature in which *these technologies* were studied *in these settings*.

C Insights of how the interactions unfolded with different technologies and how members undertook work to make these interactional projects collaborative. Crucially, through studying interaction with and around such devices, this thesis makes the case for further CSCW studies of such settings given their nature of being sites for technology use, and for HCI to critically examine and ameliorate the challenges members attend to in using these technologies for their interactional projects.

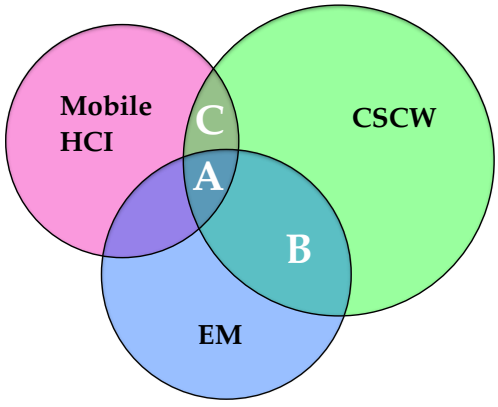


Figure 1.1: Venn diagram of the influence (size) of the different relevant research areas this thesis draws upon, and how the different multidisciplinary contributions of this work are positioned in relation to each area.

1.6 STRUCTURE OF THE THESIS

There are eight chapters structured over three parts in this thesis, as listed below in Table 1.1. A brief summary of each chapter's contribution to the thesis is included in the *Summary* column.

NO.	SUMMARY
Part I: Background and Approach	
2	Continuing the work begun in the Introduction, this chapter lays the groundwork for the thesis by surveying existing literature relating to the use of technology within social and collocated settings, from the perspectives of socio-technical studies, *Mobile* HCI, and CSCW.
3	The methodological approach adopted in this thesis is introduced in this chapter. Ethnography as a research method is described, and included amongst a brief timeline of the development of ethnomethodological tradition. This provides the context with which the fieldwork and analysis were conducted, and will allow the reader to understand the lens with which this thesis has been produced.
Part II: Empirical Work	
4	This chapter presents the study of naturally unfolding interaction with a touchscreen-based portable device within a group of friends socialising together.
5	This chapter presents a similar study of a group of friends socialising, but where the device interaction is achieved using the Voice User Interface (VUI) on the portable device.
6	This chapter unpacks how families and friends talk to a VUI smartspeaker in the home, with data collected as part of a longitudinal study.

Part III: Synopsis	
7	This chapter discusses the findings from the three empirical studies, bringing the findings into the context of existing literature, and how the three independent studies correspondingly reveal the collaborative practices of conversationalists in social settings.
8	This chapter summarises the contributions of this thesis and provides a number of conclusions.

Table 1.1: Structure of Thesis

Part I

BACKGROUND AND APPROACH

2

LITERATURE REVIEW

This chapter introduces existing research on the use of technology while we are engaged in face-to-face interactions with others. This literature provides the motivational foundation that led to the development of the research questions posed, as outlined previously in 1.3. In this regard, this chapter will synthesise the literature that frames the socio-technical and design backdrop that informs the current understanding of device use in and around conversation and face-to-face interaction. Understanding how technology fits in and around our interactions with each other, which remains "the most human thing that we do" (Turkle, 2011, pp. 3), is the key objective of this thesis. This literature review introduces work in three key areas in relation to technology use, starting from a 'big picture' topic, through to work that this thesis is very much situated amongst: (1) studies of technology use in society, (2) studies of systems designed for collocated interaction, and (3) studies of collocated interaction:

STUDIES OF TECHNOLOGY USE IN SOCIETY

This first field of work, found in 2.1, unpacks the critically reflective literature on widespread device ownership and use in society from a socio-technical perspective. This work examines the role that devices play in public and private settings, such as pubs and the home, and crucially focuses on the *impact* of device use, and people's *reflections* of it, helping to establish the backdrop to which the work in this thesis takes place.

SYSTEMS DESIGN FOR COLLOCATED INTERACTION

This second field of work, found in 2.2, will synthesise literature in which systems are designed for supporting and augmenting interaction with others while we are collocated, with

aims to support co-operative working. This section will briefly introduce 'groupware', before discussing how design work in Human-Computer Interaction (HCI) and Computer-Supported Cooperative Work (CSCW) has moved from meeting rooms to more diverse settings, supporting new portable technologies.

CLOSE STUDIES OF COLLOCATED INTERACTION

This final field of work, found in 2.3, will introduce detailed studies of face-to-face interaction upon which the methodological approach of this thesis draws. This literature seeks to examine device use as a matter of course in and through everyday life, by studying users' interactions in-the-wild, or rather, *in vivo*.

By briefly unpacking this literature, this thesis' contribution will be situated as part of this broad and interdisciplinary programme of discourse on the social practices of technology use in casual-social settings. Moreover, this contribution is not based on any theoretical or technical considerations of the design of devices, but on the practices of members in a setting. As such, this work is *indifferent* to models or theories of interaction (see 3.2.3 for an expansion of this point on indifference). In this regard, this thesis' focus is on people and how they accountably attend to device use naturally: it is *a study of people and their interaction*, and not technology and its uses.

This thesis studies three main technological developments: touchscreen smartphones, Voice User Interfaces (VUIs) on touchscreen smartphones, and 'screenless' 'smartspeakers' that only have a VUI. The commentary in this chapter primarily focuses on portable devices such as smartphones and tablets and includes what little nascent literature exists in relation to 'smartspeakers' in the home[1].

1 This is, in part, because the devices only became commercially available in the last two or so years of this thesis being produced.

2.1 TECHNOLOGY USE IN SOCIETY

In this section, a socio-technical perspective of technology use in everyday life is developed to frame the problematic motivational backdrop for the approach and empirical work in this thesis. The utopian view of how technology is briefly introduced, drawing upon Weiser (1991)'s vision of the computer in the 21st century, before literature that reveals a *societal impact* of technology use is discussed. This section then progressively unpacks and synthesises literature that examines and offers a critique of device use, and its 'impact' and influence on everyday life.

2.1.1 *Ubiquitous computing*

Portable devices take many forms, but presently the most prevalent are smartphones and tablets, which have, for the large part, led the charge in realising Weiser (ibid.)'s prospective vision of device ownership and "invisible" use, as outlined in the introduction of this thesis. In his work, Weiser (ibid.) set out a vision of (work) environments, where technology of different shapes and sizes is ubiquitously available and always within reach—its use so finely woven into everyday interactions that it "disappears". This realisation of this is ostensibly led by rapid growth in ownership of smartphones and mobile Internet usage (Poushter, 2016)[2]. More recent innovations, such as voice-activated 'smartspeakers' have also seen rapid growth in the last few years, and can now be found in 32% of US homes (Adobe Inc., 2018).

Yet, of course, the vision is utopian and with all prospective utopian visions, should be treated with caution (Bell and Dourish, 2007). However, much of what Weiser (1991) projected bares hallmarks of the reality of today, and as noted by Bell and Dourish (2007, p. 135): "ubiquitous computing is already here; it simply has not taken the

2 The growth of smartphone ownership is remarkable with annual growth rates in ownership of 10% or higher.

form that we originally envisaged and continue to conjure in our visions of tomorrow". Simplistically, it is noted that our work environments feature lots of technology, and owing to the ever-increasing capabilities of wireless communication technologies such as Wi-Fi and Bluetooth, we can interact with devices such as tablets in meetings for multiple tasks. Of course, outside of the workplace, technology has also become ubiquitous, thanks to the portability and affordability of smartphones and smartspeakers. The next section introduces the socio-technical work that examines the use of ostensibly ubiquitous computing devices in society.

2.1.2 *Togetherness and isolation*

There is much praise for the use of portable devices such as cellphones and smartphones, especially with the notion that mobile devices provide or *enhance* our daily lives by helping us to shape our experiences of the world around us, although such critique typically comes with caveats. Consider social anthropologist Sherry Turkle, for example, who is perhaps one of the most commonly referenced cases. In her oft-cited work "Alone Together: Why We Expect More from Technology and Less from Each Other", which offers a critique of the societal use of technology and weakening 'desires' to interact with each other, she finds merit in the unique qualities and "enhanced experience" provided by the Internet connectivity of modern devices:

> [...] connectivity offers new possibilities for experimenting with identity, and particularly in adolescence, the sense of free space, what Erik Erikson called the *moratorium* [...] [r]eal life does not always provide this kind of space, but the Internet does.
>
> — Turkle (2011, p. 152)

Others have found similar effects, such as notions of togetherness and dwelling established through interviewing participants regarding their use of instant messaging platforms, and specifically Whats-

App (O'Hara et al., 2014). Turkle's praise of technology use remains guarded, however, as she then turns to question the negative effects that mobile device use has in everyday life by noting that, in her view, interactions between people are made problematic by the mere presence and use of devices:

> [...] face-to-face conversations are routinely interrupted by incoming calls and text messages [...] when someone holds a phone, it can be hard to know if you have that person's attention.
>
> — Turkle (2011, p. 161)

She also extends her criticism of devices in later work by claiming that, with the pervasiveness of devices, we lose the sense of wanting to communicate, and in separate work remarks that we must "reclaim conversation" as if it were a dying art form (Turkle, 2015). This criticism rests on notions that we lose the ability to be empathetic because of our use of technology to mediate communication. While it is easy to disregard Turkle's problematisation of devices as those of a pessimist[3], her views have held stock in work elsewhere, across different disciplines, and are, of course, relatable to most people[4].

Indeed, numerous surveys and interviews identify the increased portability and functionality of mobile devices as encouraging the acceptability of their use in many settings, such as pubs and social environments. With this finding comes the implications that the use of devices in public spaces becomes derided and charged as annoying or rude by co-present others, and that interruptions from devices and extended mobile search tasks are a distraction from an ongoing conversation (Ames, 2013; Church et al., 2012; Campbell, 2007). Contradictorily though, Ames (2013) also identifies through their analysis

3 Indeed, many before her have charged other technologies with a similar critique, e.g. consider McDonagh (1950)'s critique that television has transformed conversationalists in the home into mere spectators.

4 In other words, for most people, it is relatively easy to recall a situation where someone was using a device while you were talking to them.

that "while students often expected others to be constantly connected, they were not always available themselves" (Ames, 2013, p. 1494). Ames (ibid.), drawing upon Turkle (2011)'s findings, goes on to highlight the difficulty and ongoing contradiction that exists in relation to device use, as people attempt to balance digital and physical social obligations such as 'staying connected':

> Many expressed concerns about being tethered to "electronic leashes," able to be yanked at any time out of the present [...] [o]thers adopted the values of those around them: when their families or friends derided them for not being fully present or even just having their phone out, they chose to yield to this social pressure. However, most also felt increased anxiety about what their extended network thought about them as a result.
>
> — Ames (2013, p. 1494)

In another case, Humphreys et al. (2013), also through interviews, highlight findings that suggest that the ease of using the 'mobile Internet' potentially exacerbates the problem of "mis-prioritizing communication through their mobile device over and above face-to-face communication" (ibid., p. 501). Additionally, in orienting to the use of mobile devices in public places, and in particular pubs, Su and Wang (2015) state that technology can "threaten conversation by creating the present-but-absent, anti-social, and app-addicted patron" (ibid., p. 1667). The next section turns to a specific matter of being connected at all times, in which devices 'encourage' their use: device notifications, and the research that has examined their occurrence and influence in our daily lives.

2.1.3 *Device notifications*

Moreover, there is a substantive body of work investigating how to 'better' deliver mobile notifications to individuals in the face of the

potentially disruptive nature of such interruptions from portable devices (Cutrell et al., 2000; Fischer, 2010; Lopez-Tovar et al., 2015), and now smartwatches (Cecchinato et al., 2017), drawing on different methods including the observation of groups completing tasks (Fischer et al., 2013) and conducting contextualised interviews (Hudson et al., 2002). Certainly, the existence of such work to tackle interruptions from devices raises the prospect that these interruptions themselves, and the use of devices, in general, led to the issues such as 'social isolation' (Turkle, 2011).

This literature on device interruptions should, therefore, establish that many people see the use of technology in everyday settings as *problematic* and *disruptive*, in part because their use is brought about by interruptions *arising from* devices. But, for example, multiple studies have established that although interruptions from devices, such as notifications, are characterised as problematic, many still prefer to receive notifications than not, for myriad reasons, including "awareness" (Iqbal and Horvitz, 2010) and to avoid the feeling of "being cut off" (Mark et al., 2012, p. 560). Pielot and Rello (2015), by asking participants to disable notifications on their devices for 24 hours and interviewing them afterwards, identified that "many participants were anxious to miss information from significant others and superiors" (ibid., p. 1765).

In reading this literature, it becomes clear that the accepted view is that there are also positive perspectives in relation to the portability and flexibility of mobile devices. This advantage allows for their greater use within many different settings, including conversation, and allows such devices to provide a utility in which people can remain in touch with their extended network. This raises expectations that we should quickly respond to our contacts from our friendship groups, just as we expect them to respond (Ames, 2013). This immediacy provides users with the sense of being "always connected, to be accessible at all times and places" (Peters and Allouch, 2005, p. 240) and removing the "binding between a fixed space and a person's in-

formation and communication resources" (Perry et al., 2001, p. 324). It seems that people find device use problematic because of the burden of managing constant availability (Sadler et al., 2006), yet find it indispensable and the notion of 'untethering' as undesirable because of this very same quality.

Even regarding 'non-notification-instigated' use, devices are identified as potentially being a *beneficiary* to everyday life in the home. For example, Lanigan (2009), in a study of the use of technology in family life, had what they even considered a "surprising finding". The research found that "that the more time families spent engaged with the computer, the higher their level of communication, cohesion, and adaptability", with a home computer encouraging "more frank communication", an increase in "family time [...] due to efficiencies gained through computer use", and "a source of mutual interest" for the family (ibid., p. 603). It seems, then, that devices are likely to remain present in our everyday social interactions. The next section progresses on to examining literature that unpacks how our experience of space is shaped by this device use.

2.1.4 *Device use and space*

Furthermore, exacerbating this point, the use of mobile devices in public settings has been well documented in literature for a variety of purposes, from how an iPod allows an individual to reshape their experience of time and space (Bull, 2005), to how individuals use new technologies such as cellphones to adapt their social perspective (Humphreys, 2005; Oksman and Turtiainen, 2004; Peters and Allouch, 2005), and to the enjoyment and ludic pursuits people explore with devices (Brown and Juhlin, 2015). In relation to how we now make use of devices anywhere and everywhere we go, Geser (2006) presents a sociological review of whether the mobile phone is 'undermining social order'. He concludes that, through review of how 'time-based scheduling and coordination' has declined because of the ready

availability of devices, "a new, more fluid culture of informal social interaction therefore can emerge" (Geser, 2006, pp. 5–6). Furthermore, Campbell and Y. J. Park (2008) argue in an essay that "mobile communication around copresent others [...] personalizes the communal experience of being in that space" (ibid., p. 379); which also supports work by others of the practice of using technologies to create private spaces in public places (e.g. Ames (2013) and Wei and Leung (1999)).

This literature, which draws on different approaches, establishes that given the desire, or in some cases, compulsion, to remain connected, there is a need to understand the complex factors around the co-management of both the virtual and physical interactions. This is, in part, due to the relative ease for individuals to retreat to their phone and "shield oneself from wider surroundings" (Geser, 2006, p. 4), which may unfold as individuals with anxiety use the device to shield themselves from unmanageable situations (Wei and Lo, 2006). Thus, the fact that mobile devices are always connected, and that devices can provide notifications at any point, a situation may become engendered where virtual interactions can potentially rub up against collocated physical interactions.

However, many of these arguments are derived through sociological critique or interviews, opening a gap in the literature for an examination of device interactions from an observational perspective. Such a perspective avoids being caught up in narratives of perception and feeling, instead orienting researchers to the observable matters of device use, to identify specifically what is *practically done* when people use a device in conversation.

2.1.5 *Summary*

This brief introduction should highlight and synthesise the examinations of technology use, especially from the perspective of the 'impact' of the use of technology in everyday life. Crucially, this work draws upon reflections, interviews, and sentiments about how technology

is used. While this thesis adamantly does not dismiss the validity of such findings—indeed they tell us a lot about the world we inhabit—much of it does not examine how device use *practically* unfolds, as *situated action* (L. Suchman, 1985), and in doing so loses its "utility in design" that ethnography, and specifically ethnography informed by ethnomethodology, proffers (Crabtree et al., 2009, pp. 879–880).

In the context of this thesis, then, it is pertinent to consider that as individuals gather to socialise, device use can impact an individual's orientation to space and other co-inhabitants. This thesis does not adopt a moral or sociological critique of this use but draws upon this multi-faceted argument to consider the rudimentary notion of how such interactions practically unfold (i.e. *what do people actually do with their devices?*). As opposed to attempting to gloss the use of mobile devices to generalisable problems or benefits, or relying upon reflections and interviews to guide understanding, this work is interested in explicating the unfolding and nuanced nature of this practice, as it happens *in vivo*, and how it can be used for design. This thesis will add to this existing debate by addressing a gap in literature through the provision of *empirical data* that reveals the interactional work people do to use devices in and through interaction. The next section introduces academic and design efforts to build technologies that support existing collaborative efforts as well as supporting new collaborations using portable technologies.

2.2 SYSTEMS DESIGN FOR COLLOCATED INTERACTION

This second tranche of literature introduces the long-standing focus of designing systems that allow multiple co-present users to work together under the label of 'groupware', originally primarily addressed in CSCW. More recently this work has moved out of the office settings and become concerned with other public and private settings, and examining both sedentary and *mobile* device use. With this, the work now spans the overlapping disciplines of HCI and CSCW (as does this

thesis). HCI is perhaps best described as an "eclectic interdiscipline" that was initially primarily concerned with notions of the *user experience* of technology, although now encompasses "all aspect of human life, from birth to bereavement, through all manner of computing, from device ecologies to nano-technology" (Y. Rogers, 2012, p. vii).

Research in this area has many interchangeable labels, such as collocated (Lucero et al., 2013), co-located (Jarusriboonchai et al., 2014), or co-present (Cole and Stanton, 2003) interaction, and same-time same-place (Fischer et al., 2016) research, although here the first of this list is selected for consistency with publications supported by the work in this thesis. Essentially, in each case, the premise is identical in that work examines the use of technology in situations where there are two or more people present.

2.2.1 *Groupware*

Early work in CSCW examined how to develop groupware systems (Ellis et al., 1991) and other systems to support multiple co-present users interacting with technology and working together, often also referred to as *collaborative software*. Primarily, technologies were designed for workplace settings such as meeting rooms that were already sites for collaborative action. Groupware as a term covers a range of software—from supporting collocated interaction and meeting facilitation, e.g. slideware (Chattopadhyay et al., 2018), through to technologies to support distributed working, such as email-based technologies. The motivation behind the development of groupware systems was not just the increasing availability of technology during the 80s and 90s, but also the potential benefits of augmenting existing practices with technology. For example, systems designed to support decision making were found to *increase* decision quality and equality of participation (Olson et al., 1993).

Progressively, the settings which collaborative systems were designed for moved "out of the meeting room" (Bergqvist et al., 1999) to

classrooms (Abowd, 1999), museums (Ciolfi and Bannon, 2003), public spaces (Reeves, 2011), air traffic control (Hurter et al., 2012), and the home (Edwards and Grinter, 2001; Crabtree and Tolmie, 2016). In addition to new spaces, new technologies became the focus of new interrelated fields, such as tabletop interactions and interactive surfaces (Gjerlufsen et al., 2011; Jones et al., 2012) through to mobile devices (Bellotti and Bly, 1996), such as smartphones and tablets (Lucero et al., 2012). It is the latter of this list that the next section focuses on—this field of work, sometimes referred to as *mobile collocated interactions* or part of the field of *Mobile HCI*, is occupied with designing technologies for interactions with personal devices while we are collocated with each other, i.e. the situation for which this thesis seeks to understand the social organisation.

2.2.2 *Mobile collocated interactions*

This section introduces work from the domain of 'Mobile HCI'. There is a duality to the definition of mobile in the sense of 'Mobile HCI', in that it refers to interaction and physical mobility; or the design and use of mobile devices and mobile device applications, i.e. *portable* devices but in potentially sedentary settings (Church and Oliver, 2011)[5]. The work in this section focuses on this latter definition, considering literature that explores interaction in *collocated settings*, i.e. when multiple people are physically collocated together in the same setting, and often sedentary, but use portable technologies such as smartphones and tablets.

Mobile HCI literature is replete with use cases of collocated mobile device interactions, spun out of academic-led design work, such as photo sharing (Counts and Fellheimer, 2004; Durrant et al., 2011), video watching (O'Hara et al., 2007), and collaborative searching tasks

5 Other work also adopts differing definitions of *mobility*, for example, Luff and Heath (1998) use the term in relation to the micro-mobility of artefacts in face-to-face interaction.

(Church et al., 2012; Cole and Stanton, 2003; Brown et al., 2015), and often involving interaction with additional screens or multiple mobile devices (Bergstrom-Lehtovirta et al., 2013; Lucero et al., 2013). This literature implicitly attempts to demonstrate the beneficial uses of technology in collocated interactions, which expediently refutes—or at least qualifies—simplistic popular views that mobile devices create 'social isolation' (Turkle, 2011). The growing body of literature is exemplified through the generation of design frameworks for the curation of collaborative collocated experiences with technology (e.g. Lundgren et al. (2015)). In this, mobile devices are examined as artefacts that can be brought into everyday cooperative interactions, with the work of designers transforming their "features and functionalities [...] into resources for action" (Salovaara, 2007, p. 1117).

Much of the work within collocated interactions literature, as a subset of existing Mobile HCI research, has challenged the *single-user* nature of *personal* devices, to provocatively explore how the use of mobile devices could instead be designed as *shared* devices that support *multi-user* interaction. This is best elaborated in an *interactions* article on *'Mobile Collocated Interactions'*:

> When using their mobile phones, people have a tendency to hold their devices with one or two hands, with the screen facing toward them. People will usually adopt a particular device position, combined even with a second hand to cover the screen, either to browse private content, such as a confidential email, or to avoid glare [...] For people to fully benefit from mobile collocated interactions, they must open up and start seeing their personal devices as shared, public devices. In mobile collocated interactions, phones are at the intersection of fully personal and fully shared use.
>
> — Lucero et al. (2013, p. 28)

Many examples of this work, which seems to hark back to elements of Weiser (1991)'s vision[6] (see 2.1.1), operate by adopting existing *implications for design* from previous studies to guide the design work of a new prototype. The prototype developed within mobile collocated interactions research is typically tested with 'real-world' users in a field trial (e.g. Lucero et al. (2012)), and the resulting analysis used to generate new implications for design. Some work also specifically follows routes of creating provocative prototypes, not to solve users' problems, but to find ways of evoking critical reflection by users (Redström, 2006). For example, Lundgren and Torgersson (2013) explore ideas to "design interventions that investigate how apps for mobile devices can make people interact directly in co-located space instead of enclosing themselves with their own digital device" (ibid., p. 1), primarily using the idea of games to prompt users to reflect and consider interacting with collocated others.

In the same spirit of creating enjoyable interfaces. but with the idea of using the device as a "resource" for conversation, Porcheron et al. (2016b, p. 232) explored the idea of allowing individuals to collect and share photos and notes for a design project with their mobile phone. The app allowed users to collocate and share the collections but fundamentally required the group to converse face-to-face with each other to use the application successfully and navigate its features. This work, along with others such as Lucero et al. (2012) that allowed for large display and cross-device interactions in public settings from mobile phones, provides a semblance of supporting the notion that device interactions can be curated, can be enjoyable (Brown and Juhlin, 2015), and can be used to enhance people's experiences of space (e.g. as in Bull (2005)) and conversation with each other (e.g. Lundgren and Torgersson (2013)).

6 Indeed, the notion of devices switching from 'personal' to 'shared' use is an embodiment of this vision.

2.2.3 *Designing the context-aware device*

Related to this, the notion of designing for space and the context within which the device is used is explored elsewhere, with a considerable body of work that looks to make device interactions more sensitive to their environment and context of the world around the device. For many years, a holy grail of Mobile HCI research was the "context-aware" (Schilit et al., 1994) smartphone, with many designs explored in literature; one such influential example is 'ContextPhone' by Raento et al. (2005) although there are many more (e.g. Siewiorek et al. (2003) and Gellersen et al. (2002)).

The purpose of this research is to establish ways of making devices sensitive to the environment, with Gellersen et al. (2002) drawing on Weiser (1991)'s vision to note that "[i]n the mobile device user interface, context can be used to facilitate a shift from explicit user-driven to implicit context-driven interaction" (Gellersen et al., 2002, p. 341). In the frame of the 'contextually aware device', the view of what *context* is differs perhaps from the sociological or interactional definitions, and is broadly classed as:

> [...] any information that can be used to characterize the situation of an entity. An entity is a person, place or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves, and by extension, the environment the user and applications are embedded in.
>
> — Dey and Häkkilä (2008, p. 217)

The ambitions to create devices that are sensitive to anything that is 'relevant to the interaction' have generated numerous design and research challenges. Many of these ideas draw on the notion that context can be used to ameliorate problematic interactions with devices, and are spurred on by the critique that devices are invasive in everyday life (see 2.1). One of the practical ways in which notions of

"contextual awareness" have been realised in research is through notification sensitivity to the environment, i.e. a device will determine when to deliver a notification at the "opportune moment" (Fischer et al., 2011, pp. 57–64). Work in search of the contextual device has progressed, and has considered and explored design ideas for reducing the impact of notifications on many different portable devices such as smartphones (e.g. Okoshi et al. (2016)), smartglasses (e.g. Kern and Schiele (2003) and Lucero and Vetek (2014)), and smartwatches (e.g. Lee et al. (2016)) by withholding notifications until a later time.

Furthermore, ideas also surround the use of the 'continuous speech stream' in design to detect the context. In other words, devices which listen to the stream of speech around them would make use of a continuous and live transcription of conversation to prepare or sensitise interactions to the context (McMillan et al., 2015). In a prototypical trial of this, Schulze and Groh (2016) found their system to be on-the-whole suitable, but then attach caveats that "if conversational context is employed to determine interruptibility, the prior characterization of the conversation is essential" and "concerns of participants that go beyond our measures to preserve privacy and beyond a lack of their being informed about those measures, can't be addressed by design" (ibid., p. 9), exemplifying that the challenges of the context-aware smartphone are socio-legal as well as technical.

However, in spite of this progress addressing the challenge of devices that tailor their interactions to our setting, manufacturers have been slow to adopt such features. Some smartphones feature options that enable a silent mode when they are placed face down; however, typically devices defer to the user to configure settings for notification management. The result of this is that the decision of when to deal with interruptions from a device is left to the user.

2.2.4 *Summary*

This section has established the tradition of designing systems to support co-operative working within CSCW and (Mobile) HCI. Progressively, the focus of development expanded from meeting rooms through to other settings and technologies, including mobile settings *and* mobile devices. The literature on creating collaborative experiences with mobile devices while we are collocated with others was summarised. Then literature was introduced that examines the parallel efforts to ameliorate the problematised nature of device-triggered interruptions to create mobile interfaces that are sensitive to 'contextual' information. Although notifications play a big part of 'life with devices', it is expected that they form only part of the occasioning of devices in social settings. This is reinforced by Sohn et al. (2008) and Church and Smyth (2008), whom both used diary studies to reveal other many reasons for the use of the Internet on a portable device beyond mere device-instigated notifications. The next section will pivot from matters of designing technologies to studies of how mobile devices are used in everyday life, from a perspective focused on close studies of face-to-face interaction.

2.3 CLOSE STUDIES OF COLLOCATED INTERACTION

This final section will synthesise the literature that examines interaction amongst people while they are face-to-face, especially from the HCI and CSCW domains. These are typically 'close' studies of interaction, i.e. they take an approach that orients to observable matters of technology use to address specifically what is practically done when people use of technology 'in the wild', with empirical data being the commodity that establishes the findings of the work. The first section will discuss the *turn to the social* in HCI.

2.3.1 *Turn to the social in HCI*

L. Suchman (1985)'s work represents some of the earliest and oft-cited work in HCI on the study of the social organisation of interaction with technology in this vein. Her work, undertaken in a lab-based setting, examined the use of agent-based photocopiers, adopting ethnographic approaches and drawing on the ethnomethodological perspective (this is unpacked later in Chapter 3). This work was undertaken in parallel to other key works in CSCW studies around the same period which examined settings such as air-traffic control rooms (Bentley et al., 1992), drawing upon similar perspectives, to unpack the cooperative working practices in order to support design activities (Bannon et al., 1993). The domains of CSCW and HCI were increasingly focused on studies of control rooms; in a range of settings from Heath and Luff (1992) in a London Underground control room, C. Goodwin and M. H. Goodwin (1996) in air traffic control rooms, and Watts et al. (1996) in NASA mission control rooms. L. Suchman (1997) characterised such settings as "centers of coordination", in which "participants' ongoing orientation to problems of space and time" is crucial in attending to matters of deployment of people and equipment in response to a planned timetable or emergent requirement (ibid., pp. 41–43).

As discussed above in relation to design efforts shifting out of meeting rooms, ethnographic studies too shifted "out of the control room" and other constrained settings (J. Hughes et al., 1994), as part of a "turn to the social" in HCI and CSCW. This turn was "a primary point of view for analysing the design space under the auspices of groupware and cooperative systems" (Crabtree, 2003, p. 28). In this regard, the practice of studying interaction in constrained settings became re-purposed for examining matters of "everyday life". The next section introduces literature that details studies of face-to-face interaction and technology use. Suchman's work, and that of the Lancaster CSCW studies, transformed the ethnographic practices, and the use of

the ethnomethodological perspective, into utilities for design in HCI and CSCW. This transformation of studies for design did so to support adoption of an "analytic orientation to fieldwork, which seeks to uncover the locally organized character of action and interaction [...] [that is] is essential to the ongoing development of computing systems that resonate with, support, and enhance what people actually do in new design contexts and how they organize what they do" (Crabtree et al., 2009, p. 881). The next section will synthesise and discuss the literature from HCI and related disciplines on the use of technology while we are collocated with others.

2.3.2 *Using technology while collocated*

Early observational studies of portable device use[7] examined aspects such as how devices were shared amongst groups of friends. For example, Weilenmann and Larsson (2002) discuss a study of young people in Sweden making use of phones, with observations collected anonymously in public spaces focusing on how sharing is done between collocated friends. Elsewhere, Murtagh (2002) describes the grossly observable features of mobile phone use during observations of people on train carriages. Krehl et al. (2013) followed travellers on public transport journeys and categorised their mobile device use. This categorisation and resulting model was based on contextual factors relating to the use of devices—such as location, task, and technical details. Such findings provide a comprehensive insight into the activities of device users while mobile, again, providing rich, actionable foundations for work to design a contextually sensitive device (see 2.2.3). Conversely, through the collation of existing literature, Nakamura (2015) presents a model of the actions people employ in looking at mobile phone displays in everyday life as non-verbal communication,

7 Note that these studies were conducted in the early 2000s before the widespread commercial availability of smartphones, and thus they are studies of what may now be considered 'traditional' mobile phones, or cellphones, or non-smartphones.

bringing to the fore some nuanced and complex issues around mobile device use in public life. This work positions the "phone user behind something akin to the 'fourth wall"' (Nakamura, 2015, p. 74), which while beneficial for making sense of generating generalisable models of interaction, also does the work of detracting from the situated and contextually-shaped (and contextually-shaping) nature that interaction with a mobile phone occasions and is occasioned by (discussed later in 3.2.2).

There is also a growing body of work that reveals the intricate ways in which device use becomes interleaved in a variety of settings, adopting methodological approaches that examine interaction moment-by-moment, ranging from kitchens (DiDomenico and Boase, 2013) and living rooms (Rooksby et al., 2015) through to collaborative photo-taking activities (Fischer et al., 2013). On people using mobile devices while watching television, Rooksby et al. (2015) remark that "[w]e should not view the mobile device as being brought into television viewing, but the use of mobile devices, the watching of television, and so on as things being brought into leisure" (ibid., p. 17). The sense here is that the use of mobile devices is part of leisure activities in the home, rather than an isolatable activity. In another study that recorded and studied participants use of smartphones over an extended period, the authors identified how 25% of all portable device interactions took place while participants were co-present with others (Brown et al., 2014), illuminating the notion of how the use of devices routinely occurs while we are around others. Through the use of screen-recording and fitting participants with portable "camera bags", Brown et al. (2013) provide empirical and rich insight into the daily use of smartphone users. The work also goes on to introduce sequential practices of how devices are occasioned through activities such as route finding and how web searches are initiated: by others in talk, by events taking place, and by features of the local environment. The research also brings to the fore how people make use of multi-modal features (e.g. gaze, orientation) to use the device in

conversation with others successfully. Pizza et al. (2016) follow a similar modus operandi to capture the use of smartwatches, revealing what smartwatches are used for and that people embed the interaction with the smartwatch in conversation, which they demonstrate through detailing "users' everyday activity [...] [in one situation] the participant is talking about last night, arranging some ingredients for cooking, and quickly reading a notification on the watch" (ibid., p. 5464). This attention to the minutiae of interaction provides a detailed and rich insight into the interactional accomplishments of people, as they embed device use in other ongoing activities.

Furthermore, as Isaacs et al. (2012) remark, "people attempt to blend their local and remote worlds into coherent interaction when sharing content and one experiences with friends through their devices [and that these] behaviors are becoming more varied, and possibly more common, because of the prevalence of ubiquitous devices and bite-sized content" (ibid., p. 625). Work in HCI and CSCW establishes the case that device use becomes embedded within and is treated as part of activities in everyday life, from television watching (Rooksby et al., 2015) through to searching the web, as occasioned in and through conversation with co-present others (Brown et al., 2015). These practices include people embedding the device in interaction through collaborating on web searches (ibid.) or making the screen visible to others during use (Raclaw et al., 2016).

Overall, these studies took place by observing people 'in the wild' naturally (i.e. by orienting to the naturally accountable methods in which members socially organise their interaction), with the studies taking place in perspicuous settings (Garfinkel, 2002a, pp. 181–182) to the technology being considered. This is not to say that such settings could not be constructed as part of the research, for example as laboratory settings (Rooksby, 2013), and it is the defence of such studies by Rooksby (ibid.) that is crucial to this thesis: "[i]f there is to be a "turn to the wild" in HCI, this should not be a turn away from the laboratory but a turn away from research methods that ignore human

practice" (Rooksby, 2013, p. 1). In other words, the *turn to the social* is not, per sé, a preference for conducting studies outside of laboratory settings, but ones in which the social organisation of the setting and the members within that is crucial to design (Crabtree et al., 2009), is the locus of the researcher's concern. Any study that seeks to examine interaction and technology use for design must attend to the matters of the social organisation of the setting as such details "[...] matter and not as analytic phenomena" but as matters which the members of the setting understand and must be designed for (Crabtree et al., 2012, p. 137). This thesis takes a hybrid approach, insomuch that the participants are recruited for research and asked to attend a social setting, but the setting is one which they would typically visit together for a research study, and is of perspicuity for the technology interaction of concern. Participants were made aware that interaction around device use was being studied in each case, with the second study asking participants to preferentially use the VUI on their smartphone instead of the touchscreen and the third study including the provision of the technology as part of a deployment to participants' homes. The focus of the work in this thesis is *the naturally accountable ways in which members attend to device use as it unfolds* (i.e. to reveal *how* it unfolds), and not to establish *why* device use unfolded.

Indeed, most studies of voice-based interfaces have, partly for technical reasons, consisted of Wizard of Oz studies in which the 'computer' or Voice User Interface (VUI) has been driven covertly by a human rather than a computer agent, in effect introducing a simulated element to the interaction[8]. Such practices are long-standing with VUI-design related work and studies, with early research on simulating VUIs used to demonstrate or provide software to implement such studies (Klemmer et al., 2000) or how to undertake a conversation analytic approach to analysing such interactions (Fraser and

8 This is not to say that such an approach would impede a study of attending to the social organisation of device use; indeed, it is argued as much on page 34. This is merely to remark that there are nascent studies of commercially available VUIs.

Gilbert, 1991; Wooffitt, 1994; Wooffitt et al., 1997). Specific examples also include studies such as how VUIs can be sensitively designed for use in safety-critical situations such as during driving (Martelaro and Ju, 2017), for use in mixed-reality games (Dow et al., 2005), and action games (Höysniemi et al., 2004). These authors all show how members make use of devices precisely, alter their voice for the device (Pelikan and Broth, 2016), and manage the interaction so that it unfolds when opportune (Martelaro and Ju, 2017). Furthermore, each one adopts the perspective of analysing situated action, to reveal details of how users attend to the devices in a particular context, eschewing notions of a generalisable model of interaction. These studies demonstrate the suitability of adopting observational perspectives to studying device use in order to reveal nuanced interactional practices, but also demonstrate the relevance of studying device use in a particular context—it is not possible to transgress findings from studies across different contexts. Given the recent commercial availability of VUIs on smartphones and standalone VUI devices, the necessity to study interactions with them through the Wizard of Oz technique is mitigated, however, the work in the latter two chapters remains among the first to study the social organisation of the use of actual systems in the wild.

2.3.3 *Summary*

This section has introduced literature that has studied the lived experience of device use through observational studies. This work, in juxtaposition with the design work in mobile collocated interactions (see 2.2.2) that intentionally prototypes designs for precisely these *sorts* of settings, demonstrates a shortcoming in literature. This section has brought together literature that has examined collocated interaction while mobile and sedentary, with different devices such as smartphones and smartwatches. However, at the time the work was conducted, there was no study of how device use unfolds in social settings such as public places like a pub, which bring with them 'cer-

tain ways of behaving' given the specificity of the setting (discussed later in 4.2.1). Furthermore, there remains nascent literature on the use of smartphone-based VUIs and smartspeakers given the recency with which they were introduced, with most studies to date consisting of Wizard of Oz-based studies. The empirical work in this thesis will address these gaps.

The turn to the social in HCI centred around the import of ethnographic practices and the ethnomethodological perspective. This expansion of ethnographic practices into HCI, in turn, allows research to support design efforts by generating significant implications for design, which as Crabtree et al. (2012) remind us, allows designers to ground decisions in 'facts' (ibid., p. 137). Although it is not the purpose of this thesis to design new technologies, through the documentation of the social organisation of conversation with technology interleaved, a greater understanding of how future technologies could meet the needs of users could be designed.

3

APPROACH

This chapter brings to the fore and discusses the conceptual side of the methodological approach adopted in this thesis to collecting and analysing empirical data of members' actions in the three settings studied. This chapter will discuss how this thesis' approach develops an understanding of the interactional accomplishment of conversing and socialising while interleaving the use of a device. This chapter will introduce and situate ethnography as a method of academic inquiry, used across multiple disciplines, and of its relevance and propriety for the studies in this thesis. Ethnomethodology "is the study of the methods people use for producing recognizable social orders [...] to discover the things that persons in particular situations do, the methods they use, to create the patterned orderliness of social life" (Garfinkel, 2002b, p. 6). This chapter will establish this thesis' specific form of ethnography—ethnomethodology—and how the analysis is informed by the ethnomethodological perspective, through which an understanding of how those conversing interleave device interactions with talk.

Of the three studies in this thesis, the first two draw upon video-supported ethnography to aid explication of what is done by members of the settings as they use a device during a gathering in a semi-public setting (see Chapters 4 and 5). The third study continues to draw upon the same analytic perspective, although consists of automatically captured audio data from interactions with a device in the home (see Chapter 6). The analytic perspective remains the same across all three studies, however, and this chapter helps frame this perspective but will refrain from discussing the practical matters of how each study was conducted 'in-the-wild', in part because they vary across each study; such details are included in each empirical

chapter, so as to allow for the explanation of how data were collected and analysed in the context of the specific technology. In this thesis, the analysis was not undertaken as a distinct post hoc event but instead occurred throughout the fieldwork-based studies—a practice that this chapter will establish as core to ethnographic work. Therefore, it becomes all the more relevant to understand Ethnomethodology not just as a tool for making sense of the data collected and presented, but to understand the practical aspects through which these data were collected and selected for presentation in this thesis.

3.1 ETHNOGRAPHY

This thesis adopts an interactionist perspective to understanding everyday interaction, drawing upon ethnographic practices and eschewing other methods of work, such as *a posteriori* interviews or self-report methods. *Ethnography*—an approach to the study of social life developed by Malinowski (1922)—takes many forms and is now found in many different disciplines from sociology through to computer science. Practically, however, the perspective through which the ethnography is conceptualised and established varies. There are many different approaches and perspectives under the banner of *ethnography*, this thesis adopts the perspective of ethnomethodology (Garfinkel, 1967). In this chapter, the *ethnomethodological* perspective will be introduced, and it is through this analytical lens that the primary objective of this thesis—to reveal *how* individuals conversing with others interleave the use of a device in conversation—will be met. First, however, a summary of the development of the ethnographic approach is included here to situate and rationalise the adoption of the ethnomethodological approach to ethnography.

Malinowski conducted an ethnographic study of the native inhabitants of Guinea and exposed their practices to others through his influential work *Argonauts of the Western Pacific*. Through this, he influenced the anthropological study of communities and settings by

shifting such studies from efforts that focused on the collection of artefacts, stories, and measurements into a pursuit of *immersion* in the setting; to "grasp the native's point of view [...] to realise his vision of his world" (Malinowski, 1922, p. 25). His work, through circumstance, was achieved not through mere observation of the inhabitants or through a retrospective analysis of collected artefacts, but by immersing himself in the culture of those he was studying, to *experience what they experienced*. Through his study, Malinowski established ethnography as a prominent scientific method which relies upon not just observation or fieldwork taking place, but one in which the ethnography itself chronicles the *behaviour* of those under study, and not just the tools they used or a description of the environment in which they were used:

> In Ethnography, where a candid account of such data [a detailed account of all the arrangements of the experiments; an exact description of the apparatus used; of the manner in which the observations were conducted] is [perhaps even more] necessary, it has unfortunately in the past not always been supplied with sufficient generosity, and many writers do not ply the full searchlight of methodic sincerity, as they move among their facts but produce them before us out of complete obscurity. [...] In ethnography, the writer is his own chronicler and the historian at the same time, while his sources are no doubt easily accessible, but also supremely elusive and complex; they are not embodied in fixed, material documents, but in the behaviour and in the memory of living men.
>
> — Malinowski (ibid., pp. 3–5)

In this sense, Malinowski established that ethnography is not just observation or the analysis of a corpus of exhibits (Bittner, 1973), but more so requires the ethnographer to understand the phenomena under study as if one were a member, and could see and make sense of the work in the setting from the members' point of view. Therefore,

ethnography is *not just* fieldwork but also the resulting analysis (Anderson, 1997), in which a demonstrably useful account of the lives of others is established (Button, 2000; Crabtree et al., 2012). His work required immersion, time, and devotion to understanding language and practices, and a willingness to partake in the society as if he were a native. Malinowski's work posited this, and demonstrably revealed how it was necessary to make sense and understand the perspective of the 'natives' and chronicle their lived experience.

Of course, Malinowski's study was patently different to the work of this thesis: here the people under study are not in a remote society with which there was nascent knowledge in the Western canon. Ethnography, as understood in the perspective of this thesis, was first pioneered by the Chicago School of Sociology, which took the matter of *everyday life* to be its locus of study—not "non-western societies and cultures" but instead "the city as its subject matter, and through numerous extensive and detailed ethnographic examinations of urban life subjected the city to an order of examination previously reserved for 'other' societies and cultures" (Button et al., 2015, p. 112). In this, one such pioneer, E. C. Hughes (1958), spurred ethnography from a study of another's culture to a study of one's own society, asking his students to study their taxi rides, cleaners, and so forth. However, one critique of the work from this era was that such ethnographies failed to explicate the 'interactional work' (this is discussed later in 3.2.1) of members of the setting; i.e. they failed to reveal the social phenomena but relied upon 'scenic' features of action and in doing so, failed to allow the reader to understand the 'work' of the setting (Crabtree et al., 2009). With this, it becomes evident that it is not enough to be able to speak the language, or to be vaguely familiar with 'what is done'. Conducting an ethnography—even if you 'know the language'—still presents many challenges, not least access to the setting, which may prove challenging as many settings are not readily observable by the public. With this challenge comes issues of securing the consent of participants to collect data and the development of the researcher's

competence to make sense of the work under investigation (Crabtree et al., 2012, pp. 89–95).

Bringing the influences of the Malinowski and Chicago School together: ethnography, which can involve the study of things that 'might seem familiar' or 'common sense' (ibid., p. 160), requires an ethnographer to embed themselves in the setting to understand the experience of the members of the setting from their perspective. Moreover, the production of an ethnography is *not only* fieldwork but consists of fieldwork and analysis in unison, in which the fieldwork is guided by the analytic perspective to make sense of members' actions. Critique of the work of the Chicago School's influence on ethnography varied, but the next section of this chapter expands upon one element of this critique—of a reliance upon scenic descriptions—and introduces how the tradition of ethnomethodology, and its orientation and stance to ethnographic work, can furnish readers with a richer ethnographic record of members' methodical accomplishments.

3.2 THE ETHNOMETHODOLOGICAL PERSPECTIVE

This thesis adopts the perspective of ethnomethodology in its approach to ethnography. A study that is ethnomethodological in character focuses on the ongoing ordinary primordially social features of everyday interaction (Schegloff, 1987); in other words, an ethnomethodological study reveals "the techniques and strategies members of society use in making sense of one-another's subjective perspective on everyday experience, and through these methods, achieving a significant measure of shared understanding" (Reeves, 2011, p. 30). It is through this orientation to everyday routine practices of individuals that ethnomethodology develops its concern with the accountable ways in which members organise their conduct, moment-by-moment, relevant to their context. There are several key tenets to understanding the ethnomethodological perspective that are introduced here: the explication of the *work* and *interactional what*, the sequentiality of ac-

tion, the policy of *indifference*, and the notion of *vulgar competence*. Each point is addressed in turn throughout the remainder of this chapter, and through these points the analytic lens in which each study in this thesis was conducted is assembled.

3.2.1 *'Work' and 'interactional what'*

In this thesis, interaction is treated as the locus of study, with accounts of what is done by the members of the setting being the primary resource used to establish the ethnographic record. This record consists of *thick descriptions* (Geertz, 1973) that unpack and reveal the interactional *naturally accountable* methods of members (i.e. *the accountable character* of work in the setting). Ethnomethodology, it is argued, suspends the assumption that social "order is a rare beast to be found in only a few places" (Crabtree et al., 2013, p. 6) but is instead a constituent feature of the ordinary activities and common-sense reasoning that inhabits and animates it, and it is this that an ethnomethodological lens illuminates. In this regard, the ethnographic record, then, will reveal the social order of the phenomena under study, or in other words, will allow readers to make sense of the actions of members as they converse and interleave device use within this conversation. The thick description, which will be assembled as a result of the fieldwork and analysis, is produced through attention to "to what is done in the doing of action [through] 'thicken[ing] up' the thinnest level of description to make its accountable character visible and available to others" (Crabtree et al., 2012, pp 117–118). With this, accountability is defined as an action that is observable and reportable (Garfinkel, 1967), i.e. what it is done is observ-*able* and tell-*able* by the other parties who are present (Crabtree et al., 2012, pp 117–118). It is this thick description that documents the accountable work of members of the setting. *Work*, in the sense of ethnomethodology, is not treated as equivalent to paid labour but is considered the achievement of mundane naturally occurring activities (Schmidt,

2011; Crabtree et al., 2009; Crabtree et al., 2006; Button, 2012), with Sacks succinctly regarding it as a culmination of 'everyday stuff that is done' in and through a person living their ordinary routine:

> Whatever we may think about what it is to be an ordinary person in the world, an initial shift is not to think of an 'ordinary person' as some person, but as somebody having as their job [...] doing 'being ordinary'. It's not that somebody is ordinary [...] it takes work, as any other business.
>
> — Sacks (1992b, pp. 215–221)
> edited by Crabtree et al. (2012, pp. 23–24)

Although there have been many ethnographic studies of 'ordinary activities', the ethnomethodological orientation to ethnography also embellishes qualitative participant-observation approaches with attendance to revealing the "interactional what" (Garfinkel, n.d.). This notion of 'what' was developed in a commentary by David Sudnow and Garfinkel in response to studies of Jazz singers by Howard Becker, whose work featured heavily in the development of ethnography in the Chicago School of Sociology. Becker, amongst his work, provided accounts for "the career structure of the jazz musician, the fraternal organisation of work it gave rise to, the pressures of work and playing to the audience, the dilemma of commercialism versus prestige, and the impact of family on the musician's life and the conflict it generates" (Button et al., 2015, pp. 116–117), yet Sudnow and Garfinkel argue he did not reveal the circumstances in which music was collaboratively accomplished—his work was informative, and well developed, but did not reveal the *interactional work* of a Jazz musician. These studies, although ethnographic, were found to merely provide "scenic descriptions" of what is done and were of limited use in understanding how interaction was specifically achieved as a situated and coordinated action (the notion of *situated action* is elaborated upon later in 3.2.2). The tradition of ethnomethodology is not only crucial to understanding and making available what is done as a gloss, but principally what is done *in interaction* (i.e. the interactional work of Jazz

musicians as they did it, from their perspective). In assembling the ethnographic record of what is done—to unpack this gloss—it becomes necessary to understand and relate the actions of members as a series of particular activities. As Crabtree et al. (2012) elaborate:

> We need to be able to see the activities that produce sequential order in the 'lived' details of their production – i.e., in details of the particular things that members do to accomplish the component activities of a sequence.
>
> — Crabtree et al. (ibid., pp. 103–106)

With this perspective it becomes evident that an ethnographic record informed by ethnomethodology should reveal the activities of members through some form of sequential order, to allow the reader to understand the member's perspective and actions, and of how such actions are constituent in the sequential ordering of an activity. In this, the actions of members become assembled as a series of *sequential accomplishments*, to thicken scenic descriptions so that they allow the reader to make sense of the members' practical reasoning and practical action. It is this notion of *sequentiality* and the situated-*ness* of action that the next section details.

3.2.2 *Sequential organisation of situated action*

Firstly, *sequentiality* is defined as "any kind of organization which concerns the relative positioning of utterances or actions [...] turn-taking [in conversation] is a type of sequential organization because it concerns the relative ordering of speakers" (Schegloff, 2007, pp. 1–3). With this definition, it is important to note that sequentiality differs from mere temporal ordering (although it can take advantage of it), not only in that it encompasses actions that occur temporally in tandem (such as overlapped talk), but that the sequential coherence of conversation is a continuous achievement by conversationalists, who are seeking to assemble the retrospective-prospective sense of those actions which are often outside a basic temporal order. For instance, a

speaker might answer a question several turns subsequent to it being posed in a conversation (which might be accounted for by a speaker in various ways, e.g. prefacing "before I answer your question…" to their turn). The notion of retrospective-prospective is key here, as Garfinkel notes:

> Many expressions are such that their sense cannot be decided unless one knows or assumes something about the biography and the purposes of the speaker, the circumstances of the utterance, the previous course of the conversation, or the particular relationship of actual or potential interaction that exists between speakers. The sensible character of an expression requires that we wait for what a speaker or speakers say next for the present significance of what has already been said to be clarified. Thus, many expressions have the property of being progressively realised and realisable through the further course of the conversation.
>
> — Garfinkel (1967, pp. 35–75)
> edited by Crabtree et al. (2012, pp. 122–123)

With this, the case is established that action is both context-shaped, in that to understand it one must know the context within which it unfolded, and also context-shaping, in that each action carries implications for future actions, and is only realised through those future actions. In this, actions become coherently and sequentially *organised*. It is this feature of interaction as being sequentially organised, and further so locally and longitudinally managed by members that provides the basis for Suchman's notion of 'situated action' (i.e. the arrangements of this organisation of action are negotiated and established only in and through their production and the context of the interaction (Button et al., 1995; H. T. Nguyen, 2008)). Suchman's analysis draws upon observation of everyday interaction with an agent-based photocopier at Xerox PARC, and by drawing upon ethnomethodology, she was able to explicate not only issues with the design of the

hardware but also fundamental notions of the mundane achievement of work in using the device. On situated action, she notes that:

> That term underscores the view that every course of action depends in essential ways on its material and social circumstances. Rather than attempt to abstract action away from its circumstances and represent it as a rational plan, the approach is to study how people use their circumstances to achieve intelligent action.
>
> — L. Suchman (1985, p. 35)

Suchman's definition builds in the notion that people are 'everyday sociologists', and that members of settings can observe and recognise what other members of the setting are doing, and that this stems from the natural accountability of members' actions (Berger and Luckmann, 1966). Natural accountability is the notion that the 'members' of a setting can observe the work of others around them in that setting, and crucially, *know* what it is that they and others involved in that work are doing (Garfinkel, 1967, pp. 1–34). With this, members can unproblematically offer an account of what they are observing, and that the other members of the setting will recognise this account (ibid., pp. 1–34). Specifically, members' actions are naturally accountable in terms of their *practical action and practical reasoning* (Garfinkel and Sacks, 1970)—ethnomethodology is not concerned with 'activity', 'action', or 'agency', but with how these notions are "ordinarily understood by the members of society from within the settings in which they operate" (Crabtree et al., 2012, p. 29). As Crabtree et al. (ibid.) remarks: "[t]he naturally accountable character of everyday activities is an achieved outcome of their conduct, which is to say that in making their activities happen—in the work of assembling and accomplishing them—members attend as a matter of course to making them naturally accountable" (ibid., p. 25). Moreover, not only does Suchman's work provide the practical methodological approach for this thesis, but the premise of action as established in and through its achievement as a product of the context within which

it is done, is imperative in bringing the empirical findings into the context of design in this thesis. In other words, Suchman's influential work scopes out a field of work in which research in interaction with systems focuses on the practically and accountably done actions as opposed to theoretical assumptions of action, and so sensitises researchers to the need to include, not abstract, context-shaping and context-shaped implications of action.

In this thesis, the work of members in the setting will be chronicled through the presentation of series of excerpts of data. Members' interactional accomplishments will be analysed with respect to the coherence and situated nature in which they occur, enabling the analysis to reveal the *interactional what* of how device use is interleaved within conversation.

3.2.3 *Ethnomethodological indifference*

Thus far, this section has detailed the ethnomethodological orientation to the interactional work of members in settings, and how this is revealed through an attention to the sequentiality through which their accountable actions are conducted. Through inference, it should also be clear that this thesis is not concerned with theories of 'why' something happened, or indeed theories of interaction or work in general, but rather focuses on the practical situated accomplishment of action. In this sense, the thesis adopts the notion of "ethnomethodological indifference" (Garfinkel and Sacks, 1970). As summarised by Lynch (1993), this consideration allows researchers to pragmatically study the work of people "[r]ather than addressing whether sociologists ever can achieve adequate or acceptable accounts of the phenomena they study" (ibid., p. 141). In other words, what matters in this research is explicating the *members' methods* of interaction without *a priori* models of how such interaction unfolds (Livingston, 1987), i.e. there need not be a theoretical unpinning of understanding in how

people use mobile devices because this work is primarily concerned with the *accountable* interaction of the setting.

Button et al. (2015) argue that an account of the setting imbued with interpretation detracts from the work of the setting, transforming the thick description of members' actions into an interpretation of both members' actions and interpretations. This claim rubs up against others who argue that "indifference" is not necessary to derive a valid account. Dourish (2014) argues that the allowance of researchers to draw upon epistemological notions does not impede analysis, as such analysis is rooted in the researcher's immersion and experience of the setting and that the researcher's account of the work of the setting is not invalidated as such. However, for this thesis, indifference was adopted insomuch that the goal of the work was to 'take a step back' from the critical assessments of device use found in existing literature—both academic and popular press—and instead practically study how such device use is interleaved within conversation, and through this explicate the *members' methods* of how this use is achieved. Applying *a priori* understanding to the analysis would instead pivot this work from an examination of how device use unfolds as an accomplishment and instead project the existing rhetoric of device use upon the analysis, and in turn, diminish the contribution and motivations behind the thesis. Thus, in accordance with practice guided by others, the data collection and analysis in this thesis was based on studying settings and analysing data without *a priori* frameworks of what constitutes interleaving practice, and instead allows such notions of how members' actions unfold to be guided by the data (Crabtree et al., 2012; Heath et al., 2010).

3.2.4 *Vulgar competency*

The final matter to address in this brief summary of the development of ethnomethodology and this thesis' methodological approach is to consider the issue of *vulgar competency*, which is fundamental to how

researchers reliably make sense and present the social order of the setting, i.e. of how members make their actions accountable to each other (Garfinkel and Wieder, 1992a). The notion of competency is the antithesis to the interpretation of findings and is developed through the ethnographer attaining a position in the setting in which they not only understand the routines of the setting—a *gloss*, if you will, of what is done—but how the members accomplish that routine. Button and Sharrock (2009) solidify the necessity of developing competency: "even if it appears to an outsider that nothing is going on, there will be something that is being done" (ibid., p. 86). Slack (2000) further argues that vulgar competency is intrinsically connected to the type of account produced in the analysis: if an ethnographer does not adequately attend to members' everyday work practices through the perspective in which the member lived and undertook them, then those accounts falter and potentially become mere interpretations of phenomena. In other words, if one cannot see it from the members' perspective, then any account cannot be a true reflection of the members' situated action.

Furthermore, the notion of competency stands in unison with that of ethnomethodological indifference: such competency should come from an understanding of the member's methods rather than through the use of formal sociological methods of inspection. Through the development of vulgar competence, it becomes possible for the analyst to understand the 'inner-workings' of the setting and of the members' perspective, allowing the "analytic and member concerns [to] merge [such that] the very distinction between the [. . .] analyst and member is obliterated" (Pollner, 2012, p. 15).

The specific approach to developing competency, as with all ethnographies, depends upon the setting under study. This thesis, perhaps more so than others that adopt an ethnomethodological perspective, studies an activity which intrinsically motivated the thesis: the initial ideas for this thesis came from remarks by others of the author's use of a device while they were socialising together in a pub. This candid

observation from a friend was transformed into a research proposal that spawned a thesis that examined just how this device use unfolds in social settings. This thesis studies settings in which the researcher was already a part of and had a competency: i.e. people socialising together and who ostensibly use technology in such gatherings.

Part II

EMPIRICAL WORK

4

STUDYING PUB TALK AROUND SMARTPHONE USE

This chapter presents a study of people using their personal mobile devices while they are socialising together with friends in a local pub. This work was undertaken to unpack the interactional accomplishment of device use in and through an ongoing face-to-face conversation of three or more people in a 'casual-social' setting. The study was naturalistic—participants were recruited as groups of friends with the intent of being accompanied by a researcher and video-recorded during a social gathering in a local pub. They were not at any point asked to use their mobile device, and the use of their devices that ensued, all of which drew upon the devices' Graphical User Interface (GUI) touchscreen, was entirely coincidental and arose out of external factors or the unfolding conversation between the friends.

This chapter was previously published and presented at the Computer-Supported Cooperative Work & Social Computing conference[1]—a number of changes have been made to ensure this chapter addresses the research questions of this thesis.

4.1 INTRODUCTION

Pubs are sites for jovial interaction amongst friends and strangers (Fox, 1996), are described as "focal stages of sociability" (Törrönen and Maunu, 2005, p. 25), and provide a casual and social setting in which people can relax with friends while drinking, talking, watching sports, and so on. The very nature of socialising in a pub imbues an informality to the interaction amongst those present, commonly referred to as patrons. Through the collection and analysis of ethnographic data

1 See Porcheron et al. (2016a).

of groups of friends socialising together, this chapter will unpack the ways in which the use of mobile devices such as smartphones is accomplished as a mundane feature of conversation in a pub (referred to as pub talk herein). It is in this sense that pubs are established as 'casual-social settings', and from existing literature, are identified as places where mobile device use already naturally occurs. Therefore, their selection as a site for this ethnographic study will be demonstrated to be a suitable and adequate setting for unpacking how devices are used to address members' problems in interaction.

There have been numerous studies into the use of mobile devices by individuals in situations where the user is collocated with others (as discussed previously in Chapter 2). Mainly, these follow in the lines of 'reductionist' approaches of HCI to identify 'causes' and 'solutions' to problems, in contrast to the holistic ethnographic approach adopted within this thesis. These studies often ascribe device use to different factors such as *boredom* (Pielot et al., 2015), *habit* (Oulasvirta et al., 2011), or *interruptions* or *notifications* originating from the mobile device (C. Park et al., 2017). Su and Wang (2015) present findings from a study based on observations of friends using mobile devices in pubs, through which they remark upon how mobile devices "alter[...] the prime activity" (ibid., p. 1659) of places such as pubs because of new activities taking place; by this they specifically orient to the use of the devices by pub patrons. Furthermore, work within academia has explored numerous psychological and emotional factors relating to what leads to device use (Kushlev et al., 2016) and how receptive people are to interruptions in different situations (Fischer et al., 2010; Mehrotra et al., 2016).

The conclusions drawn in the literature that relate the use of mobile devices in a face-to-face conversation ascribe elements of negative impression formation and 'interaction quality' (Vanden Abeele et al., 2016), with devices seen as having a profound negative impact upon face-to-face encounters (Nakamura, 2015). Even more, a perspective of faltering societal development has been used as a rallying cry in

the widespread critique of the use of mobile devices during encounters with others (Turkle, 2011) to encourage people to "reclaim conversation" with each other (Turkle, 2015). This thesis differs in its orientation to the study of device use by adopting an ethnographic perspective to illuminate how people practically accomplish device use in social settings. With this approach, this thesis eschews the demarcation of causes and solutions, as well as the contention of the 'morality' of the actions of patrons in using a device, and adopts a holistic approach to revealing the interactional accomplishment of how people use devices in a casual-social setting.

By studying natural interactions of groups of friends socialising together in a pub, the concept of pub talk with and around smartphones will be explored. Pub talk is perhaps best described as informal chatter that consists of "repetition, rhetorical questioning, and apparent irrelevance" (Mass Observation, 1943, p. 241), or in other words, pub talk could be construed as informal and relaxed conversation and could be regarded as the overall *accomplishment* of inhabiting a pub to socialise. This chapter will seek to explicate the gloss of 'using a mobile device in conversation' and crucially reveal the methodical actions taken to use a mobile device while also engaging in social interaction with co-present others.

4.2 STUDY DESIGN

This section outlines the design decisions made with respect to the study. Summarily, a video-supported ethnographic study informed by ethnomethodology (Garfinkel, 1967) was undertaken, in which groups of friends socialising together in a naturalistic manner. To do this, the friends were recruited to go to a pub together to socialise, and be video recorded for the duration of the gathering—groups were not recruited to use devices or guided to complete a given task. The collected data were reviewed and analysed with an orientation to making sense of the 'members' accomplishment' and orientation

of device use within the setting. The videos were viewed, with periods of mobile device activity ('fragments') catalogued for further in-depth review and remarked with notes upon the primary activity or 'purpose' of the device use. These fragments were then viewed and described in a more fine-grained manner—detailing what was done with the device in terms of movement, the visibility of the screen, whether someone mentioned or brought the device or purpose of the device use up in talk, and so on—in order to provide an index into the various members' practices. This provided a comprehensive oversight of the data and supported the selection of fragments as *vivid exhibits* (Crabtree et al., 2012, p. 111–112) of how members organised device use within pub talk. The resulting analysis rests upon the assumption "that knowledge and action are fundamentally social in origin, organization, and use, and are situated in particular social and material ecologies" (Jordan and Henderson, 1995, p. 41), and that by orienting to the actions of members, and the sequentiality in which they are organised, the practice of how device use is begun, carried out, and ended in the course of conversation can be revealed.

4.2.1 *The pub as a study setting*

A motivating factor for the studies in this thesis is to explore the interactions that have led to the rhetoric around the impacts of technology use in everyday situations, and especially when friends are socialising (see Chapter 2) in settings that could be considered both 'casual' and 'social'. This thesis defines such a place as one where individuals purposefully co-inhabit with the purpose of socialising in a relaxed and unimposing environment and where conversation is the main activity. This stance also does not place restrictions on a venue that is exclusively public or private, or the type of venue.

The definition is perhaps most closely relatable, but not entirely congruent to the notion of "third places" proposed by Oldenburg (1989). Third places are spaces that are outside the home or work-

place, where people can gather, socialise, and relax. The definition of *casual-social* augments this notion by allowing for places that may be considered homes, or perhaps even social environments in or near to a workplace. For example, the common-sense experience is that it is routine for groups of friends to meet in public plazas, cafés or restaurants, homes, libraries, and comfortable spaces around work environments to 'catch up' and socialise with each other. Each of these places can be a relaxing and social forum supportive of group conversation where the purpose of the gathering in such a place is to support leisure-time.

Laurier et al. (2001) demonstrate how settings such as those explored in this thesis still exhibit organisational traits, although they typically lack "complex articulation and coordination work" (ibid., p. 222) found in more formal settings. However, they still show that cafés and "places of that type" (ibid., p. 199) provide a "common code of conduct" (ibid., p. 210) that is informal yet provides guidance of behaviour that is adhered to by members. This conduct is found to exhibit elements of informality and engenders expectations of how members demonstrably and competently perform the work of socialising together. However, the work in this chapter will attend to how members' articulation and coordination work still occurs within members' interactions with, around, and through mobile device use and conversation.

In selecting a fieldwork setting, consideration was given to a variety of venues including cafés and public squares. A pub was selected for a number of reasons, some logistical and others sentimental. Su and Wang (2015) remark that with their observational study of mobile phone use in Irish pubs: "without doubt, the mobile phone is ubiquitous in pubs [...] [participants] overwhelmingly acknowledged that mobiles, used tactfully, was not a breach of etiquette" (ibid., p. 1663). Additionally and anecdotally from personal experiences, pub settings would allow for the observation of naturally occurring interactions

around mobile device use in an environment in which mobile phone use is common, and sometimes at the derision of co-present others.

The devotion of spending leisure-time in pubs and bars with friends is a popular British pastime; pubs typically open early and close late, many provide food and drink, and they serve as an environment suited to relaxing and conversing with others. In describing her observations of English culture, anthropologist and popular social science writer Kate Fox describes pubs as "a central part of English life" (Fox, 2004, pp. 88–108) and others have also highlighted pubs "as a social centre for the community" (Clarke et al., 2000, p. 693). These descriptions are also reflected in official statistics which state that 48% of people aged 16 and over would choose to go to a pub or bar in their free time; this figure is even higher for younger age groups (Seddon, 2011). Therefore, observing a gathering of friends who are leisure-time socialising in a pub would act as a suitable setting and activity in which mobile device use has previously been identified to unfold.

In summary, pubs are a *perspicuous* (Garfinkel, 2002a, p. 181) setting in which friends gather to socialise (corresponding with the premise of the study and how data was collected, discussed below in 4.2.2) and in which mobile device use is *ubiquitous*. In the words of Garfinkel, a pub makes available the "material disclosures of practices of local production and natural accountability in technical details *with which to find, examine, elucidate, learn of, show, and teach the organizational object as an in vivo work site*" (ibid., p. 181). In other words, mobile device use is routinely accomplished and regulated in and through pub talk, and by studying the use of devices *in vivo*, the phenomena of inquiry—that is, the use of touchscreen-based mobile devices in conversation—can be explicated. Based on up the above literature, statistics, and personal experience, it was natural to conclude that pubs provide a suitable and natural environment for the study of how people embed mobile device use during the conversations in a casual-social setting.

4.2.2 *Collecting data in the pub*

After finding a pub that agreed to host the research, participants from the university were recruited using email and word-of-mouth. Participants were recruited as groups of friends who felt they would "typically go to the pub with each other" and were willing to be observed for their "behaviours around mobile devices" within a pub. In total, eleven participants took part (in three separate groups); seven of the participants identified as female, with the remaining four identifying as male. Each group had at least one female and one male, although this was by chance and not intentional. Of the recruited participants, four were aged 18–23, five were 24–29, and two were 30–39. The studies were conducted over a three-month period in the UK, taking place at a time agreed with the recruited participants. The study was approved by the university's School of Computer Science Research Ethics Committee and participants were reimbursed with an online shopping voucher for their time spent during the study.

The study was *participant-observer* in practice. To achieve this in a naturalistic way, the only 'activity' asked of participants was that they converse as they normally would—no tasks were given to participants (the information sheet that was given to participants prior to the study is in Appendix A.1). Video and audio recordings of participants socialising together in groups were collected as part of the approach in an effort to allow for the explication of the interleaved use of mobile devices in the conversation, and to study the observable-reportable actions exhibited by members of the setting through video ethnography. The video data was captured with two fixed wide-angle lens cameras on tripods, and a separate audio recorder on the table for higher-fidelity audio to avoid issues of the noise of the environment 'drowning' out the sound of the participants. This allowed for video data to be collected unobtrusively during conversation without the requirement of camera operators being present. Relying upon field notes would have hindered participation, and brought attention

to the observation of the conversation while also providing a lower fidelity of data.

Questions were asked after the 'observation phase' as an interview (the primary questions are given in Appendix A.2) so as not to interrupt the flow of the conversation. The purpose of the interview was to contextualise the observations and gain an insight into the participant's perceptions of mobile device use in conversation. Given the evolving landscape of mobile technologies, this acted as a point of curiosity to understand the present situation.

Through the questionnaire, participants were asked about which technology they owned: all participants owned smartphones, and had them present, a majority (seven) also owned tablets (although six of these relied on a Wi-Fi connection); however, none had a tablet with them, and there were no smartwatches.

Overall, the ethnographic record is comprised of video recordings of the interaction, field notes made after the session, individual questionnaires completed by members, and the recording of the informal semi-structured group interview.

4.2.3 *Analysing the collected data*

To analyse the corpus of collected video and audio data, video analysis, drawing on ethnomethodology (C. Goodwin and Heritage, 1990; Heath et al., 2010) was performed. Firstly, shortly following data collection, the corpus of data was catalogued and indexed to identify episodes in which mobile device use occurred. Timestamps and descriptive language were used to construct a record of the interactions that took place, which allowed for iterative re-examining of prior data with relative ease. This was in order to aid the discovery of the observable-reportable actions performed by the members of the setting and to help gain an overall impression of the data collected across all the sessions.

In total, 51 episodes of mobile device use in the sessions were identified (some of which were overlapping), with episodes ranging from a few seconds to a few minutes in length. A substantive review of the episodes was performed to examine the interaction that unfolded, honing in on episodes that represented observable-reportable intersections of mobile device use and conversation for a more in-depth analysis. Seven fragments were selected for further analysis where instances of device use lasted for more than 'a few seconds' and where this interaction occurred in and through conversation (i.e. device use was interleaved in some grossly observable fashion with conversation). In each case, these fragments were selected in line with the aims of this research to reveal the social organisation of device use in pub talk and that were deemed to warrant further investigation, with each fragment being reviewed individually and discussed collaboratively with other researchers multiple times. These fragments were then transcribed with both verbal (i.e. talk) and non-verbal (e.g. gestures and other interactional resources) being carefully noted. Situations where, for example, mobile devices were used merely as timepieces for a split-second, were ignored and not used within the corpus.

Following multiple iterative reviews, a collaborative 'data session' was performed, with other HCI and CSCW researchers within the same research group invited to watch, review, and comment on collected video data and analysis, and to provide critical reflection on the findings explicated. In this session, observations and commentary developed through the analysis were provided by the author along with transcripts of the clips, and all were reviewed in a collaborative and reflective manner.

4.3 FINDINGS

Three fragments will now be introduced and presented over a series of 'data excerpts'—these fragments are *vivid exhibits* (Crabtree et al., 2012) of the data within the collected corpus. By this, this thesis con-

siders each of these as exemplars of the activities of members' observed practices in the collected corpus in which members' interactional projects are accomplished in and through the use of devices in a pub. In summary: the first fragment illustrates typical interactionally unproblematic use of a device to introduce new information to the conversation, the second fragment introduces a more complex case that reveals more features of interaction in which a device is used as a resource to make a joke, and the final fragment introduces an interactionally problematic case in which device use is used to contest an argument (Heath et al., 2010, p. 111). Throughout the explication of the fragments, the practice of how device use is used as a mundane activity in pub talk will be established.

In each fragment, the setting will be introduced, giving a background to the conversation that is unfolding. By this, this thesis means to demonstrate the ways in which: device use is brought about *in and through* conversation as an interactional accomplishment, how members perform actions throughout the conversation to sustain device use within it (often this involves articulation and accounting practices), and how members stop using the device.

Appendix B provides details of the transcription notation used in this thesis. All names and identifiable information within the transcripts provided are entirely fictional.

4.3.1 *Introducing new information for conversation*

In the first excerpt from the fragment, titled *Miniature Schnauzers*[2], given in Data Excerpt 4.1, two friends, Cally and Dayna, are discussing their favourite dog breeds. They are sitting across the table from two other friends who are having a separate discussion. There is a disagreement between Cally and Dayna, as a matter of personal preference, in relation to their favourite dog breeds.

2 The complete fragment is included in Appendix C.1.

01		CAL	i like miniature schnauzers
02		DAY	°how big are schn-?°
03		←CAL	it's like (.) like (.) they're
04			so: cute

Data Excerpt 4.1: Miniature Schnauzers (i)

In this opening excerpt, Cally occasions the interactional project of *providing new information to the discussion* in and through the ongoing conversation. This is to address an information deficit that develops once it is established that Dayna is unaware of the size of Cally's favourite breed. We first joined the discussion between the pair as Cally establishes her preference for Miniature Schnauzer (line 01) and Dayna seeks clarification on the size of the breed (line 02). This small snippet provisions the members' occasioning of the forthcoming device use which will be explicated.

Later device use will be occasioned by this established information deficit through the production of two key activities in this fragment. These will be unpacked separately below:

(i) Getting someone to look up new information, and

(ii) Collaboratively finding new information.

Each of these actions are unpacked respectively in relation to how they are used to accomplish the occasioned interactional project.

4.3.1.1 *Getting someone to look up new information*

The discussion continues below in Data Excerpt 4.2 where both Cally and Dayna reflect upon their favourite dog breeds. Cally previously gestured with her hands for an approximately waist-width-sized dog (line 03, above), and Dayna copies this gesture (line 05, next excerpt)

before stating her preference for "big dogs" (line 07). Cally acknowledges this remark with "i know" (line 08), and then instructs Dayna to use Google to search for the breed. Through this instruction, Cally is getting Dayna to use the device in order to provide her with the additional resources to allow her to make sense of Cally's perspective (i.e. as a request to 'try to see it from my perspective').

05		←CAL	((briefly looks at her bag to
06			her left before looking back))
07		DAY	i like big dogs
08		←CAL	i know, but google schnauzer,
09			right?
10		DAY	((gets phone out from bag))
11		CAL	((leans towards DAY))
12			the puppies (.) schnauzer
13			puppies are gorgeous

Data Excerpt 4.2: Miniature Schnauzers (ii)

Cally's request is treated as unproblematic in the routine of talk-in-interaction in a pub by the friends. Whereas it is foreseeable that such a request may be explicitly oriented to by as members *out of place* in other settings, in the pub—and as a constituent activity of pub talk—such a remark is produced and responded to as not out of place. In other words, Cally's occasioning of the device use through the request for Dayna to search for information, and Dayna's acquiescence to the request from Cally, establish that the practice of getting someone to look up new information is mutually regarded as acceptable practice as part of the social organisation of interaction in a casual-social setting such as the pub.

4.3.1.2 *Collaboratively finding new information*

Following this discussion, Cally re-orients her gaze to the others conversing at the table, responding to the continuing discussion amongst the other three people at the table. Some 12 seconds later she refocuses on the topicalised device use, as exhibited below in Data Excerpt 4.3, by shifting her body posture and gaze to look at the device screen. In the next excerpt from the data, both Cally and Dayna work together to collaboratively accomplish the task for which the device use was occasioned (introducing new information for conversation).

At the start of this fragment, Dayna tilts her screen marginally towards Cally, and Cally reciprocates by leaning towards Dayna and shifting her eye gaze towards the screen (visible at line 17). Cally re-prompts Dayna with the search term: "`miniature schnauzer`" (line 15). Dayna seeks guidance from Cally on the spelling of the terms to type into Google, although she does not complete the utterance asking for this; it is established through Dayna's phonation of the word "`schnauzer`" (line 18) and apparent inability to complete the task. Cally provides the spelling (line 20), with Dayna enacting Cally's guidance by typing the dictated letters on the device, established as a Google search through inference of the ongoing interaction. Following the action of Dayna typing in the search term, as dictated by Cally, Cally proceeds to provide further guidance by instructing Dayna to "`go look at schnauzer puppies`" (lines 23–24).

This action is imbued with collaborative efforts from the pair as device use is interleaved within the talk between the two. Cally *instructs* Dayna to search for Miniature Schnauzer, although remains attentive to the device use—shifting her posture and fixing her eye gaze to the device's screen. As Dayna seeks clarification on the spelling of the search terms, Cally responds to the clipped request "`how do you?`" (line 16) with "`erm`" (line 16), with Dayna clarifying it is the spelling of the term *Schnauzer* that she is seeking assistance with through her phonation of the word (line 18). Cally responds to this by providing the spelling (lines 20—21).

15		CAL	so it's miniature schnauzer
16		DAY	how do you?=
17		←CAL	=erm::
18		DAY	(sccchhhh) (tea) (ee) (ar)
19		CAL	oh schnauzer (.)
20			it's s-c-h-n-a-u-z- n-a-u-
21			(2.2) schnauzer
22		DAY	oh, schnauzer
23		←CAL	schnauzer, go look at
24			schnauzer puppies right↑
25			((continues to look at phone))
26		DAY	°my internet is rubbish so
27			this may take some time°

Data Excerpt 4.3: Miniature Schnauzers (iii)

Cally accountably provides support to—and engages with—the device use by Dayna, as established through her gaze and posture as to be 'watching' and participating with the device use. By participating with Dayna to complete the occasioned activity, Cally's intention of ensuring the completion of the occasioned task is made visible (i.e. Cally waits for Dayna to do the task requested, and watches her as she goes to complete it). Following Dayna's completion of the input into the device, Cally provides further instruction to Dayna to examine photos of the puppies: "`go look at schnauzer puppies`" (lines 23–24), however, Dayna rebuffs this remark by stating that her internet is slow and that it "`may take some time`" (line 27). The act of remarking that her device Internet is slow, and putting down her device, completes this period of collaborative device use, and is followed by both members re-joining the existing conversation taking place between the other members at the table. Dayna returns to the device approximately two and a half minutes later on and examines the photos with Cally, as concluded in Data Excerpt 4.4.

28		DAY	((looks down and unlocks
29			phone)) ↓oh that: thing
30		←CAL	((leans towards DAY and
31			shifts gaze towards her
32			screen)) yes look at them
33			oo:::↑
34		DAY	°schanuzer°

Data Excerpt 4.4: Miniature Schnauzers (iv)

In this final excerpt of the fragment, Dayna has looked down and unlocked her device, resuming the interactional project to retrieve new information, which was halted because her "`internet is rubbish so this may take some time`" (lines 26–27). After unlocking her phone, she utters the remark "`oh that thing`" (line 28), reporting that the search results have loaded—and through this alludes to her peripheral awareness of the breed. Cally, in turn, resumes the posture she held earlier to participate in device use, i.e. that of leaning towards Dayna and looking at her screen; then, while pointing at the images on Dayna's screen, she utters "`look at them`" (line 31), reaffirming her opinion that she considers the breed to be cute, although Dayna fails to acquiesce to this opinion. Following this fragment, Dayna locks and puts her phone down, and both her and Cally resume conversing with the others at the table.

This fragment thus far shows a typical sequence of members using a mobile device in pub talk unproblematically. In this case, it was to introduce new information for conversation. Cally did this by getting Dayna to look up new information and then collaborating with Dayna to complete the task which she instigated. This fragment should provide an initial sense of the activity of bringing a device into conversation, and how the device use does not necessarily become a topic in of itself and is ostensibly treated as perspicuous to the setting and activity of conversing in a pub. Device use within

pub talk is punctuated by pauses and halts to the device interaction as members wait for responses, or reorient between 'conversational floors' (Edelsky, 1981) (a pause of 12 seconds between Fragments 4.2 and 4.3 and over 2 minutes between Fragments 4.3 and 4.4). The next two fragments present an increasingly complex depiction of the phenomena, in which the interactional projects of members consist of using a device to make a joke and to contest a response in an argument, and in which the problems being dealt with in interaction deviate from collaborative efforts exhibited by members thus far.

4.3.1.3 *Methodical accomplishments in this fragment*

Before turning to the next fragment, a quick recap is provided on the methodical accomplishments of the members in this fragment in successfully using a device to add new information to a conversation. Firstly, conversation progresses on to different dog breeds, and it is established how Dayna is unfamiliar with a given breed. This information deficit occasions a member, Cally, to **instruct Dayna to use her device** to use Google to search for the breed. Once ready to perform the task, Dayna seeks clarification on the search terms to use. Cally **clarifies the activity** by spelling out the words to use, telling Dayna to look at images, and all the while looking at Dayna's phone screen. Due to her slow Internet connection, Dayna **temporarily puts her phone down** while the results load and then talks with the others at the table. A few moments later, Dayna unlocks her phone and **looks at the screen** to see the results, Cally joins in too by leaning towards Dayna and **examining the images displayed**. They both **acknowledge the information sought has been found**.

4.3.2 *Making a joke*

The first fragment has shown how device use can be occasioned in and through the conversation to introduce new information. The next fragment of data, titled *Font Size*[3], exhibits how members are able to retopicalise conversation and introduce device use to the floor, in this case to make a joke at the expense of another person. In this fragment, which commences below in Data Excerpt 4.5, there are five friends including the researcher. They are currently discussing Christmas food and as we join them in Data Excerpt 4.5, Lawrence has recently returned to the table from buying another drink from the bar and began using his device. Jayne is currently recalling a pub she visited in September, which had Christmas 'stuff' out.

01	←JAY	beginning of september they
02		had their (.) all their
03		christmas stuff out (.) and I
04		was °like oh my god nobody
05		( )°

Data Excerpt 4.5: Font size (i)

In this excerpt, Lawrence's device use remains not explicitly accounted for thus far, as the group continue to discuss what is deemed by the members to be the absurdity of Christmas paraphernalia set out in a pub in September. Lawrence will then introduce a new topic to the conversion to make a joke unrelated to the conversation at hand. As has been established above, conversation in pubs is informal and vacillatory, allowing members to alternate to-and-from device use. Pub talk is shown to feature disjunct topic shifts that are immediate and disconnected from each other, and ultimately, topics in conversation do not require formal resolution or agreement. Following on from this, Lawrence undertakes two actions:

(i) Preparing the floor for a joke, and

3 The complete fragment is included in Appendix C.2.

(ii) Delivering a joke.

These actions are unpacked in the next two sections to reveal how they are formed as constituent activities of pub talk.

4.3.2.1 *Preparing the floor for a joke*

While the discussion continues, Lawrence has been using his device solitarily. He then, as seen in Data Excerpt 4.6, moves the phone close to his face while simultaneously shifting his head such that he is looking closely at his device's screen and utters "jesus" (line 06).

06		←LAW	°jesus!°
07		JAY	we just booked ours (1.0) we
08			do me and liam and james and
09			malcolm do (one every year and
10			we) just booked it
11		MAL	du bois↑
12		LAW	=sorry (.) have you (.) um (.)
13			((jovially)) jonathan has sent
14			round an email (.) this is
15			great for your study isn't it?

Data Excerpt 4.6: Font size (ii)

In this excerpt, the conversation continues without remark from the others regarding Lawrence's "jesus" (line 06). He then interrupts the group, as Jayne and Malcolm are discussing their Christmas meal plans with the other people at the table, by first issuing an apology for the interruption ("sorry", line 12), and then chuckling while drawing attention to an email he has received from a third party, Jonathan, by asking whether others have seen an email that has been sent to a mailing list (lines 12—14). As the friends are all students at the same university, it is possible the email was sent to a mailing list which Lawrence assumes all others will have received too. Finally, he remarks to the researcher that "this is great for your study" (lines 14–15).

Through this interruption to the conversation, Lawrence is introducing a new topic to the group that is unrelated—he acknowledges

this with an apology—and then prepares his friends for his upcoming device use. In this, he is establishing that the new topic he is introducing is about an email that has been sent around. However, at this point and without further details, others at the table are unlikely to be able to respond to the question without further details of the email in question. Thus, this utterance, which is delivered while chuckling, occasions a device to be introduced into conversation in order to demonstrate the reason for which the topic was brought up. Lawrence provides incomplete detail of the email by remarking that it was Jonathan who "has sent round" (lines 13–14) the email, but the incompleteness of his description establishes that he will give further detail in a forthcoming utterance.

4.3.2.2 *Delivering a joke*

While Lawrence has established that the email is remarkable by bringing it up, and has alluded to a humorous aspect through the jovial delivery of his interruption, he has yet to establish the specifics of the email to others. In this final excerpt, below in Data Excerpt 4.7, Lawrence establishes and delivers the joke, using his mobile device, to demonstrate what he perceives as an absurd font size.

16		←RES	going to have to zoom in for
17			the camera (.) it's only set
18			to 720p!
19		JAY	mu:::::a::h
20		LAW	yeah (.) that's (.) that (.)
21			that's the email!-
		(...)	
26		MAL	is that him or is that your
27			phone fitting the line in?

Data Excerpt 4.7: Font size (iii)

As Lawrence rotates his device to others to show them the email, the researcher comments upon him having to "zoom in for the camera" (lines 16–17) as it becomes clear to others the remarkable feature of the email, for which is the basis of the joke, is the small font

size. Lawrence continues to joke about the font size—reiterating that "`that's the email`" (line 21). A discussion unfolds as Malcolm verifies the basis of the joke, establishing whether that it is a display issue or the email itself. The conversation, as before, then continues to be around the newly introduced topic of the email formatting. Through his preparation of the floor for his joke, and his delivery through rotating his device screen, Lawrence has established a joke by ridiculing the sender of the email and the formatting of the email message.

This fragment reveals how Lawrence, in using his device in pub talk, is able to establish a joke using the device as a resource in conversation and occasioning a conversational topic shift to his device use. This fragment exhibits how a device can be used by members as part of previously undisclosed matter in pub talk and that it is not solely a resource for previously established conversational topics. The informality of pub talk means that such occasioned interaction is not treated as a problematic interaction by other members by acquiescing to the topic change and engaging with the delivered joke. In the next fragment, this understanding of pub talk will be further extended with an exhibit of how device use is used to contest an argument, but in lieu of a resolution, a topic change unfolds.

4.3.2.3 *Methodical accomplishments in this fragment*

In this fragment, a member, Lawrence, is already using his device, although the specifics of this engagement are unaccountable to the others in the setting. Lawrence then **interrupts the existing conversation**, firstly with an apology, and then **asks an ostensibly rhetorical question** to those who are co-present regarding an email he has received. He does not provide the specifics of the email at this stage but then **rotates his phone around** so that the other members at the table can see the email on his screen, and jovially remarks that "`that's the email`". This remark is predicated on the small font size with which the email is displayed on his device, which is accountably visible to other members as a result of him rotating his phone. Through the delivery

of this joke, and given the vacillatory nature of pub talk, Lawrence **introduces a new conversational topic** about the email he has received.

4.3.3 *Attempting to contest an argument*

The next fragment, called *Shorthand*[4], focuses on what, at a glance, may seem to demonstrate device use occasioned to attend to the same matter as above (i.e. that of introducing new information to the conversation). However, through unpacking the data in this section, the actions of members will be shown to be driven by a fundamentally different members' problem. In this fragment, which commences below in Data Excerpt 4.8, the same friends are returned to; they are currently discussing the notion of observational studies, ethnographies, and making fieldnotes. This, in turn, leads to a discussion about shorthand notations. The fragment begins with Lawrence questioning Jayne whether shorthand notations are "`mainly phonetic`" (line 01), with Jayne's indirect response occasioning a disagreement between the two co-interlocutors.

01		←LAW	isn't it mainly phonetic?
02		JAY	it's like:
03			(3.2)
04		←JAY	there's various versions so
05			the one she tried to teach me
06			first so i could start going
07			is missing out all the vowels
08		LAW	((briefly looks at JAY while
09			picking up his phone, he then
10			(begins to use his phone once
11			he has it in his hands))

Data Excerpt 4.8: Shorthand (i)

4 The complete fragment is included in Appendix C.3.

In this opening exchange, Lawrence establishes his perception that shorthand notations are "`mainly phonetic`" (line 01), to which Jayne fails to provide a direct response to the question (i.e. of *yes* or *no*). Instead, her response implies *no* through the use of the opening phrase "`there's various versions`" (line 04), which she follows with an explanation of a shorthand notation based on omitting vowels (i.e. a notation that is *not* phonetic). Through her response to the question, Jayne establishes a disagreement with Lawrence that shorthand notations are mainly phonetic. Lawrence, instead of acquiescing, in turn, picks up his device and (as is revealed later on) begins to search for information on shorthand notations (lines 08 onwards), demonstrating his disagreement, or least dissatisfaction, with Jayne's as-yet-incomplete response. What follows on from this opening activity is two key activities that further demonstrate the complexity and intricacy in which device use is interwoven within conversation in the pub, and exemplifies how troubles occur as a result of device interaction:

(i) Halting the conversation to attempt resolution, and

(ii) Resuming without resolution of the disagreement.

Both of these actions are unpacked respectively, with relation to how they are formed as constituent activities while using a mobile device in pub talk.

4.3.3.1 *Halting the conversation to attempt resolution*

A continuation of the fragment is given in Data Excerpt 4.9. The conversation between the group continues with Jayne giving an explanation of a second shorthand notation; Malcolm makes a remark that cannot be deciphered but which is cut off by Lawrence through his utterance of "`hang on`" (line 22). Through this interjection, Lawrence attempts to stop the progression of the conversation between the other members. He then begins to phonate the word "shorthand" while typing on his device (lines 22—24), accountably searching for the information in response to his question.

12		LAW	↓yeah
13		←JAY	and once you get good at that
14			you just write a lot quicker
15			(0.7) but then she had one
16			which was literally like (.)=
17		LAW	↓yeah
18		JAY	=swiggles and just didn't look
19			like anything and i don't know
20			if that's phonetic or::::
21		MAL	( )
22		←LAW	hang on! ((typing on phone
23			with thumbs)) schuh::::::::ort
24			(.....) hand (..)
25			my mum's regular handwriting
26		RES	i know some people who miss
27			out vowels (.) like the e
28		←JAY	that's how i do it (.) missing
29			out vowels is very very good
30			but there's a squiggly one i
31			don't understand

Data Excerpt 4.9: Shorthand (ii)

Lawrence's issuance of a command to the co-present others to "hang on" (line 22), followed by elongation of the phonation of the word *short* (line 23) ensure he continues to hold the floor in conversation. As is revealed through a post-observation chat, he was using his device to search for information on shorthand notations in order to support his point. In other words, Lawrence stops the progression of conversation and holds the floor, to retrieve information in order to resolve the disagreement that occasioned the current conversational topic.

Through his device use and his verbal articulation, Lawrence exhibits how he is attempting to resolve the disagreement—by searching on his device for information to support his point. Yet, he does this concurrently while an answer to his question is being given; he then interrupts others' talk and holds the floor while he completes his occasioned device use. Such actions reinforce the notion that conversation within the setting is treated as an informal activity. Through

the mundanity with which he and other members treat his device use, which is ostensibly not interleaved in the ongoing multi-party conversation, and his holding of the floor while using the device, it is established how device use and conversation are treated with a sense of informality both by Lawrence *and* his interlocutors.

In this situation, not only is Lawrence's device use not ostensibly interleaved with the conversation, there is also an attempt to halt the progression of conversation while he completes the device use. Nevertheless, this request is not honoured by the conversationalists and the talk between the other members continues. Initially, the researcher first renews the topic of shorthand notation at the table before others continue the discussion while Lawrence uses his device away from the group.

4.3.3.2 *Resuming without resolution of the disagreement*

The next part of this fragment also further brings to the fore the demonstrable ways in which members' interactions establish the notion of the pub as a casual-social setting and that of pub talk as being imbued with informality. Thus far, device use is showed as being used as a resource in conversation to retrieve further information (as per the prior Miniature Schnauzers example in 4.3.1). Whereas in the prior case, device use is shown to be used to address the members' problem, in this fragment, continued in Data Excerpt 4.10, conversation is shown to continue without the intended purpose of device use being concluded, with a disjunct topic shift occurring.

32	MAL	this is why i didn't do
33		ethnography (1.8) just get the
34		participants to fill
35		everything out
36	LAW	i didn't do ethnography either
37		either!
38	MAL	yeah↑ you do it- i'm not- i'm
39		not doing=

40		←LAW	=oo↑ that's got (2.0) that's
41			cinnamon in it or something
42			something (..) smells amazing

Data Excerpt 4.10: Shorthand (iii)

Malcolm's remark that he "didn't do ethnography" (lines 32–33) is responded to by Lawrence with a claim that he "didn't do ethnography either" (lines 35–36), at which point he stops using his device (although still continues to hold it). Malcolm responds to this retort but is interrupted by Lawrence who shifts the topic of conversation on to Zoë's food by shifting his posture towards her, leaning in towards her plate, and remarking that "oo that's got cinnamon in it" (lines 40–41). Following this, a conversation about food, and Christmas food in particular, unfolds, with the disagreement remaining unresolved and not returned to by any members.

This fragment, in its entirety, establishes how using a device while someone is talking to you, and interrupting a conversation to use your device, although highlighted as potentially problematic, may not be treated as such by members in and through pub talk. The first fragment of interaction in this chapter, which focused on Miniature Schnauzers, concluded with the members returning to the device to complete and resolve the problem at hand. The second fragment further brought into play how conversation topic changes occur within pub talk and render prior discussion points amongst friends as unresolved. Moreover, this fragment further demonstrates that given the fleeting nature with which conversation progresses and how the topic being discussed changes, reasons for which a device use was occasioned may not be satisfied—resolution of topics or discussions may not necessarily occur in pub talk, and throughout the corpus, was not treated as problematic.

Furthermore, the continuation of talk, and the later retopicalisation of the conversation on to Zoë's food demonstrate the vacillatory

nature of pub talk, i.e. the topic at hand in a conversation can change in rapid succession, with members changing between activities such as device use and talking with relative ease and wavering between the two. Although Lawrence began using his device to resolve the disagreement in talk, he then changes the topic readily, without recourse, and without others calling for him to resolve the disagreement. In other words, the topic of shorthand is not returned to—Lawrence provides no results from his device use, and nor is this sought for by other members.

4.3.3.3 *Methodical accomplishments in this fragment*

In this final fragment for this chapter, a conversation about shorthand notations is ongoing. Lawrence asks Jayne a question regarding different notations but **commences to use his device** as an indirect answer is given by Jayne. Lawrence interrupts Jayne's answer by **requesting that she waits for him to use his device** and then he **phonates the word shorthand** while typing on his phone's screen, accounting for the specifics of his device use being to search for information to contest his point. After his phonation, Lawrence continues to use his device however the **conversation amongst the others at the table continues** without waiting for him to complete his self-occasioned task. Lawrence eventually **returns to talking with others** without resolution of the contested point by **stopping the use of his phone**.

4.4 CHAPTER SUMMARY

This study not only provides a basis for understanding how mobile device use is embedded within a multi-party conversation in a pub, but also for how the interactional accomplishment of sustaining the device use is achieved when the mobile device relies entirely on line-of-sight and touching the device screen.

These three fragments of interaction have increasingly introduced the intricacy through which mobile device use unfolds as an inter-

woven activity within pub talk. Across each fragment, device use is shown to move quickly, with conversational topics changing quickly, to and from device use, and disjunctively to new topics. This device use is interleaved within this conversation—members were shown to bring a device into conversation to introduce new information, to resolve disagreement, or to make a joke; and to stop using a device due to sluggishness of the device, once the problem for which the device use was occasioned is 'resolved', or as a new topic is introduced. Crucially, this interleaving is shown to be treated as a consistent and inseparable activity of pub talk and that the device use by members is interleaved within the sociality of the setting. The device is used *as part of* conversation in a pub, and is used in and through that conversation by members.

Furthermore, device use was shown to be engaged with collaboratively in searching for new information together, or individually and separate from conversation. While this latter case could be construed as problematic (and indeed has been by others as discussed above in 4.1), it is noteworthy how in this video-supported ethnography that members performing such actions were seemingly not identified as problematic in interaction by others. In the three fragments presented, as exhibits of the complete corpus, device use was not explicitly oriented to by members as unnecessary, rude, or distracting in talk, even in cases where conversation was interrupted to allow for device use to unfold. While of course, some people will argue that device use is any or all these things (again, see 4.1), the people observed in this ethnography reveal that this is not a universal case. The resulting conclusion of these observed practices is that device use, as supposed prior to data collection, is treated as perspicuous to the pub as an activity interleaved within conversation.

In summary, device use was occasioned as a result of—and in spite of—conversation, and treated by members as a mundane activity within conversation in a pub, even if that use was not topicalised in conversation. Furthermore, the informality of the setting and that

of pub talk further mean that resolution of the occasioned device use is not necessarily made accountable.

4.4.1 *Methodical accomplishments*

This chapter has stepped through three fragments introducing the pub talk in which members have occasioned and brought device use into the conversation. This occasioning was done to introduce new information to the conversation, to make a joke, and to attempt to contest an argument. Practically, this occasioning practice took the form of instructing others to use their device, choosing to use the device oneself, or using a device as ostensibly unrelated to the conversation at hand. Through the interleaving of device use in interaction, members accounted for their device use—i.e., they made the specific nature of their interaction observable and reportable—by rotating their phone, by holding it between people to make the screen visible, or by phonating what they were doing on the phone. Device use ended (temporarily or entirely for the occasioned task) as an interleaved activity as the conversation topic changed, as devices took time to respond, or that the interactional project of using the device was accountably completed.

4.4.2 *Outlook*

Study two, in Chapter 5, considers how such gatherings unfold when the use of the device occurs through speech using a Voice User Interface (VUI) on the smartphone in a similar setting: a café. VUIs are different from Graphical User Interfaces (GUIs) in that they provide a mechanism for people to interact with a device through their voice, although in the case of smartphones the VUI also displays its computation and output on the screen of the device. It is so posited that the consequences of interaction occurring through this paradigm are that the tasks being completed using a VUI become hearable to those within earshot, and that through this, the device use interleaved in conversation in a casual-social setting is made naturally accountable. While interactions around the device use in this chapter were shown to be collaborative at times, it is proposed here that by requesting members to make a preference for voice-based device interaction, the interaction with the VUI may lend itself to occasioning further collaborative efforts amongst the co-interlocutors. How members attend to conversing with each other while device use unfolds through voice as well as the use of the touchscreen, and how members attend and orient to this device use, will become a focus of the next chapter.

5

STUDYING CAFÉ TALK AROUND SMARTPHONE-BASED VUI USE

This chapter presents the study of friends socialising face-to-face as a group in a neighbourhood café, and examines their interactions around the use of the 'natural language' Voice User Interfaces (VUIs) on their personal mobile devices. The previous work in Chapter 4 explored how conversation in a pub unfolded with and around the use of the Graphical User Interfaces (GUIs) on mobile devices, and how these device-based interactions were interleaved within the conversation amongst groups of friends. This analysis revealed some pragmatics of how cooperative interaction occurred in, through, and around the use of the mobile device.

The work in this chapter pivots to considering how conversation unfolds around the use of a newer innovation, the 'personal assistant' on mobile devices, which is interacted with primarily as a VUI in combination with a GUI. It was previously posited in the outlook of the prior chapter that VUIs might lend themselves to collaborative interactions amongst the co-interlocutors around the device use due to the use of talk over touchscreen interactions. This turned upon the notion that talk to VUIs would make the specifics of members' use of devices hearable to those around the device use.

The research in this chapter was previously published and presented at the Computer-Supported Cooperative Work & Social Computing conference[1]—a number of changes have been made to address the research questions of this thesis.

1 See Porcheron et al. (2017b).

5.1 INTRODUCTION

Previously, this thesis studied groups of friends socialising in a casual-social setting—a pub—and explicated how members unproblematically interleave the use of touchscreen-based device interactions in talk to address problems at hand. In accordance with this thesis' aim to explore the broader uses of technology use and its interwoven nature in casual-social settings, this chapter progresses to studying interactions that unfold around the use of voice-based 'personal assistants' on devices. To recapitulate, this type of space provides a suitable natural environment to observe participant behaviours with mobile devices "in the wild" in a perspicuous setting (Crabtree et al., 2006). A casual-social setting forms an environment in which individuals and groups can socialise with each other, that may be in or outside of the home or workplace, and that provides a level of comfort and relaxation for those who gather there. In much the same way that a pub was previously regarded as a perspicuous setting for studying interactions around mobile device use, a café so too becomes such a setting for examination of naturally occurring interactions around device use. Turkle succinctly emphasises this notion of communal spaces where we are together with others, but where mobile devices are present and in use:

> In this new regime, a train station (like an airport, a café, or a park) is no longer a communal space but a place of social collection: people come together but do not speak to each other. Each is tethered to a mobile device and to the people and places to which that device serves as a portal.
>
> — Turkle (2011, p. 155)

The previous chapter explored how members of the casual-social setting account for and interleave device use within conversation, e.g. by bringing it up in conversation or sharing visibility of the screen. This thesis posits that by shifting the input to the device to talk, away from interactions using a touchscreen, members may simultaneously and

naturally account for the device interaction as an accomplishment of using the VUI, given how talk is the most obvious form of making one's actions accountable:

> Talk is the most obvious and pervasive way in which members conduct their work and make whatever it is that they are doing into an intersubjectively recognisable and naturally accountable activity.
>
> — Crabtree et al. (2012, p. 44)

Therefore, this study deviates from the prior examination of touchscreen-based device use by *requesting* participants to use the VUI instead of the touchscreen preferentially. Given this, it was decided to observe and record interactions in this successive study in a café due to lower background noises, although cafés remain true to the established definition of a casual-social setting: they are places where groups of friends can gather to socialise and relax with each other. As the work in this thesis is concerned not with device use as study of the technology, but rather with the study of interaction amongst multiple members of the setting while technology is being used during, asking participants to use the VUI does not influence the outcome of the study. Indeed, the focus is of the *naturally occurring* interactions *around* the device use.

As the name suggests, these VUIs are interacted with primarily through the user speaking, with software running on the device 'recognising' the words spoken, and attempting to run the command or request made by the user. Typically VUIs on personal mobile devices 'listen' to the user and display the 'understood' words as well as the course of action being taken by the device, or an error message. In this sense, a VUI on a portable device presents a composite, or *hybrid*, of both VUIs and GUIs, in which users must provide commands or instructions through talking to a device without *a priori* display of possible options. In turn, users are presented with a GUI to accomplish and understand the processing of their voice and the action taken by the device.

From a technical perspective, a factor behind the introduction of VUIs is that talk remains a flexible and pervasive method members routinely employ to accomplish a task[2], i.e. the scope and capability to communicate through talk is much greater than that through a pre-designed touch-based GUI, especially on small-screen devices given the potentially limitless use of vocabulary within human language. Marketing materials from companies such as Apple Inc. have promoted their respective VUIs as allowing users to 'get things done' with ease, allowing the user to talk to them as they would 'to a friend'. Moreover, research into talking with computers is long-standing, with work as far back as Licklider (1960) remarking that "there is a continuing interest in the idea of talking with computing machines" (ibid., p. 10). More recent work has pursued the ideas of talking machines (i.e. conversational agents) that act as companions for the elderly (Vardoulakis et al., 2012), or virtual museum guides (Kopp et al., 2005). This thesis, however, is primarily concerned with another form of conversational agent—the 'virtual butler' that helps people 'get things done' (Payr, 2013). In particular, the study in this chapter will explore the practical use of the virtual butlers readily found on peoples' personal smartphones and tablets to make sense of just how these 'butlers' are practically used in interaction to 'get things done'.

While there are a plethora of commercial products as well as active research projects examining various aspects of creating and enhancing VUI design, existing systems come with their own restrictions in terms of capability in translating spoken words and sounds into text, in parsing the text into machine understandable commands, and performing the desired actions with the understood commands. These technical limitations are not the focus of this thesis but remain a factor in considering how members attend to and accomplish their interactional projects with the devices.

The marketing materials for these VUIs, marketed as Intelligent Personal Assistants (IPAs), suggest that they can be interacted with like

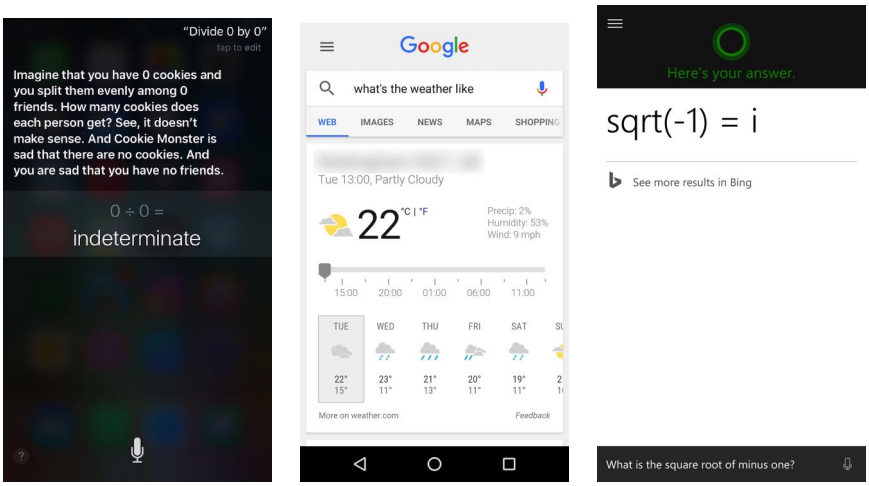
2 Beyond, of course, mere reasons of selling more products.

any person might, and can respond to natural human talk. For example, two such systems, Siri (Apple Inc.) and Cortana (Microsoft Corporation), both appear to exude humour in response to general 'conversational' input, questions, and instructions. In turn, their responses to the user might be seen as sarcastic or entertaining. While VUIs may provide the veneer of conversational intelligence, this study examines the interactional accomplishment of using a VUI in and through interaction, grounded in the empirical evidence derived from a video-supported ethnography, and reveals the cooperative nature of interaction in, through, and around the use of the device.

On a large proportion of current mobile devices, the VUI may be triggered through one of two means: by pressing a physical or on-screen button, or by the utterance of a 'hotword' that enables the interface (e.g. "Hey, Siri"). The human interlocutor (i.e. the user) then talks to the assistant, and is able to engage in dialogue and ask questions (e.g. about the weather), or give instructions (e.g. to call someone); the VUI responds either by speaking back or by displaying a response on the device's screen. In essence, the assistant is a natural language interface to the device's existing functionality. Figure 5.1 presents three screenshots of the most popular commercial VUIs on smartphones[3] responding to different types of questions. As shown in the dialogue with Siri, responses may contain humour in addition to factualness. Furthermore, in addition to task-oriented questions and instructions, some commercially available VUIs also respond to general questions such as "how are you?" and "what's your favourite colour?", further anthropomorphising the assistant.

Early iterations of VUIs were focused on single tasks, such as Zue et al. (2000)'s JUPITER that was capable of providing weather information. This particular system, as with others at the time, relied on people making telephone calls to interact with it, with the system engaging in dialogue with the interlocutor by talking back in a 'conversational' manner. As network connectivity and accuracy with auto-

3 This is anecdotal based on smartphone ownership numbers.



(a) Siri (b) Google Now (c) Cortana

Figure 5.1: Example screenshots of different 'Intelligent Personal Assistants' on commercially-available smartphones. The screenshots of Siri, Google Now, and Cortana are copyright Apple Inc., Google Inc., and Microsoft Corporation respectively. These screenshots were taken in 2016.

matic speech recognition improved, VUIs, such as InCa (Kadous and Sammut, 2004), were able to operate on portable mobile devices by making use of remote computing power and wireless communication technologies. VUIs are now readily found on many devices such as smartphones, tablets, watches, and even televisions. Additionally, although such systems fail to mimic human talk fully, Pelikan and Broth (2016) were able to reveal the succinctness of how people adapt their talk to an assistant's needs and capabilities, making their interactions more successful. Their work focused on a dyadic face-to-face conversation with a humanoid robot and was able to reveal a number of difficulties individuals face in such talk. In this work, there is a pivot to considering how this talk unfolds as a situated action within a multi-party conversation. A number of pieces of work focused on VUIs have suggested a number of positive aspects in order to justify their development further. In one case, Jones et al. (2014) describe how a voice-controlled personal assistant could be used to support collaboration amongst those gathered around an interactive smart

table, or for use in hands-free or 'eyes-free' interaction while driving a car (Cycil et al., 2013). Others such as Luger and Sellen (2016), however, paint a more challenging picture. Through interviews, they found that there still exists a "gulf between user expectation and experience" (ibid., p. 1) with existing conversational agents and user expectations. This gulf stems from people's perceptions that such systems should deliver more than they presently do and for issues with the VUI communicating system functionality. Innovations to address this gulf include features such as displaying understood text on a screen, voice typing (Kumar et al., 2012) (i.e. live dictation), and the grounding (i.e. affirmation) of spoken input through responses (Clark and Brennan, 1991; McTear et al., 2016), although peoples' reported experiences suggest that numerous problems still remain. By exploring the interactional accomplishment of using VUIs on smartphone devices *in vivo* within a social gathering, and by employing a conversation analytic approach, a rich description of the collaborative work performed by members that occurs in, through, and around the VUI use can thus be explicated.

5.2 STUDY DESIGN

A brief description of the setting in which the observations were undertaken is provided below, including details about the participants, and also the rationale for continuing to adopt an ethnomethodological and analytic orientation. The study was approved by the university's School of Computer Science Research Ethics Committee.

5.2.1 *The café as a study setting*

This chapter extends the work begun in the study of pub talk around touchscreen-based device use in the prior chapter by pivoting to exploring talk in a café around smartphone-based VUI use. In order to situate the study, a casual-social setting was chosen (see 4.2.1) to con-

duct a number of observations of friends socialising together. As discussed, this study deviates from the prior examination of touchscreen-based device use by *requesting* participants to use the voice-based personal assistant on their device instead of typing. All participants were required to have used the assistant before the study. It was supposed that given the identical premise of participant activity (i.e. of a gathering to socialise and relax), that the difference in setting would not be influential on the participants' use of devices (nevertheless, such a concern was not a primacy in this research given the analytic orientation, as discussed in the next section).

Although the particulars of activities that take place in a café may differ from that of a pub, to those who choose to gather there as a group of friends, the purposes remain the same: to socialise and relax together. In other words, although the food may be different and there may be less sport or alcohol, the purposes for being in the setting as a group remain consistent. Indeed, cafés were also included within the definition of "third places" (Oldenburg, 1989), from which this thesis' more encompassing definition of a casual-social setting was formed. Oldenburg defined third places as spaces that are outside the home or work environment that support gathering, socialising, and relaxation for groups and individuals, and with this definition, a café is the epitome of a such a setting. Work that has studied interaction in cafés remarks upon how they provide a "common code of conduct" (Laurier et al., 2001, p. 210) that is both informal yet still provides a guidance of behaviour that is adhered to by members, but typically lack "complex articulation and coordination work" found in more formal settings (ibid., p. 222). Laurier, a number of years later, in studying interactions in a café, introduces cafés in the UK and USA as "familiar nodes in the network of gathering places that remain a necessity for the accidental tourists, be they business executives, chefs or mathematicians, who shuttle back and forth stitching regions together" (Laurier, 2008b, p. 5). It is through this text that the significance of cafés is realised as spaces in which different people in

different situations gather for various purposes. Indeed, during the observations that took place for the work in this thesis, the café received a constant stream of visitors, some merely taking drinks to go, and some meeting up with others, some brought children and some arrived as groups of friends. Therefore, such a setting is perspicuous to observe a gathering of friends socialising.

5.2.2 *Collecting data in the café*

In the study, a neighbourhood café was selected that served hot and cold food, cakes, and drinks. The café is in a residential suburb of Nottingham, within a pavilion at a local park and nearby to schools and a university. Suitable times for observations were agreed between the café and participants, allowing for video and audio recording of the friends talking during a gathering lasting up to ninety minutes. All sessions were recorded on weekday afternoons when the café was open to the public. Video capture was completed by two fixed wide-angle cameras on tripods with an audio recorder placed on the table to allow for clearer capture of talk between the participants.

Groups of friends were recruited via email and social media to visit the café together for the purposes of socialising. Prior to the study, participants were asked whether they had previously used a personal assistant on their mobile device, although there was no frequency or expertise required by them in order to take part. Three groups of four friends were recruited to go to the café together over a two-month period. Seven participants self-identified as male, and five as female; they ranged in aged from 22 to 37. All participants gave informed consent and were reimbursed for their time with a shopping voucher. During the studies, all participants drank various drinks, some ate cake, and one brought some light reading with them to do as they were chatting with their friends.

The study approach is most aptly described as participant-observer, with a researcher present at the table conversing with the group

where relevant. The group of friends met the researcher at the café and were asked to complete a consent form prior to data capture. They were free to move about in the café although primarily sat around a single table as they socialised, drank, and ate cake with each other. For the study, participants were asked to preferably use the personal assistant on their mobile devices instead of typing where possible[4] adjusting the study to be as close to 'natural' as feasibly possible to that of a study of interactions that are somewhat prescribed by very nature of the enquiry. However, the methodological approach to studying interaction remains steadfast given the analytic orientation to the accomplishment of members in and through interaction, irrespective of *reasons why*.

As per Chapter 4, there was no requirement to use a device, and there were no tasks set for the friends to perform during the study. The idea of curating a number of tasks for groups to perform with VUIs during the sessions was considered, however following a pilot study in which participants were given 'free reign' on what activities to perform during the study, and told to converse as they normally would, it was concluded that this was not needed—-people still chose to use the VUIs on their devices. Therefore, participants were simply asked that they socialise and when the opportunity arose, they use a VUI instead of typing, if appropriate. After the study, a number of informal questions to gauge feedback and gather personal perspectives on the use of VUIs were asked. However, this group interview was used as a debriefing exercise rather than to shape the findings.

5.2.3 *Analysing the collected data*

To analyse the collected corpus, as with the prior study, an analysis shaped by ethnomethodology (Garfinkel, 1967; Sacks et al., 1974) was performed. Through this, the orderly and situated practice of using

4 The information sheet provided to participants prior to the study is included in Appendix D.1

VUIs in conversation was explicated. This analysis required the watching of the collected corpus multiple times, in order to segment and identify relevant fragments of data consisting of VUI use. Firstly, fragments were watched (and re-watched), with the methodical actions of members within the setting catalogued and indexed to identify instances where a mobile device and a VUI was used. Timestamps and descriptive language were used to construct a record of the interactions that took place, which allowed for iterative re-examining of prior data with relative ease to help gain an overall impression of the data collected across all the sessions.

Three fragments were selected for presentation in this chapter. In line with best practice, as summarised by Heath et al. (2010), each fragment progressively reveals the organisation of interaction with and around the use of the device:

> As you build an argument the analysis should be progressively revealed and emerge by virtue of the presentation of each successive extract. Furthermore the fragments should become more delicate and complex such that the audience can learn how to see the phenomena and can follow the argument as it unfolds.
>
> — Heath et al. (ibid., p. 111)

In the three fragments to be presented, the first illustrates a typical use of a VUI on a smartphone to introduce new information to the conversation that will set the scene for the interactions that unfold. The second fragment introduces a more complex case that reveals how users perform additional actions to get the VUI to work as desired. The final fragment introduces an interactionally problematic case where a member uses the VUI to establish its capability, revealing the richness and complexity of interaction around the use of a VUI in a café.

The work was oriented to unpacking the retrospective-prospective character (see 3.2.2, Garfinkel (1967, pp. 35–75)) of members accomplishing the work of using a VUI in this setting, in and through their

ongoing social interaction. This orientation necessitated the identification of interactional accomplishments that occasioned the use of the VUI. This included how the device was introduced, the instruction or question that formed the request to the VUI, the actions (in-talk and physical movements) of members in the setting throughout the activity, and so on. In other words, this orientation, and the resulting analysis, allowed for the formation of a comprehensive understanding of the activities performed by members in using the VUI.

In total, 40 episodes of VUI use in the sessions were identified (some of which were overlapping), with episodes ranging from a few seconds to a nearly five minutes in length. A substantive review of the episodes was performed to examine the interaction that unfolded, honing in on episodes that represented observable-reportable intersections of the use of the VUI and conversation for a more in-depth analysis. Six fragments were then transcribed with both verbal (i.e. talk) and non-verbal actions (e.g. gestures and other interactional resources) being carefully noted. These fragments were selected in line with the aims of this research to reveal the social organisation of device use in the use of VUIs in and through conversation and that were deemed to warrant further investigation.

5.3 FINDINGS

Data from the fieldwork will now be introduced and presented over a series of 'data excerpts'—these excerpts form fragments that *vividly exhibit* (Crabtree et al., 2012) the actions of members within the collected corpus. Each of these fragments exhibits the activities of members' observed practices of how members made use of the VUIs on their personal devices in the café.

In total, the corpus consists of 123 utterances to VUIs by members, across 40 distinct episodes of data from a corpus consisting of 3.6 hours of video data. In particular, this chapter will reveal (1) how members perform a request with their device, (2) how members ori-

ent to and appropriately deal with the request and the VUI's response to the request, and (3) how members collaborate through the interaction with the VUI.

Appendix B provides details of the transcript notation used in this thesis. All names and identifiable information within the transcripts provided are entirely fictional.

5.3.1 *Introducing new information for conversation*

The first fragment, called *When Does the Sun Go Down?*[5], that commences in Data Excerpt 5.1, consists of four friends: Arthur, Harry, Sally, Julia, and the researcher. The friends are meeting late afternoon during winter and the sun is shining into Harry's eyes. He holds his hands in front of his eyes although he refuses to move because he will "`...be fine in like three minutes`" (line 01). The friends joke about this experience, and that this forms part of their study (lines 08–15), but this challenge of the blinding sunlight establishes the interactional project that ensues. Later in the next excerpt, Julia uses her iPad, which she had out on the table already, to find the time of sunset as a result.

```
01  HAR   i'll be fine in like three minutes ((holds hands in front of
02        eyes))
03  RES   keeps coming back as well like
04  SAL   as soon as you change it comes back
05  JUL   yeah yeaha
06      (0.3)
07  RES   there's actually just someone out there with a light!
08  ALL   ((laugh))
```

Data Excerpt 5.1: When Does the Sun Go Down? (i)

In this opening excerpt, the group jokes about the sun shining through a nearby window into Harry's eyes as a result of the time of day (approximately evening time, around sunset). This occasions, as will become evident in the following excerpts, Julia to use her VUI

5 The complete fragment is included in Appendix E.1.

Figure 5.2: HAR blocks the sunlight in his eyes (Data Excerpt 5.1: When Does the Sun Go Down? (i), line 1)

to retrieve the time of *sunset* for this given location. This work to establish the time of sunset addresses the ongoing concern established in conversation about the sun in Harry's eyes, and responds to his remark that he will "`be fine in like three minutes`" (line 01), with the use of the VUI commenced by Julia to introduce new information to *verify* his claim. The next two sections, centred around two excerpts that follow Julia as she seeks new information for the conversation:

(i) Addressing the question to the VUI, and

(ii) Collaboratively finding new information.

As will be evident through the sequential revelation of members' actions, the VUI will be introduced and used with the address of a question occasioned in and through the conversation. The nature of multi-party conversation and making a device interaction accountable will allow members to engage with the interaction at hand in the completion of one member's interactional project to verify the time of sunset.

5.3.1.1 *Addressing the question to the VUI*

The discussion continues before the next excerpt, Data Excerpt 5.2, picks up the conversation.

```
16  JUL   [ ((removes cover from device but leaves open)) ]
17  SAL   ((laughs))
18  JUL   ((presses button on device))
19  HAR   there we go!
20  JUL   what's the time of sunset?
21      (1.3)
22  ALL   ((gaze at the tablet))
23      (3.0)
24  JUL   ok! // ((device displays clock)) //
25  ART   ((leans in to look))
26  SAL   that's [ a      ] fucking analogue clock it pisses me off!
27  HAR          [ today? ]
28  HAR   ilunno (0.6) 24 hour=
29  JUL   <no no no!> it misunderstood actually (0.8) understood what's the
30          [ time ]
31  HAR     [ time ] now
```

Data Excerpt 5.2: When Does the Sun Go Down? (ii)

Figure 5.3: All members gaze at the tablet (Data Excerpt 5.2: When Does the Sun Go Down? (ii), line 25)

First, this excerpt is unpacked. As the laughter begins to die down within the group at the start of the excerpt, Julia removes the cover from her device, which she has on the table (line 16); she waits for the group laughter to die down, presses the home button (line 18), and be-

gins her utterance as Harry finishes remarking that the sun has now moved (line 19). Julia uses her device's VUI to ask for "`the time of sunset`" (line 19). As she does this, she fixes her gaze at the screen, which displays a loading animation while the device is awaiting input. At this point, all members lean in towards the device, as shown in the image, and demonstrate an awareness of an impending response from the VUI. As she began her request, Harry utters "`there we go`" (line 18) in direct response to the sun no longer coming through the window directly into his eyes (inferred by his shift in posture, including no longer covering his eyes with his hands). Nevertheless, Julia pressed ahead with the question, having begun the performance of preparing the device for the address by pressing the button.

After a few moments, the VUI returns the time for the local area as an analogue clock. A number of comments on this are passed: Sally comments on the presentation of the time (line 26) and Harry questions if that is for the present day (line 27). Julia then interrupts the talk and retorts that she has realised the device has "`misunderstood actually`" (line 29) and that the VUI is presenting the current time, not the time of sunset.

By explicating the distinct sequential actions taken by members as the device is used, the ways in which members practically reason about how a VUI responds to a request and attend to the VUI's response become evident (e.g. by leaning in, rotating gaze, not overtalking the VUI or VUI-user). In this exhibit, Julia reasons about the failed outcome of the request to the VUI through examining the displayed response, and makes this accountable to all (line 25) through her verbal report. Her position and access to the device affords her greater visibility (as seen in the image in Figure 5.3) of the response from the device, which typically displays the 'transcribed' text of what the device 'understood'. Through her vocalised interpretation of the output of the VUI, she provides an explanation for the problem source—or rather, starts to—as she realises it "`understood what's the--`" (line 29) and Harry, who seemed to question the answer (line 27)

completes her sentence with "time now" (line 31). Harry's completion of Julia's utterance exemplifies that he is in accord with her reported interpretation of the source of technical trouble.

Through the ongoing interaction, as will be revealed in the next excerpt, members collaboratively reason that the response was not as expected and that this must be because the transcription of the request by the VUI was wrong.

5.3.1.2 *Collaboratively finding new information*

Julia's previous assessment that the device was showing the current time, as opposed to the time of sunset, (lines 30—32) leads to a proposal to ask a different question (line 36) to the device from Julia. In turn, the members collaboratively find words to return a successful result, as examined in this next excerpt, with Julia proposing a follow-up question, and Harry providing a suggestion of the wording she should use (line 35).

```
32  JUL   so-
33  ART   soaoah yeah↑
34  JUL   shall i ask (1.6) um:=
35  HAR                        =what time will the [ sun set? ]
36  JUL                                            [ ((holds button)) ]
37  JUL   // ((audible chime)) //
38      (4.0)
39  JUL    // ((on screen text: go ahead i'm listening...)) //
40      (0.3)
41  JUL   when does the sun go down?
42      (2.9)
43  JUL   sunset will be at [ seventeen thirty two ]
44  ART                     [ ther:::e you go      ]
```

Data Excerpt 5.3: When Does the Sun Go Down? (iii)

In this final excerpt, Harry proposes a slightly different question (line 35) although ultimately Julia asks "when does the sun go down?" (line 41), to which the VUI provides an accepted answer as revealed in Arthur's comment of "there you go" (line 44). Although the time is visible on the screen and all members look at it, Julia also provides an aud-

ible report of the time displayed (emblematic of the *hybrid* nature of VUI interaction on a smartphone also making use of the screen). This excerpt reveals how members can collaborate on a project by working together to use a device, exemplified in this fragment through the demonstrable reasoning of suggesting the cause of trouble and through reformulating the request by the members. In the fragment, across the three excerpts, Julia interprets the initial result from the VUI as incorrect (line 29), but then reasons about the response, reveals her reasoning to the group, and then asks the VUI the same question with a different lexical construction (line 41). In this, she does not just retry or repeat the same request, however, she rephrases—with the presumed aim of soliciting a successful answer from the VUI, as per the occasioned purpose of her actions.

Rephrasing a request, as occurred here, was a practice used by members to attend to and deal with troubles with VUIs responding incorrectly. While in this case, the original interlocutor rephrased the request, on other occasions members other than the original VUI-user may also have performed a rephrased version of the original request on their own device. Therefore, it is posited that there is a distinction in the occasioning of rephrased requests and repeated requests. Rephrasing occurs as members attend to a VUI not completing their request as a result of the VUI not completing a transcribed request as expected, e.g. as Julia informs others in the setting (line 29). Given members' use of VUIs is for a specific purpose, members undertake and extend their occasioned activity of using the VUI until receiving a response that accomplishes their goal.

Finally, this fragment concludes by remarking upon the collaborative and coordinated activity that occurs throughout this fragment. The members collectively shift their body posture so that they are looking towards the device being used by rotating their torsos and leaning across the table. Additionally, they mutually pause their talk while requests are performed and responses computed, they gaze at the tablet, and they attend to the answer as soon as it is provided—i.e.

they work together to complete the request. In this entire sequence, the request is interactionally occasioned in and through the conversation about the sun shining into Harry's eyes. The other members then witness the request being performed (line 20), and the failure of the device to respond appropriately is made accountable by making the screen clearly visible to all members, such that the other members can see and practically reason about the result. This, in turn, allows for members to collaboratively reason about the grounds of the VUI's failure (lines 29–35). In attending to the failure, the members then construct a further request which leads to a satisfactory result. Given that members accomplish the natural accountability of performing a request with a VUI through conversation, it appears that the practice of rephrasing a request lends itself as a resource to support collaborative activity amongst the co-present members.

In this fragment, a device was brought into the conversation in order to introduce new information to the group in relation to a problem established through interaction. Although the VUI user encountered *technical* trouble, through suggestions from others, the member was able to complete their request and introduce the new information to the conversation. Given the hybridity of interactions with VUIs on smartphones also featuring a GUI, this introduction of information occurred through members looking at the information displayed on the screen of the device. The next fragment moves beyond an example where new information is introduced into conversation to a situation where members are trying to answer a proposed question in talk. At face value, this fragment seems to occur in the same manner as the first one, although through unpacking the data, interaction will be shown to be replete with challenges as members attend to problematic VUI interaction.

5.3.1.3 *Methodical accomplishments in this fragment*

With this fragment, the work of introducing new information for conversation was unpacked over a series of excerpts. The question was occasioned in and through the conversation, as a result of the sun shining into Harry's eyes, although Harry dismissed the problem as he expected the sun to set shortly. Julia ostensibly uses this moment to ask the VUI on her device the time that the sun will set, to determine the veracity of Harry's claim. She does this by **preparing to use the VUI** by removing the device's cover and holding the button down to activate the VUI. She then **addresses the VUI** through talking to the device with her question and then **looks at the device** as it computes a response. The co-present others also **lean in to look at the device screen** as an analogue clock is displayed, reasoning about the perceived incorrectness of what is displayed. Harry and Julia **propose that the incorrect information** is displayed as a result of the device 'misunderstanding' the request, and **propose a new wording**. Julia **re-addresses the VUI** through talking to the device again with a rephrased request, **looks at the device** along with the others as the response is computed, and then **provides a verbal report** once an answer is given. This successfully completes the interactional project which was occasioned.

5.3.2 *Answering a question in conversation*

The opening excerpt from the second fragment, titled *Do Animals Have Accents?*[6], is given in Data Excerpt 5.4. This excerpt unfolds as four friends: Lily, Gary, Karl, Antonius, and the researcher, are socialising together. The group, which consists of members from the UK, Romania, and Austria, have been discussing the different onomatopoeic sounds that various animals make and how these sounds vary by country and language.

6 The complete fragment is included in Appendix E.2.

```
05   KAR    do cats acth- (0.5) can you work out whether it's french because
06          because its talking in a- doing a french cat impression
07   LIL    i::::: think some animals you can
08      (1.9)
09   LIL    ((picks up phone from table))
```

Data Excerpt 5.4: Do Animals Have Accents? (i)

There are presently two conversation floors[7] taking place in the conversation: in the floor focused on here, Karl asks Lily about animal accents before recounting scenes from a television show to Lily (omitted from this thesis for clarity), and in the other floor Antonius is recalling the sounds different animals make when uttered in Austrian German. Just before Karl begins to recount his story, Lily picks up her smartphone (line 09) and begins to type with the on-screen keyboard throughout the story. In the case of this fragment, the actions of members will be shown to consist:

(i) Addressing the question to the VUI,

(ii) Repeating a request to the VUI, and

(iii) Getting others to perform the request.

These three activities are now unpacked below. As will be evident through the sequential revelation of the action, members collaboratively work together to undertake action in their orientation to the problem that occasioned the use of the VUI. Given the nature of multiparty conversation, these activities will be shown to be recurrent and overlapping with each other, rather than discrete temporally-ordered accomplishments.

7 In other words, within the group there are two discussions continuing in parallel with those co-present orienting to one conversation at a time (Edelsky, 1981).

5.3.2.1 *Addressing the question to the VUI*

After the story in which Karl recounts an episode of a television show, both he and Lily laugh and then he orients to and engages with the other floor; he does this by shifting his gaze to look at the others in the other conversational floor (specifically Antonius, who is talking at that moment)[8]. At this point, Lily moves her smartphone closer to her mouth and asks her VUI "`do animals have accents?`" (line 42). This question was not specifically asked in talk but arises as a result of the topic that all the members have focused on in both floors at some point—in other words, the work of answering this question was occasioned in (and as a direct consequence of) the conversational topic. Following Lily's request, a short gap in talk (line 43) unfolds before Gary shifts his gaze to Lily and responds to her question, as shown in the video still captured at line 44, even though her question was aimed at her VUI.

```
40  LIL   er:::m: ((holding phone in front of her at chest level))
41      (3.7)
42  LIL   ((moves phone up to face)) do animals have accents?
43      (2.1)
44  GAR   ((shifts gaze to LIL))
45        yes they do actually! i think i've read something
46  LIL   i think i have [ too↓  ]
47  GAR                 [ yeas! ] [ (0.6) cows! i- i              ]
48                                  read about cows that they have
49                                  different accents around the world
50  KAR                             [ you missed mine- my racist joke ]
```

Data Excerpt 5.5: Do Animals Have Accents? (ii)

In this second excerpt, Lily prepares to perform an utterance (line 40–42) and performs her utterance (line 42), first by moving the phone closer to her mouth with the handset held such that the microphone is in front of her lips, and then performing her utterance while there

8 In other words, Karl shifts his gaze and body posture away from the members he was previously conversing with the others at the table to those he was not conversing with, but who were conversing with each other in parallel.

Figure 5.4: Lily performs the request (Data Excerpt 5.5: Do Animals Have Accents? (ii), line 44)

remains a gap in talk. Following this utterance, Gary rotates his head from Antonius to Lily and responds to her question: "`yes they do actually`" (line 43). Karl and the researcher also shift their gaze towards Lily, in the case of Karl by leaning back slightly and rotating his head, while the researcher rotates his head alone. Through his response to Lily's question, Gary exemplifies the manner in which interactions with a VUI are made accountable insomuch that others can observe, report upon, and respond to the device interaction accordingly—in this case, to answer the question Lily asked. In this sense, this reveals how the preparatory action of talking to a VUI on a personal device is made naturally accountable, and can be oriented to by co-present others as a matter occasioning further discussion.

Following this response by Gary to Lily's address to her VUI, Lily affirms that she too had heard something[9]. In this, Lily implies that her question was guided by her being unsure ("`i think`", line 46), but occasioned by both the conversation and recent news events that she had read about; with Gary acknowledging he was aware of the story too (lines 46–49).

9 A story regarding cows with regional accents had been on the BBC News website a few days prior to this gathering.

5.3.2.2 *Repeating a request to the VUI*

Following this brief discussion between Lily and Gary following Lily's VUI use, Lily has been looking at her device sporadically while talking to Gary. She then lifts the device closer to her face again, and performs a new request to the device, which is indicated in the commencement on the next excerpt, Data Excerpt 5.6. In this next excerpt, the work of how members *respond to the VUI's response* by repeating the request in order to accomplish their task is demonstrated.

```
51  LIL   DO: ANIMALS HAVE ACCENTS!
52      (2.4)
53  LIL   °rubbish°=
54  KAR             =parrots presumably do=
```

Data Excerpt 5.6: Do Animals Have Accents? (iii)

This short excerpt features the first attempt at responding to the VUI's perceived failure to perform as expected, which in this case is done by repeating the request with greater volume and impetus (differing from the previous fragment in which a rephrased request was made, see 5.3.1.2). As Lily does this, a gap in talk occurs (line 52) as other members look at Lily; she then quietly utters "rubbish" (line 53). In and through this utterance she further accounts that the device has failed to adequately respond to the request put forward in her address to the device. The repetition of the phrase, and increasing volume makes evident the device's failure to 'hear' what is said. Here, this demarcates a different problem with VUIs—not only do such devices have trouble responding correctly to what is said, at times they may not 'hear' words at all.

5.3.2.3 *Getting others to perform the request*

The final excerpt from this fragment, given in Data Excerpt 5.7, concludes Lily's efforts to find the answer to the question of whether animals have accents. In this next excerpt, Lily asks Karl to help her complete the request, which she does by passing the responsibility of

uttering the request to Karl by holding the device in front of Karl and questioning whether he could "ask it" (line 55, shown in Figure 5.5).

55	LIL	=can you ask it?
56		((holds phone out in front of KAR's face))
57	RES	((retrieves phone out of pocket))
58	KAR	**DO: ANIMALS HAVE ACC::ENTS!**
59		(0.9)
60	LIL	no:!
61	RES	*// sorry i'm- //*
62	RES	((RES touches screen to stop utterance))
63	RES	**do animals have accents?**
64	LIL	**do: animals have accents?**
65	RES	*// ok i've found this on the web //* (sigh)
66	GAR	do [they?]
67	LIL	[ah (.)] it's working now!

Data Excerpt 5.7: Do Animals Have Accents? (iv)

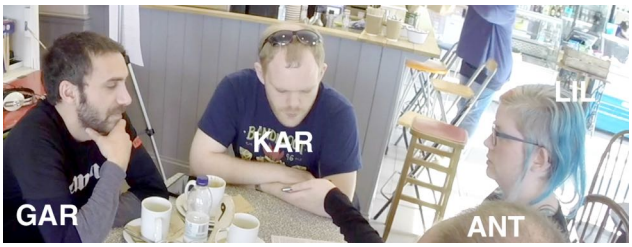


Figure 5.5: Lily asks Karl to perform the request (Data Excerpt 5.6: Do Animals Have Accents? (iv), line 56)

At the moment where Lily holds her device out to Karl and asks him to "ask it" (line 55), the researcher retrieves his phone from his pocket to perform the request (line 57). Karl's request also fails, as revealed through Lily's next attempt (line 64), which this time yields search results, as does the researcher's (line 63). Following the fragment, both begin to share information retrieved from webpages to the other members of the group. This fragment is indicative of what VUI use looks like—-that is, there are a number of grossly observable features that take place: there is ongoing selecting of speakers, repetition of requests, pauses in talk, body co-orientation and so on.

Importantly, one aspect revealed through close examination of the interaction is that once a request is performed with a VUI, members' talk to VUIs may be sequentially followed by pauses in talk (lines 43, 52, 59), as members visibly orient their gaze towards a device as a response is computed.

The excerpts in this fragment reveal how, as a practice, talking to a VUI in turn occasions the mutual production of silence by the co-present members as they re-orient to the use of the VUI, and in turn focus on the device or the interlocutor. Members do not pause their interaction or 'sit in silence' however, their embodied actions of gaze and body co-orientation furnish others with how they are focusing their attention, as they turn to the device interaction. In effect, performing a request brings about a lapse (Hoey, 2015) in the conversation: neither the member who was performing the request selects to talk next, nor does any other member. VUIs function by assuming a pause-in-talk specifies the completion of a request, thus a pause by the interlocutor is necessary. Therefore, it is noted that the data reveals that the action of performing a request with a VUI may prescribe a lapse in talk and the mutual production of silence. This action is, by definition, intrinsically and demonstrably collaborative—members collectively and collaboratively socially construct silence in orientation to the utterance to the VUI.

Furthermore, this fragment features numerous repeated requests to VUIs: Lily gets Karl to perform the request as a result of her device's repeated failure. In the fragment, Lily uses her smartphone on multiple instances to perform the request "`do animals have accents?`" (lines 42, 51, 58, 63, 64), each time with more impetus in her voice. With each repeated request, Lily accounts for the device's failure to appropriately respond to her initial request: either the device has mistranscribed, or it has not transcribed at all, and so another attempt is required to complete the task at hand. Whereas in the prior fragment, members were shown to revise or *rephrase* a request to get the VUI to work, here we see another action of members to accomplish the task at hand with

a VUI: repeating the request in cases where it seems a VUI has not 'heard' the request. Thus, a finding is that members address a problematic interaction with a VUI through the further production of talk: *they repeat their request*, and that others may assist in this completion.

Finally, Lily's actions, of talking with greater impetus and volume, and then getting someone else to talk, account that for her, the perceived source of device trouble is that the device cannot transcribe her utterances because of her diction[10], which results in her involving Karl in the task of talking to the VUI to complete the request. Her practice of involving Karl in the request has the interactional outcome of recruitment to resolve the trouble at hand with the device interaction, a fundamentally collaborative mechanism of social organisation (Kendrick and Drew, 2016)[11].

5.3.2.4 *Methodical accomplishments in this fragment*

The second fragment in this chapter reveals the work of answering a question in conversation using a VUI. Lily **prepares to use her device** to answer the question Karl has asked, **lifting her phone to her mouth** and **addressing the VUI** through talk. The device ostensibly does not respond; therefore Lily **re-address the VUI with a higher volume and impetus in her delivery but using the same question**. The device again 'fails' to respond; thus Lily **asks another co-present member, Karl, to perform the request**. He then **addresses the VUI with a loud speaking voice and different emphasis** compared to Lily, but with the exact same question—again there is no response from the device. The researcher then **retrieves his phone** and **prepares his VUI by holding the button down**. He **addresses his VUI** with the same question while **Lily also addresses** her VUI again. Both

10 Lily's reasoning of the trouble source is made visible as a common-sense understanding to other members of the setting, including Karl, through the second production of the request that is produced louder and with greater impetus.

11 Kendrick and Drew (2016) refer to recruitment not as single or class of social actions, but as an interactional outcome achieved through methodical practices such as requesting or offering assistance.

members then **verbally report that their devices have computed responses** and then go on to **discuss the findings**. In this regard, the interactional project of using a VUI to answer a question in conversation is completed.

5.3.3 *Establishing the capability of a VUI*

The first fragment in this chapter identified how members of the setting introduce new information into a conversation by using a VUI to retrieve the information. The second further identified how members are able to direct specific questions raised in a conversation to a VUI, to answer and address information deficits amongst those who are co-present. This final fragment, which commences in Data Excerpt 5.8 and is called *Hey Siri! . . . Call My Mother*[12], focuses on the researcher and Gary discussing the capability and features of the VUI on his device, again using the VUI to address an information deficit. However, whereas the first two fragments identify how members are able to use a VUI in conversation to address problems established in and through conversation, this fragment deviates from this and reveals how VUI use may be occasioned self-referentially in order to address matters of VUI capability[13]. First, this fragment will be introduced, with the problem at hand established, and then the fragment will be unpacked as two distinct activities, in which members undertake:

(i) Establishing the desired function, and

(ii) Testing the functionality by addressing the VUI.

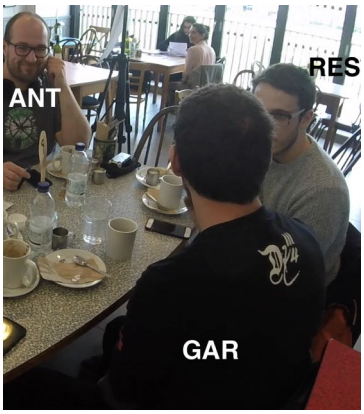
These two activities are now unpacked in the following sections to reveal the problem case of whether the device is able to deal with a similar but non-English spelling of contact details.

12 The complete fragment is included in Appendix E.3.

13 This thesis strays away from attributing a reason beyond what is naturally accountable, but the recentness with which VUIs became available at the time of this study would suggest that there is not the same level of familiarity with their capabilities as with basic smartphone interactions completed using the GUI.

5.3.3.1 *Establishing the desired function*

The premise of the problem established through the conversation between the pair is whether the device's VUI is capable of matching the non-English spelling of 'mama' in Gary's smartphone's address book, with the discussion predicated around the focus of the study. The conversation commences in Data Excerpt 5.8, in a separate floor consisting of just Gary and the researcher, who are both sitting next to each other. The other members of the setting are conversing while the two discuss Gary's interactions with his device.

01		GAR	i'm curious if I say in
02			romanian (.) to call my mother
03			(0.7)
04		GAR	it will actually find the
05			contact for my mother is (.)
06			mama in romanian (.)
			(...)
12		RES	cos you can also tell people
13			who they (.) like you can say
14			like
15		←GAR	**hey siri=**
16		RES	=my mother is this
17			person (0.8)

Data Excerpt 5.8: Hey Siri! . . . Call My Mother, part (i)

As this exhibit commences, Gary questions, by way of asking the researcher indirectly, whether if he asks his device to call his mother, the device will recognise the name in his contact list (the contact's name is spelt 'mama' in Romanian). The researcher responds in broken English, alluding to a feature with VUIs and smartphones that allow for pseudonyms to be allocated to contacts, although his utterance

is punctuated by Gary's initial commencement of device use, using the hotword "hey siri" (line 15). Without shifting his gaze from the researcher who is speaking, Gary lifts his phone and performs this phrase and then moves the device back to chest height between him and the researcher, interleaving the opening utterance of the hotword within their conversation.

Gary's initial proposition (lines 01—06) establishes that he is curious about the capability of the VUI and its ability to function as desired. This is followed by the researcher discussing a feature that would potentially allow Gary to request a device to call his mother, using an alternative name to the name given to the contact in his phone's address book (lines 12—14). As such, the researcher does not directly respond to Gary's question, but makes a proposition of a technical solution to the problem at hand, by adding an English pseudonym to the contact.

5.3.3.2 *Testing the functionality by addressing the VUI*

Following the continuation of the researcher's utterance regarding the pseudonym functionality (lines 16—17), Gary glances at his device and re-performs the hotword to activate his VUI. This next exhibit, Data Excerpt 5.9, reveals how Gary tests the functionality of the VUI with his desired capability, which allows him to demonstrate, either successfully or not, his suspicions of a limitation of the device to deal with alternative spellings of words such as *mamma*.

18	GAR	((glances down at screen))
19		((moves device in front of
20		mouth)) **hey siri**
21	GAR	((moves device to chest
22		height between the two))
23		(1.0)
24	RES	i'd press the button
25		(1.2)
26	GAR	((moves device in front of
27		mouth)) **hey siri**
28	GAR	((moves device to chest
29		height between the two))

30			(2.4)
31		GAR	((moves device in front of
32			mouth)) **call my mother**
33	←		((GAR and RES look at screen))
34			(5.9)
35			*// what is your mother's*
36			*name? //*
37		RES	((points towards screen)) yeah
38			but then
39			(0.9)
40		GAR	**my mother is mama**
41		GAR	*// i can't find anyone called*
42			*mamma //*

Data Excerpt 5.9: Hey Siri! . . . Call My Mother (ii)

In this excerpt, Gary glances down at his device, with the screen still dark, and so he lifts his device again and re-utters "`hey siri`" (lines 18–20). He holds his phone in a position that the researcher can see at waist height in front of him, although this time his gaze remains on the device awaiting a result. The researcher proffers advice based on his personal experience that he would "`press the button`" (line 24), although a moment later Gary (successfully) activates the VUI through a further utterance of "`hey siri`" (line 27) and then asks his VUI to call his mother (line 32). Consider the sequence of Gary's performance of this hotword: there are repeated pauses after his utterances to the VUI (lines 15, 20, 27, 32) as Gary provides the utterance and waits for the device to respond by looking at the screen.

On completing his 'test' request to the device, Gary returns to holding the device at chest level between him and his co-interlocutor, as can be seen in the image within Data Excerpt 5.9, allowing both members' direct line of sight of the screen. After nearly six seconds of both partners looking at the screen between them, the device seeks further information of the name of his mother in his address book—Gary provides this by telling the device that the contact card for his mother is "`mama`" (line 40). This, however, does not work as the device searches for contacts named "mamma" and does not seemingly look for altern-

ative spellings such as "mama". The use of his VUI is ended after this, underscoring the problem at hand that occasioned the use of the VUI was to *test and verify the functionality of the VUI* rather than to actually call his mother.

5.3.3.3 *Methodical accomplishments in this fragment*

This final fragment reveals the work of establishing the capability of a VUI through interaction with it. Gary establishes the capability of which he is curious, i.e. of whether the VUI can call his mother. He does this by questioning whether the VUI will be able to respond to the request in conversation. His co-interlocutor beings to respond to this question, although Gary prepares the use of his VUI nevertheless by lifting his phone to his mouth and performing the hotword. He does this several times as the device ostensibly does not respond to his attempts to activate the VUI. Once activated, he then tests the functionality of the VUI by addressing it with the request. The device seeks clarification of which contact details correspond to his mother, which Gary provides through a further address. The device returns a result audibly that confirms the device cannot do as Gary requested. Although the VUI 'failed' to meet the expectations of completing the task of calling his mother—the question addressed to it—for the interactional project of establishing the capability of the VUI, the interaction was ostensibly a 'success' to the members involved.

5.4 CHAPTER SUMMARY

This study differs from the first study of mobile device use in Chapter 4 by prescribing that members adopt a preference for using the VUI on their device instead of interacting using the touchscreen. The analysis of the video-recorded interactions, however, were, for the most part, identical in practice. The first study identified how members used their device to respond to problems arising in conversation, such as to introduce new information to the conversation (see 4.3.1) or to

make jokes (see 4.3.2). Furthermore, although it may be argued the use of VUIs was 'not natural', as studied, it becomes a moot argument to consider—the work in this chapter, and indeed this thesis is concerned with *naturally occurring interaction around the device use*, i.e. of how members bring the device into conversation, how others orient (or not) to this, how they attend to the device through interaction, and so on. During the studies, participants were not guided or scripted in how to 'deal' with device interaction as individuals or as a group. Therefore, were this to be a study of *whether* or *why* such voice-based personal device interactions occur, this would certainly be a legitimate limitation. However as this thesis concerns itself with *how* such device use is interleaved and oriented to in conversation, there is no such limitation to consider.

To summarise the findings, by pivoting to an exploration of how friends socialised around the use of the VUIs on their personal devices in a casual-social setting, this chapter revealed how they used the features to introduce new information (see 5.3.1) or answer questions raised in conversation (see 5.3.2). Furthermore, however, an apparent unawareness of capabilities of VUIs was also shown to, in-part, occasion the use of the VUI to explore and test its capabilities (see 5.3.3).

Moreover, as with touchscreen-based device use, these activities were often accomplished while interleaving the use of the VUI with conversation—by performing requests to the device within the conversation, by members in conversation orienting to the device through gaze and production of silence, and by collaboratively dealing with troubles that arise as a result of technical problems with the VUI. What becomes evident, as each successive fragment is unpacked, is that the members make the use of VUIs naturally accountable as part of the ongoing conversation in the setting, in and through the use of the device. Additionally, the hearable request to the VUI occasioned other co-present members' engagement with the VUI user as a result of establishing them accounting for device use in interaction. In contrast with touchscreen-based use, technical troubles with the VUIs

were made accountable in the setting as a result the work of device users' actions, and co-present others responded to these troubles. In this sense, using a VUI was shown to facilitate other members attempting to complete the task on another member's behalf when technical trouble arose (see 5.3.2) or proffer suggestions on addressing technical challenges (see 5.3.3).

In conclusion, as per Chapter 4, the use of devices in the casual-social setting was occasioned in and through conversation and remained perspicuous to the members. Additionally, VUI use was not oriented to as problematic when occasioned, with this again being easily attributed to the very nature of the setting and purpose of the gathering (i.e. an *observational study* of friends socialising). However, in contrast to the prior study of device use, the VUIs on the smartphones were addressed as a topic in-of-itself during observations, with these conversations leading to users to explore the functionality of such systems, revealing challenges that remain at the time of study in relation to users' understandings of the capability of VUIs.

5.4.1 *Methodical accomplishments*

This chapter has revealed the methodical accomplishments of using a VUI within a multi-party conversation. VUI use was occasioned in talk to introduce new information into the conversation, to answer questions in talk, and to establish and test the capability of the VUI itself. Practically, this was done by ostensibly selecting to use ones own device, or asking another person to use the VUI. The device user then prepared the VUI by pressing a button or using the hotword as an initial address to the device. The VUI was addressed in talk through the utterance of a request. If no response was computed by the device, the user reattempted their addressed request with a different impetus and volume. Members would practically reason about the response from the VUI, and if it was not 'correct' would re-address the VUI with a rephrased request. Once the purpose of the VUI was completed,

members would either use their portable device as a normal touch screen or cease the use of the device.

5.4.2 *Outlook*

A number of 'smartspeakers' were released by major technology companies and became commercially available in the UK during the analysis of this study. These devices, which require mains electricity and thus are stationary shared devices rather than portable personal devices, only support interaction through voice and do not have a touchscreen. The VUI in these devices responds to questions, and can 'talk back' to the user, and can also complete basic tasks such as timers, newsflashes, and music playback, as well as connecting to the Internet of Things devices. The next chapter, Chapter 6, will explore how interaction unfolds *in vivo* around the use these of voice-only devices that are communal rather than personal in nature, revealing the details of how family talk around such a device ensues.

6

STUDYING TALK AROUND STANDALONE VUIS IN THE HOME

This final chapter of empirical data presents the study of conversations in households around the use of a voice-only Voice User Interface (VUI) device. The previous work in Chapter 5 explored conversations amongst groups of friends socialising together in a neighbourhood café, where conversations that included the use of VUIs on portable devices such as smartphones and tablets were the focus of the study, and of how this use was interleaved with the conversation. The previous chapter demonstrated how accountable interactions with VUIs led to other users becoming involved in interactions with the device—this chapter further explores how interaction unfolds when that device is communal and all interaction with the device, including its response, is made accountable to those within the setting. This chapter considers how interaction unfolds around a non-personal device (i.e. the device is 'shared' amongst users in a multi-party environment), and without a screen (i.e. it draws entirely upon voice interaction to complete tasks).

To achieve the goal of this thesis, this chapter will study how families make use of a 'smartspeaker' VUI device in their home. Given the variation in setting and interface being studied, this chapter adopts the approach taken with regards to data collection in the prior two chapters by recording audio from participant homes over a month-long period.

This work was previously published and presented at the Computer-Human Interaction (CHI) conference[1]—a number of changes have been made to address the research questions of this thesis.

1 See Porcheron et al. (2018).

6.1 INTRODUCTION

This chapter proceeds in the tradition of Human-Computer Interaction (HCI) and Computer-Supported Cooperative Work (CSCW) research that deploys technology to study the situated and emergent lived experience in the home (Tolmie and Crabtree, 2008; Rooksby et al., 2015). In this way, the work continues upon recent work emerging in CSCW that has begun to examine VUIs in collaborative action[2] for social settings such as meetings (McGregor and Tang, 2017). The study here reports findings from month-long deployments of the Amazon Echo with the 'Alexa' voice agent in five households. Audio capture was selectively performed by a separate device, a Conditional Voice Recorder (CVR). Over six hours of verbal exchanges involving the VUI in some way were collected using this purpose-built device,

A number of studies in homes, primarily focused on the use of mobile devices while another activity is taking place, have followed a variety of different practical approaches to data collection and focus. For example, Schirra et al. (2014) used interviews to examine television watching and the use of Twitter, whereas Jokela et al. (2015) employed a combination of interviews and diary studies. More recently, Ferdous et al. (2016) incorporated home visits and self-controlled video capture of families' technology use during mealtimes. In a similar vein to the approach taken by Ferdous, Rooksby et al. (2015) set up device screen capture technology on homeowners smartphones and installed video cameras in living rooms to capture television watching and the use of mobile devices—families were asked to turn these cameras on when they felt like it. In this latter case, the analysis oriented to the sequentiality of action and methodically revealed how members routinely embed their use of mobile devices to enhance leisure time socialising around the television. Both studies reveal crucial

2 As part of the PhD work, a workshop was organised at Computer-Supported Cooperative Work & Social Computing conference on this topic, which went on to provide much insight and guidance for the work that was undertaken subsequently. The workshop outline and focus is presented in the publication Porcheron et al. (2017a).

details about technology use in homes: that the home is a site for *multi-activity* and that the use of technology fits into this as members engage in another activity while using a device and, in the case of television watching, as part of the same activity. However, as will be discussed later in the chapter, practical and ethical issues preclude capturing video data or asking participants to assist in capturing data in this study. Other approaches exist, such as relying on automatic devices logs from the VUI device, as others have used (e.g. Ammari et al. (2019)), however this would have failed to provide access to the data this thesis needs: conversations around the VUI itself. Therefore, the study in this chapter adopts an approach used in other in-the-wild video ethnographic studies, such as by Pizza et al. (2016), who created a secondary device to be used in addition to the device being examined, which captures and records the surrounding conversation to aid analysis but that does not require activating or controlling by the user.

Previously, how the use of voice-controlled interfaces, which have become a staple feature of commercially-available smartphones, tablets and other portable devices, are managed in conversation was explicated through the study of friends socialising in a café. More recently, voice has become the primary interface with standalone screenless devices such as the Amazon Echo and Google Home. These devices are marketed as 'smartspeakers', are cylindrical in design (see Figure 6.1), and are operated using voice only. As with the VUIs found on portable devices, these devices are also referred to as 'conversational agents', intelligent or virtual personal assistants, and so on. Researchers have also adopted the term "*conversational* interfaces" (McTear et al., 2016, emphasis added), which resonates in many ways with the advertised user experience of such devices: specifically, these are technologies that it is possible to 'have a conversation' with, and which you can 'just ask' questions. In addition, the devices under study in this chapter are pitched as being especially suited for use in the home for a variety of purposes: to help with cooking, play music,

access news and information, or play games. Nevertheless, here the marketing- and perspective-agnostic term Voice User Interface (VUI) is used here in continuity with Chapter 5.

(a) Amazon Echo (b) Google Home (c) Apple Homepod

Figure 6.1: 'Smartspeaker' devices are typically self-contained cylindrical speakers designed for the home. The images of the Amazon Echo, Google Home, and Apple Homepod are copyright Amazon.com, Inc.; Google Inc.; and Apple Inc. respectively. Images of models available in 2017.

Despite the wealth of enabling research in computational linguistics such as natural language processing, dialogue systems, and computational sociolinguistics (D. Nguyen et al., 2016), research that empirically examines the social and interactional issues of VUI use in an everyday home setting is lacking. In other words, with a few exceptions, little is known about the practical accomplishment of interactions that solely take place with VUIs, or the articulation of just how those interactions unfold in the everyday lives of their users. The prior work in this thesis (see Chapter 5) unpacked this interaction in relation to use that takes place when the device interaction involved a portable personal device with a touchscreen, however, here this chapter instead explores how members of a home make use of a device when the device use is done entirely using voice. This absence of literature is significant since the work outlined here sug-

gests a range of conceptual shifts that might need to be taken into account when designing VUIs for home settings and, more broadly, multi-party interactions. This work goes on to reveal further details of collaborative efforts by members to get the device to work in and through everyday talk, in line with the research questions posed in this thesis.

6.2 STUDY DESIGN

A brief description of the setting in which the studies were undertaken is provided below, including details about the participants involved. This study deviates from the modus operandi established in the prior two empirical chapters in that the collection of data in this study is an audio-only longitudinal study in a non-public setting. This section will introduce the rationale for this design as well as the reasoning for the continued adoption of an ethnomethodological analytic orientation (see Chapter 3). The study was approved by the university's School of Computer Science Research Ethics Committee.

6.2.1 *The home a study setting*

Whereas the studies in Chapters 4 and 5 focus on studying interaction in a semi-public 'casual-social' setting (a pub and a café respectively), this exploration of standalone VUIs is conducted in participants' homes. This is driven by the design of the device under study—whereas smartphones are engineered to be portable devices that allow users to achieve tasks irrespective of their location, 'smartspeakers' are typically not designed to portable, need to be plugged in to mains electricity[3], require configuration using an app to connect to a Wi-Fi connection, and are typically pitched by manufacturers

3 Some portable standalone smartspeakers have since been released, such as the Amazon Tap, although none were available at the time of study and the Amazon Tap has since been seemingly discontinued.

as devices for the home. With smartphones, their portability affords their use and thus the possibility of *studying interactions around their use* in a range of different venues, with the ones chosen done so to explore how devices are used and interleaved within a multi-party conversation of friends socialising in a relaxed manner in a setting perspicuous to their use. However, VUIs are not marketed as portable or *personal*, but instead as communal devices that can be 'installed' in a fixed location in the home, and usable by any person living in—or passing through—the space.

Despite the differences in design, including the lack of Graphical User Interface (GUI), the concept of the VUI on a standalone device intersects with that of the VUI on the smartphone: users use a 'hot-word', they deliver their request 'using natural language'[4], and the VUI responds through a synthesised voice that sounds humanlike[5] (note that in the case of smartspeakers, this synthesised voice is always produced whereas it is optional on portable devices). Therefore, the lack of portability, personalness to an individual, and no GUI establish the smartspeaker as a distinct but interrelated product to VUIs on smartphones, and establishes the need for studying the use of such a device in a home environment.

6.2.2 *Collecting data in the home*

In order to capture naturalistic use of a VUI in the home, five households were recruited to take part in a month-long longitudinal study. The desire to adopt month-long longitudinal study approach instead of observing a gathering of friends as in the prior two studies stems from the both the design of VUI smartspeakers being pitched as home-based devices and the recentness of the introduction to market of standalone smartspeakers. While smartphones had been around for

4 This phrase is in quotes because its veracity is not established but proclaimed in the marketing materials of device manufacturers.

5 Relatively speaking, to an untrained ear, and within limits.

more than six years at the time of study in Chapter 4 and VUIs had been found on smartphones for more than three years at the time of study in Chapter 5, smartspeakers were less than two years old, and in the UK less than one year old at the time of this study. A longitudinal study would potentially allow users to configure and develop *some* competency in using the VUI device rather than only focusing on their first encounters with a device[6].

Of the five households recruited for the study, three were inhabited by couples, while two households were families consisting of two parents and two children. The age range of the adult participants spanned from late-20s to mid-50s. Each participating household was given an Amazon Echo, configured with a household member's personal Amazon account, and the Alexa companion app was installed on one of their personal smartphones. The purpose of this is to identify how the conversations unfold around the use of the VUI in the designers' intended setting. Households freely selected the positioning of their Echo and could relocate it when and as desired. Four of the households placed the VUI in a kitchen or dining area, while one placed it in a living room. These sorts of places, through the presence of the VUI smartspeaker, are made into *activity centres* and thus *ecological habitats*. Crabtree et al. (2003) define the former as "places where media are actively produced and consumed and where information is transformed" (ibid., p. 215) and the latter as "places where communication media live and where residents go in order to locate particular resources" (ibid., p. 215). The multiplicity of functions proffered by VUIs facilitates household members to make these devices 'at home' (i.e. for their use to be part of the routine lived experience of those in

6 The only real requirement was that users had developed enough competency to be able to use the device, no measure of competency was taken or sought as the goal was not to determine or evaluate use over time.

the home rather than of some unusual occurrence)—to produce and consume media, as well as a site for communication[7].

6.2.2.1 *The Conditional Voice Recorder*

Given the sensitivity and technical/ethical challenges of collecting data in the home over an extended period, this study further deviates from the previous studies by opting for audio collection only instead of video. It was considered that recruiting participants to allow video capture in their home for an extended period would generate vast amounts of video data, including of many moments participants did not want to be recorded, as well as potentially being an invasive way of collecting data. Solutions to this were considered, such as requiring participants to start and end recording, however, this would potentially impede the capture of any spontaneous interactions with the VUI device. Therefore, given the desire to study interaction in the home over a month-long period, this study further specifies that only interactions temporally close to the use of the VUI should be studied, given the sensitivity to—and infeasibility of—analysing all interaction in the home.

To capture the use of the Alexa VUI in the home, a second purpose-built device—termed the Conditional Voice Recorder (CVR)—was designed, built, and deployed with the Amazon Echo. The CVR is activated when a proximate Echo is used (the design of the CVR is depicted in Figure 6.2) by continually capturing audio using a conference microphone, i.e. it records voice *conditionally* upon some event occurring. The device always keeps the last minute of audio in a temporary buffer in the non-permanent memory of the device while active. When the hotword, 'Alexa' (that activates and begins the use of the VUI) is detected, the CVR saves the prior minute into permanent storage and records one further minute of audio (this period is extended if the

7 Although the Amazon Echo does include features for communication, such as voice calling, none was observed in these studies, perhaps given the relative nascency of the devices in the UK.

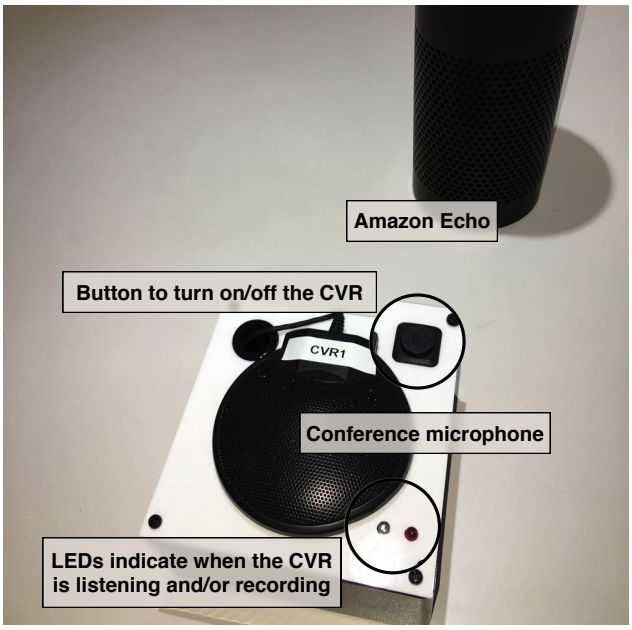


Figure 6.2: The Conditional Voice Recorder (CVR) is a small (12cm by 12cm) black box with a white lid. On the lid is a conference microphone, two LEDs (one blue, one red; to indicate when the CVR is actively 'listening' for the hotword and when it is recording, respectively), and a button to turn on, or off, the CVR. Pictured also is the bottom part of an Amazon Echo. Version 2 of the CVR is shown (Version 1 being an initial developmental prototype).

hotword is heard in the subsequent minute). The CVR also features a button to turn off audio capture, and two LEDs (blue and red), that indicate when the CVR is 'listening' for the hotword (blue) and when it is recording (red). The CVR is more sensitive than the VUI[8], to ensure talk to Alexa was captured.

6.2.3 *Analysing the collected data*

The resulting corpus of data from the deployments consists of over 6 hours of recorded data, where each audio clip consists of one or

8 Voice recognition systems often have an accuracy or threshold variable that allows a 'best guess' of whether a particular sound fits a given word or not. In this case, the CVR was prone to false positives such that it may sometimes activate recording when the user had not uttered the hotword. This was seen as favourable to false negatives, where the CVR would not activate when the VUI was used.

more requests to the VUI in the home, as well as the preceding and succeeding conversation around the request. Overall, 883 distinct 'request' utterances have been identified, where a request is talk that is directed to the VUI in a seeming attempt to get it to 'do something', e.g. answer a trivia question, play particular music, or set a timer. Often these requests formed part of a larger sequence which might encompass various other requests that are temporally and/or topically related. The corpus contains 185 of these (i.e. there are 185 audio clips containing the use of the VUI).

As with the previous chapters, this work takes on the perspective of ethnomethodology and is interested in how members organised their actions with and around the VUI. Specifically, this work examines how members, as conversationalists, analysed moment-by-moment unfolding interactions with and around the device and with one another to accomplish and address the problem that occasioned the use of the VUI.

A substantive review of the audio clips was performed to document the requests to the VUI and who made them. Twelve fragments were then transcribed carefully using the Jeffersonian transcription system (see Appendix B). These fragments were selected in line with the aims of this research to reveal the ways in which members use a standalone VUI device in and through conversation (in other words: each clip focused on the use of the VUI while other people where audibly co-present). Although there was interest in situations where the VUI was used solitarily, this is beyond the focus of this thesis' aim to study the use of technology during gatherings of multiple people and of the naturally accountable interactions around this use (see 3.2.2 on natural accountability). It was determined that each fragment selected warranted further investigation, with each fragment being reviewed individually and discussed collaboratively with other researchers multiple times. The fragments selected for inclusion in this thesis 'vividly exhibit' (Bannon et al., 1993) the interactional accomplishments of the members in selecting to use the VUI, as exem-

plars of the data in the corpus (Crabtree et al., 2012), but are not representative of how all instances of VUI use may unfold.

All names and identifiable information within the transcripts are, as before, entirely fictional. Talk that appears primarily to be addressed to the VUI by members of the setting is highlighted in bold text. In turn, the synthesised voice produced by the VUI is identified by the label 'ALE' (i.e. ALExa) in transcripts. The inclusion of synthesised voice output as part of the transcript should not be seen to suggest any conceptual equivalence among members and the VUI, but merely constitutes a convenient way of presenting the temporal organisation of device output as it appears in interaction.

6.3 FINDINGS

Data from the fieldwork will now be introduced and presented over a series of data excerpts. The fragments revealed in this chapter are taken from two different households over three different occasions. The first fragment examines how users test the *functionality* of a VUI, as per 5.3.3 with smartphone-based VUIs (see 6.3.1), and provides a clear sense of the nature of interaction with a VUI device. The second further introduces how the use of a screenless-VUI device is done accountably to the normative moral order of the setting, in this case to ask for background music suitable to a New Year's Eve party. Whereas the first fragment reveals little of the social interaction around the VUI, this second fragment progressively emphasises the intricate nature in which VUIs are used, with the different sorts of activities that take place around their use. The third fragment introduced in this chapter reveals how the design of a VUI smartspeaker supports their use for long-running activities while those collocated are also performing another task together (in this case, eating a family meal while playing a game together), bringing in the richness of the setting and how VUI devices are used as part of the multi-activity home. Interaction with the device will be shown not to take place as

a singular indiscriminate event but rather is achieved as a situated action as part of—or rather, *interleaved* within—the already ongoing activities that unfold within the home. In this sense, and by design, the home is regarded as—and will be shown to be—a perspicuous setting (Garfinkel, 2002a, p. 181) for the use of VUI devices and thus the examination of their use.

As before in the prior empirical chapters, this work is not interested in claiming that these are 'representative' or 'generalised' findings of how all interactions can or may unfold; rather, it is given that members continually try to make their own interactions orderly and rely upon the orderly features of others so as to each analyse what the others are doing and thus 'go on' (the *ethnomethodological perspective*, see Chapter 3). This means that this chapter, as with the prior two, seeks to exhibit how members bring the VUI device practically into that interaction order.

6.3.1 *Establishing the capability of a VUI*

This first fragment, called *Where is Greece?*[9], commences in Data Excerpt 6.1. This excerpt unfolds as the two homeowners, Nikos and Isabel, are entertaining their neighbours, Leah and John, who are chatting and socialising around the bar in the kitchen area of their flat. Nikos introduces the Amazon Echo and then Leah and John take turns using the device, asking it various questions. All participants in the conversation are Greek, and at times converse in Greek and sometimes in English (primarily towards the Amazon Echo). Talk that was Greek in these conversations was translated into English by Nikos following the study.

Previously, in 5.3.3, establishing the functionality of a smartphone-based VUI was shown to consist of two actions by members, and indeed the same actions are revealed to take place with the screenless VUI device when unpacking the actions of members:

9 The complete fragment is included in Appendix G.1.

(i) Establishing the desired function, and

(ii) Testing the functionality by addressing the VUI.

These activities are now explicated in the following two sections, revealing the problem case of whether the device is able to deal with requests for the details of places.

6.3.1.1 *Establishing the desired function*

We join the action in Data Excerpt 6.1 as the friends are sitting down to drink coffee in the kitchen area of the flat. John, one of the guests, has asked a number of questions to the device regarding definitions of words, a well-advertised feature of the Amazon Echo. After requesting a few words, Leah takes a turn in conversation and introduces a request for a different type of request: one for details of a specific location.

```
01  LEA   alexa (.) where is greece
02      (2.0)
03  ALE   // greece is a un-recognised country in the northern hemisphere
04        (.) it shares a border with turkey, albania, bulgaria
05        and macedonia= //
06  ISA               =[ that's it ]
07  LEA               =[ that's it ]
```

Data Excerpt 6.1: Where is Greece? (i)

In this opening excerpt, Leah makes the request to the VUI for "`where is greece?`" (line 01). Given the VUI is screenless, this request is positioned not as a 'show me on a map', but rather a request for a description of the location of the country Greece. The members do not talk as the VUI delivers its response, listing the neighbouring countries of Greece and upon the device remarking that Greece neighbours Macedonia, both Isabel and Leah simultaneously remark "`that's it`" (lines 06–07). All four friends in the room, of course, *know* the location of Greece—as they are all Greek—the request is merely to explore the capability of the device to support sense-making in how the device re-

sponds to various requests. The remark of "`that's it`", delivered jointly by Leah and Isabel latches on the completion of the name "`macedonia`" (line 05) by VUI, and turns upon a geopolitical dispute between the nations of Greece and the Republic of Macedonia, and brings to the fore how members of the setting are waiting for the name to be produced as part of the list of neighbouring countries[10].

In performing this request, Leah establishes the premise that her request is to test the capability of the VUI to respond to questions about the location of places. In choosing to request details of a location of which she is acutely aware, she reveals that her request is both to determine whether the device supports *this kind of request*, and further enable her to establish the veracity of the response from the VUI.

Of course, Leah does more than establishing the desired function of the VUI, she *tests* its functionality too, by addressing the device with her request. In other words, Leah's address to the VUI performs both the work of *establishing* and *testing* it. The next section expands upon the notion of testing a VUI by examining two further questions put to the device.

6.3.1.2 *Testing the functionality by addressing the VUI*

Following Leah's opening request to establish and question the device to give information regarding the location of places, John proceeds to ask a further question in the same vein. In this next exhibit, found in Data Excerpt 6.2, John asks the VUI for the location of *Amfissa*, a city in Greece.

```
09   JOH    alexa where is amfissa
10        (2.0)
11   ALE    // amfissa is a city in phocidos (..) greece (.) it is 82 miles
12          133 kilometres west of athens and 26 miles 42 kilometres south
13          of lamia //
```

Data Excerpt 6.2: Where is Greece? (ii)

10 This was verified through discussion with the participant household members following the study.

In this second excerpt, John follows the same structure in his request to the device as Leah (i.e. *where is...*), but this time asks for the location of a city in Greece, and could be considered to be 'upping the ante' by asking a question that is ostensibly more difficult given its specificity of being a city rather than country. In this instance, as with the prior request, the device responds seemingly correctly with the members in the setting allowing the device to complete its response.

Following this response, John makes a further attempt at *testing the functionality of the VUI*, as depicted in Data Excerpt 6.3. As before, he does this by following the same format of request as established by Leah (i.e. *where is*), but this time opts to ask the VUI device for the location of "delphi" (line 15)[11].

```
15  JOH   alexa where is (0.3) delph-ee
16      (7.0)
17  ALE   // delph-i is a village in carroll county indiana (.) indiana (.)
18        it is 62 miles 99 kilometres north of Indianapolis and 87 miles
19        140 kilometres= //
20  NIK                 =alexa stop
```

Data Excerpt 6.3: Where is Greece? (iii)

In the first excerpt in this fragment, Leah asks the VUI device for the location of *Greece*, the country she and other co-interlocutors are from. In this, she establishes her desire to determine how the device provides information about the location of places and chooses a place of which she is aware—in this case, her country of origin. This notion is further realised through the successive utterances produced in the setting (the retrospective-prospective character of action, see 3.2.2) The second request to the VUI smartspeaker, by John, her partner, further tests the VUI by increasing the specificity of the locale put to the device to locate. In both cases, ostensibly, the device responds as members seemingly expect. In this final excerpt, John asks for the location of "delphi" (line 15), which is an ancient monument

11 This is spelt *delph-ee* in the fragment transcript to distinguish it from the different pronunciation used by the VUI device in its response of *delph-i*.

and a registered UNESCO World Heritage Site in Greece. With this, it seemingly becomes evident that John's follow up question is asking for the location of something more precise than the prior two questions (country and then city), and further establishes the ratcheting of specificity over the three requests to the device in this fragment.

After John's request, the VUI device takes 7 seconds to respond, during which the members remain quiet[12]. The device begins responding, and commences its response with a different pronunciation of the place requested by John (delph*ee* vs. delph*i*), and gives the location of a village in Indiana, USA. Given the coherence and context established through the production of the prior two requests to the device, it is determined that this response from the device is not the intended location of which John was seeking details. Nikos cuts the VUI device off as it continues to produce information about Delphi, Indiana by producing the request "`alexa stop`" (line 20).

The group discussion then moves on to new topics after this response, affirming this sequence of events was around the purpose of *testing the functionality of the VUI* rather than actually seeking information about the location of a country, city, and landmark. With smartphone-based VUIs, establishing the capability of the VUI was positioned as a case that deviated from actions performed typically with touchscreen-based devices, as the VUI features were made available to users in the past few years. Moreover, in the context of the screenless devices, which were released within the UK within the prior three months before the recording of this fragment occurred, establishing the capability of the VUI was a recurrent practice amongst all the households studied.

12 Given VUIs devices of this nature do not verbalise what-it-is-doing, it is unclear as to the cause of this elongated time between request and response.

6.3.1.3 *Methodical accomplishments in this fragment*

In this first fragment, the work of establishing the capability of a smartspeaker VUI to respond to "`where is`" questions is unpacked. Leah initially opens the **address to the VUI by using the hotword, followed by her request**. Through the member's perspective, this request is establishing and testing the functionality of the VUI, given that Leah knows the answer to the question. The VUI computes a response for a few seconds, and the co-present members **do not talk until the VUI has fully delivered its answer** through a synthesised voice. A second request is produced by another member through **address to the VUI**, using the same wording but with a more specific location (i.e. a 'harder' question, and again this turns upon the members of the setting being knowledgeable of the correct answer). Again, the VUI computes and responds to this request through the synthesised voice. A third request that is ostensibly more difficult again is **addressed to the VUI** by the same member. The VUI takes seven seconds to compute an answer and begins to deliver its response, using a different pronunciation of the locale given by the user. The coherence of the questions establishes that this answer from the VUI is 'incorrect' (i.e. for a different location with the same spelling), and is reinforced by another member of the setting **cutting off the VUI's response delivery by addressing it with a stop command**. Through the perspective of the members, this sequence of requests is about establishing and testing the capability of the VUI—not to literally find the location of Greece, Amfissa, and Delphi. The interactional project was a success from this perspective.

6.3.2 *Asking the VUI to play music*

This second fragment, called *New Year's Music?*[13], commences in Data Excerpt 6.4. This excerpt unfolds as the same two homeowners as

13 The complete fragment is included in Appendix G.2.

above, Nikos and Isabel, are hosting a New Year's Eve party. The party has been going for some time and during a conversation about the Amazon Echo and the CVR, an attendee commences the interactional project of playing some background music using the Amazon Echo. One of the key marketed features of the Echo is the ability to perform long-running tasks such as playing music or setting timers.

In this fragment, a guest, Anna, will be facilitated in making a request to the device by Nikos, however, the request ultimately fails and the two users attend to dealing with the outcome of the device's computation. This fragment will progressively demonstrate the ways in which interaction with a VUI occurs within and is accountable to others within the setting.

Overall, the members' problem of getting the device to play music in the background will be shown to consist of two core activities that will be unpacked across a number of excerpts:

(i) Requesting the VUI to play a category of music, and

(ii) Responding to the VUI's choice of music.

6.3.2.1 *Requesting the VUI to play a category of music*

Firstly, the request to the VUI device to play music is examined. This request is presented in Data Excerpt 6.4 and occurs amongst the hubbub of the party in the background. Nikos, the homeowner, opens the interaction with the Echo by producing the hotword, before Anna performs a request for the device to play music.

```
01  NIK   alexa
02      (2.6)
03  ANN   play some new year's music
04      (1.8)
05  ALE   // here's a station for jazz music (.) instrumental jazz //
06        ((begins playing jazz music))
```

Data Excerpt 6.4: New Year's Music (i)

The first consideration in examining this fragment is the categorisation of music that Anna uses—specifically that of "`new year's music`" (line 03). This categorisation does not a carry specific genre or type of music connotation, yet of course, remains a normatively understood request to be for music suitable for a New Year's Eve party. This illuminates a key consideration of how users approach such devices: given there is no reference for correct use of the device[14] or input to the device (i.e. there is no *a priori* information as to what works, or does not work), users must produce utterances that may or may not work to determine the capability[15].

Specifically in this case, "`new year's music`" turns upon various socially shared and culturally situated assumptions about what constitutes relevant music to play for New Year's Eve. As members of society, we routinely deal with and attend to such complexities of categorisation[16], yet such challenges are not pre-determinedly defined, such as a codified genre or specific artist or song, but rather music relevant to a season or holiday. The device responds to this request by playing *instrumental Jazz music*, although in the opaqueness of the device, it remains unclear whether the device has understood 'correctly'. A later examination by the researcher of the web-based logs available to Amazon Echo users revealed a request for "play jazz music", suggesting[17] that the device did not correctly transcribe the spoken input.

6.3.2.2 *Responding to the VUI's choice of music*

The next element of this fragment is to consider how Anna responds to the device's next action to play "`instrumental jazz`" (line 05) music. The next excerpt, Data Excerpt 6.5, commences as she does so.

---

14 By *correct use* this thesis means *gets the device to do the desired function*, i.e. it is the correct outcome from the user's perspective.

15 In contrast to a GUI which could display possible next options.

16 Nikos does not challenge or guide Anna to expand upon her choice of music.

17 Although not *confirming*, as the veracity of such cannot be established.

12	ANN	**alexa this is not what we wanted**
13		[((laughs))]

Data Excerpt 6.5: New Year's Music (ii)

This second excerpt brings to the fore a key element of interaction with VUIs through deepening Anna's response to the device's next action. First of all, it becomes evident that to Anna "`new years music`" does not include the "`instrumental jazz`" station the Echo has opted to play. She attends to this matter by using the hotword to activate the device and says "`this is not what we wanted`" (line 12). In furthering the previously established point regarding the use of voice-based interfaces being used with various socially agreed categorisations, it may be suggested that the current design of VUIs necessitate a try-and-see approach. Consider this instance in the context of both the prior fragments in this chapter (see 6.3.1) and the use of the smartphone-based VUI to explore device capability (see 5.3.3): in all three instances, the practice adopted by users is that of making an attempt at doing something to *understand if an option is possible*. This underscores an intrinsic difference in the way devices get to be used for various tasks: typically GUIs present available options to users through menus, graphics, icons, and so forth; however a VUI provides little in the way of *affordance* to guide the user[18] resulting in a *trial-and-error* approach: users must issue requests to determine *if and how the device would respond* in order to determine *what the correct input is for the device to respond the way they want, if it exists*.

A third excerpt, Data Excerpt 6.6, is now introduced that incorporates the prior excerpt (for readability), but also includes the interaction that follows Anna's remark that the music is not what she desired, followed by laughter.

18 Here the use of *affordance* is more in accord with Norman (1988)'s use of the term, i.e. an affordance provides "strong clues for the operation" of the item (ibid., p. 9), rather than Gibson (1979)'s original definition.

12	ANN	**alexa this is not what we wanted**
13		[((laughs))]
14	NIK	[(1.2) **alexa** (1.1) **shut**] **up!**
15	ANN	hey::↑(.) **alexa nikos apologises for being so rude**
16		(0.3)
17	ALE	hi there
18		[*((resumes playing jazz music))*]
19	NIK	[(2.4) **alexa stop**] **stop!**

Data Excerpt 6.6: New Year's Music (iii)

In this final excerpt from the fragment, the device does not respond to Anna's request, following which Nikos then instructs the VUI to "shut up" (line 14), a command that stops the playback of music on the device[19]. Given the accountable nature of interaction with a VUI device—in that talk to and from the device is audible and reportable by those present, talk to the device exists and can be called to account within the normative moral order of the setting. In the excerpt, Nikos tells the device to "shut up" (line 14), to which Anna produces an ostensibly ironic apology to the device for the rudeness of his response "nikos apologises for being so rude" (line 15), and in doing so establishes the viewpoint that Nikos' request to the device breached the normative moral order—i.e. the socially shared and agreed-upon sets of ways of acting—against which members of the setting are held to account. In this indirect rebuke, Anna enforces the notion that the use of the VUI occurs within this normative moral order, and in turn, that the device itself becomes embedded within the fabric of the home, through the established and expected moral organisation of social conduct. In other words, the VUI smartspeaker does not exist as a personal or private device, but one for which its use is considered an activity that is accountable to all in the home.

19 The VUI responds to this command in much the same way as the command *stop*. The two most common VUI devices, Amazon Echo and Google Home, now include a child-friendly mode that primarily only responds to 'polite' requests from users, although this did not exist at the time of the study.

6.3.2.3 *Methodical accomplishments in this fragment*

This second fragment examines the interactional project of requesting a VUI to play some music. During a party, a discussion ensues about playing some background music. The homeowner, Nikos, **addresses the VUI by uttering the hotword** to activate it, following which a guest, Anna, **addresses the VUI with a request for appropriate music**. The VUI computes and then responds to the request with a synthesised voice describing its next action—to play instrumental jazz music. Anna again **addresses the VUI**, by producing the hotword and **stating that the outcome was not as they desired**. The VUI does not respond within a second or so to this address, following which Nikos then **addresses the VUI by instructing it to "`shut up`"**. This address leads to Anna **reprimanding Nikos** by rhetorically **addressing the VUI again that Nikos "`apologies for being so rude`"**. The VUI produces a response that does not seem to coherently follow from any of the three prior requests addressed to it. After the VUI resumes playing music, Nikos **cuts off the VUI's playback** through a further **address of instructing it to stop**, thus ending the interactional project to play music using the VUI.

6.3.3 *Using the VUI to play a game while eating*

To briefly recap, the first fragment introduced how VUI devices are used within conversations in homes to respond to various requests for information, as users test and explore the functionality of the device through use. The second fragment deepened this explication of how VUIs are used within the home, revealing elements such as how users may select socially-established definitions in their requests rather than typical *a priori* discrete categorisations; and further how VUI use is held to account within the setting.

The final fragment, called *Beat the Intro*[20], is from a family consisting of two parents, Susan and Carl, and two children around ten years old, Liam and Emma; and further demonstrates how VUI devices are brought into an ongoing activity in the home, but more so function as multi-user devices within the multi-activity home. In this sense, the VUI device is shown to be used in and around conversations in the home. In this home the VUI is placed on the top of a bookcase that is used as a sideboard in the dining room. The family have been using the Amazon Echo for approximately a week, and have developed a reasonable familiarity and competence in its use, with each member of the household having used the Echo for most days at least once or twice. They are eating an evening meal all together at the dinner table on Mothers' Day.

This problem of getting the device to play a game while the family are eating a meal is unpacked over the following three activities, with the former establishing how the VUI becomes introduced into an ongoing activity in the home:

(i) Preparing to address the VUI,

(ii) Requesting the game, and

(iii) Responding to the VUI's action.

6.3.3.1 *Preparing to address the VUI*

As we join the family in the first excerpt, Data Excerpt 6.7, Susan, the mother, announces to the others that she would like to play Beat the Intro "`in a minute`" (line 01).

Beat the Intro is a game available for the Amazon Echo that the family have previously played together; it involves listening to a few seconds from the start of a song and then players must guess, by announcing, the song and the artist. The game is a 'Skill'—an installable feature developed by a 3rd-party for the Amazon Echo.

20 The complete fragment is included in Appendix G.3.

```
01   SUS    i'd like to play beat the intro in a minute
02   LIA    [ oh no:: ]
03   SUS    [ alexa   ][ (1.1)  ] beat the in[tro
04   CAR               [ °yeah° ]
05   LIA                                    [°no:::...°
06      (0.6)
07   CAR    it's mother's day?
08      (0.4)
09   SUS    it's (    ) yep (.) listen (.) you need to keep on eating your
10          orange stuff (.) liam
11      (0.7)
12   CAR    and your green stuff
```

Data Excerpt 6.7: Beat the Intro (i)

Susan announces to the family that she would like to "play beat the intro" (line 01), and in doing so, prepares the family to play a game together using the VUI. Liam produces an assessment of this ("oh no", line 02) and then an elongated "no" (line 05) as Susan then instructs the VUI to play the game. Carl mentions Mother's Day, while Susan instructs Liam to eat his food.

The first observation is that addressing the VUI—here located in instructions to play the Beat the Intro skill—is interleaved amongst *multiple activities*, or 'courses of action', that the family are working to accomplish together. For instance, the family are eating dinner together, and they are talking about that eating (lines 09–12 particularly). Requests for compliance from Liam are produced by Carl amongst Susan's initial instruction to the VUI (line 03), where Carl counters Liam's negative response to Susan's preparatory utterance "i'd like to play beat the intro in a minute" (line 01) with the reminder that "it's mother's day?" (line 07). Activities that might be glossed broadly as 'parenting' turn on establishing appropriate ways of behaving during mealtimes particularly for younger members of the family, such as the instruction to Liam to "keep on eating your orange stuff" (lines 09–10). All the while, these other concurrent activities are closely geared into the organisation of Susan's further requests to the VUI.

It is through these 'other' utterances—not to the device, but to each other—around which the VUI is used, that establishes the family's treatment of playing a game with the VUI during dinner with perspicuity. In this sense, the activity is not oriented to as unusual or out-of-place, merely *unwanted* by some members because of their inevitable involvement. Furthermore, here, the action of preparing others for the use the device as a family for a cooperative activity demarcates this *type* of device as different to smartphones, insomuch that here the smartspeaker is to be used *together* by the family, and that this preparatory account is used to ready the members for the next action Susan is to perform (i.e. that she, and the family, are to play Beat the Intro together).

Another issue to consider is how the VUI device responds to the user, and whether this is treated as a *success* in the course of action by the users. In the first fragment examined in this chapter, the device provided information on the incorrect *Delphi* (USA, as opposed to Greece) and in the second fragment, the device provided the wrong sort of music as an issue in transcribing spoken words into text; in this excerpt, however, a different type of technical problem occurs in comparison to the prior two fragments: no-response. In this sense, it remains unclear as to the specific nature of the problem at hand and provides no information as to the actions the user (or as shall be revealed, *users*), should take.

In many ways, these initial observations offer a consonance with prior studies of technology use in the home and how such technologies get drawn into the organisation of home life as resources for action (e.g. see Rooksby et al. (2015)). Empirical accounts such as these present a more nuanced perspective to the conceptualisation of such technologies like the VUI as disruptive to established moral order by drawing attention away from interaction with co-present others (Turkle, 2011)—rather, here it can be seen that what is unfolding is the use of the VUI alongside other ongoing activities, suggesting that VUI devices get recruited into the lifeworld of cooperative and colloc-

ated activities in the home (Rigby et al., 2017; Rooksby et al., 2015; Tolmie et al., 2008). In this sense, the use of these devices becomes regulated *in* those activities.

6.3.3.2 *Requesting the game*

The next matter to turn to is the address to the VUI, and for this a further excerpt of data is presented in Data Excerpt 6.8. A request has already been made on line 03 above, although this has ostensibly 'not worked' by virtue of the device not responding to the request. In this next excerpt, the matter of how members in the setting embed this request amongst the ongoing activity in the home is examined.

```
13  SUS   alexa (1.3) alexa (0.5)=
14  CAR                          =°and your brown stuff°
15  SUS   play beat the intro
16  EMM   °and the yellow stuff?°
17  LIA   °and the meat stuff°
18      (0.9)
19  ALE   // resuming the music //
20  EMM   ((laughs))
21  ALE   ((music plays))
22  SUS   oh no::!
23  EMM   ((laughs))
24  CAR   alexa stop:
        (...)
32  EMM   alexsa [ (1.0)                        ] play beat the intro::
33  CAR          [ is it called beat the intro? ]
```

Data Excerpt 6.8: Beat the Intro (ii)

This excerpt commences with Susan's repeated request given the VUI devices non-response. No response from a device also occurred with smartphone-based VUIs in the café where members would repeat a request if the device seemingly did not 'hear' a request made to it (see 5.3.1.2). Susan's request (lines 13 and 15) is again interleaved with the ongoing parenting activity by Carl (line 14), and on this occasion the device responds to her request (line 19) by "`resuming the music`" (line 19). In both of Susan's requests to the VUI (lines 03 and 13—

15), Carl talks during the request, on the first occasion in agreement with the game and in this latter case to instruct Liam to eat his food. Following Susan's second request, the device performs an undesired action: it resumes music that was played previously. The music is stopped through a request to the device, and Carl then attempts to start the skill again (omitted in this chapter, although can be found in Appendix G.3). We then hear Emma take an attempt to start the skill (line 32). As she does this, Carl inserts a question between Emma's utterance of the hotword and the main request, suggesting that perhaps the family are using an incorrect name of the Skill.

For users of VUI the data show the ways of addressing the device provide for certain conversational structures that members can orient to in interaction as a request is made. Consider for example Carl's questioning of the name of the Skill, "`is it called beat the intro?`" (line 32), and just how he inserts it sequentially into Emma's utterance (line 31). Carl produces this question precisely in the 1.3 second gap between Emma's production of the hotword "`alexsa`" (line 06) and subsequent request to the device "`play beat the intro`". Consider also the request performed by Susan on line 03 of the prior excerpt (Data Excerpt 6.7), where she utters "`alexa (1.1) play beat the intro`" (line 03) while Carl quietly says "`yeah`" (line 04) during the 1.1 second pause. Carl's "`yeah`" provides a counter to Liam's rejection of Susan's preparatory utterance in line 01, and, importantly, this "`yeah`" is positioned at the precise moment after Susan's production of "`alexa`"—Carl appears to be orienting to this regular pause. The syntactically formulaic nature of input production to a VUI device, i.e. that of *hotword-gap-request*; enables competent device users to project this gap, to constructively minimise silence, and to therefore offer the possibility of taking advantage of the gap to take a turn-at-talk. Often this also leads to the original requester interacting with the VUI then selecting to resume talk following this interweaved utterance (DeVault et al., 2014, pp. 302–304), re-emphasising the nature of VUI devices being used as an activity alongside other activities.

6.3.3.3 *Responding to the VUI's action*

Before examining how members sought to remedy these problems, it is necessary to look at a related issue: how responses themselves are treated by those using the VUI to accomplish a task, and in this case, as suggestive of trouble. Whereas in VUIs on portable devices, as unpacked in Chapter 5, voice-to-text transcription is often displayed on-screen, users of screenless devices have to rely solely on the audible response (although they may find more clues as to what went wrong in the companion app supplied with most screenless devices). The analysis of interaction with the VUI reveals a significant mismatch sometimes between the ways in which designed responses from the VUI appear to integrate indicators of the form of trouble, and how members dealt with them. Although it is tempting for simplicity's sake to call certain responses from the VUI 'error messages', this would not be correct, as these responses are not always the result of a computational error, e.g. they may be due to the VUI device mistranscribing the request. Nevertheless, these responses are a resource for diagnosing and resolving the trouble. This point forms the central concern with the final excerpt, to be presented in Data Excerpt 6.9.

```
35  ALE   // you want to hear a station for b b intro [ (0.5) ] right? //
36  EMM                                                [ °no:° ]
37      (1.1)
38  EMM   no: (.) i don't alex:a (0.5) no!
39      (1.3)
40  ALE   // alrig↑ht //
41      (0.7)
42  CAR   we played it the other ni:ght! the game we played
43        the [ other night ((laughs)) ]
44  SUS       [ yeaherr:: alexa        ] skills (.) beat the intro
45      (4.5)
46  SUS   °uh::↓:°
47  EMM   she didn like tha:↓:t
```

Data Excerpt 6.9: Beat the Intro (iii)

In this final excerpt, the VUI responds to Emma's request to "play beat the intro" (line 32), by questioning whether the user wanted to "hear a station for b b intro" (line 35). In this response, the device ostensibly incorporates a partial transcription of the request into a question to the user, implying there is some uncertainty in the device's processing as to the user's request. Seemingly, the device transcribes "beat the intro" as "b b intro"[21], with the word "play" taken by the device to be a request for music playback. Given the device's response to the request made by Emma, which consists of an uncertain next candidate action the device could take, the device is 'actively listening'[22]. Emma retorts "no, i don't alexa, no" (line 38), to which the device ends the interaction and returns to its initial state of listening for the hotword. Given the device's 'failure' to respond to the requested action, Carl makes a response that suggests exasperation with the device ("we played it the other night!", line 42), before Susan then takes another attempt, using a different verb: "alexa skills beat the intro" (line 44)[23]

In this fragment, the family take it in turns to repeatedly rephrase the request as slight variations: first without a verb at the beginning of the request (line 03), before incorporating the verb "play" (lines 15, 32) into the request, and later swapping "play" for the noun "skills" (line 44). These requests are also varied through differences in the prosody (cf. lines 15 and 32 in the prior excerpt). Both of these differences in request production make available to the observer the collective demonstrable reasoning of the cause of the trouble by the user, i.e. that it is the words in the request, or the utterance of those words,

21 This transcription was confirmed from *a posteriori* examination of the logs of the Amazon Echo, and is assumed to be in relation to the opening sequence music for the television show *Big Brother*.

22 In other words, its microphone remains active thus the next utterance by the user does not require the hotword to activate the device.

23 This request also fails, and it takes a further minute or so before the family are able to start the Skill using the verb *start*. While each successive request could be iterated through in this chapter, it becomes superfluous given each successive attempt follows the same actions of members in revising of their request.

at fault for the failure of the device to respond as desired, and that a different request is needed. Indeed, Carl's questioning of the Skill name demonstrably affirms this. Further, Carl's remark that the family had previously played the game (line 42) suggests that the family are treating this as a problem of getting a game they have previously played to start.

Through these requests to the device, taken by different members, without specific invitation from Susan who instigated this sequence of activity with the device, the suggestion follows that the communal nature of VUI devices lends itself to collaborative efforts in which multiple members may work together by using the device to start an activity in which all can engage. Of course, this is insomuch that anyone present can make a request to the device, and that such interactions are guided by the normative moral order of the setting in which the members and device co-exist. Furthermore, this again echoes prior work in this thesis that demonstrated that collaborative interaction with VUIs is replete with such repetitions and rephrasings (e.g. see 5.3.1.2) when responses from the VUI are made accountable. In this case, this is done in and through the interaction with the device.

Overall, this excerpt was included in this thesis as an exhibit of how the use of VUI devices are also demonstrably used alongside and during other activities in the multi-activity home, in this case, while eating dinner. A single member, Susan, announces that she wants to play a game using the VUI while eating dinner, and other members, given her prerogative to play the game as it is Mother's Day, ostensibly acquiesce to this decision. However, the members are all recruited into the attempt to start the game, given Susan's failure to get the VUI to start the game for her on the first two attempts. Members take it in turns to use the VUI, all the while having a separate parallel discussion regarding dinner. Ultimately, the family eventually succeed in their problem of getting the device to start the game, and play the game before later choosing other games to also play. What becomes clear through the explication of activity in this fragment is how the

VUI use is interleaved within the ongoing conversation and in combination with prior fragments, and how this use is regulated as part of the moral order of the home.

6.3.3.4 *Methodical accomplishments in this fragment*

This final fragment focuses on conversation to and around the smart-speaker as members of a family attempt to play a game together while eating dinner. In this fragment, Susan first **makes a preparatory account to the collocated others** of her desire to play a game together using the VUI. She then **addresses the VUI amongst the ongoing conversation** between members of her family by instructing it to "`play`" the game. The VUI does not respond, and the adults at the table guide the younger members to comply with Susan's desire given her prerogative of it being Mother's Day. The members of the setting then make a joke about Susan's use of language as she instructs Liam to eat his food. Following this, Susan makes a second **address to the VUI using the hotword (twice) and the same request** as before. On this occasion, the VUI responds by resuming the playback of music to which **Susan makes a negative remark** of "`oh no`". Carl **cuts off the VUI through addressing it with "`stop`"**. Emma **addresses the VUI using the same request but with different prosody** to Susan. The VUI responds by offering a next candidate action, although this is not the 'correct' answer and thus Emma **responds to the VUI by dismissing its offer**. Susan makes another attempt at **addressing the VUI with a different word at the start of her request** ("`skills`" vs "`play`"), although this also does not result in the desired outcome. The family take a few further turns at **addressing the VUI with different opening words** in their requests before ultimately succeeding in starting the game.

6.4 CHAPTER SUMMARY

This chapter concludes Part II and the empirical work of this thesis. The approach taken to data collection and analysis within this thesis, and outcomes of the analyses that have attempted to answer the research questions posed, produced three packages of thick description of members' activities during interaction with devices in multi-party casual-social settings. Whilst the first study in Chapter 4 sought to unpack the methodical approach to how members converse in a pub and make use of touchscreen-based interactions, it also brought to bare findings that provoked studies two and three. For example, members were shown to visually and verbally account for private device interactions in a public manner given the personal and discrete nature of interactions on a small touchscreen device, and through this accounting work, supported collaborative actions. The second study in Chapter 5 augmented these interactions with portable devices by examining how members embedded and oriented to interaction accomplished primarily using the VUI on the device. These findings revealed how members still did accounting work for interactions, but also crucially chose to involve other members in their device interactions by sharing the device screen, or through ceding control of the device itself. Both of these studies, which drew upon portable devices with touchscreens, involved collaborative efforts among members to complete the task.

The analysis presented in this chapter attempted to examine interaction around the use of a screenless VUI device, marketed as a smart-speaker. By drawing upon fragments from the corpus of recordings of Amazon Echo use collected from multiple homes, these findings reveal how the use of the VUI is interleaved with ongoing activities in the home. A number of factors of VUI use are revealed, such as the nature of VUIs requiring users to adopt a trial-and-error approach to getting the device to work, resulting in the user having to make

sense of and deal with errors before attempting a further request to the device.

The use of the device was shown to be held to the normative moral order of the setting, with specific requests held to account by co-present others. Furthermore, requests to and responses from the device were used by others to join in and use the device as part of the singular effort to complete the occasioned task with the device. In this regard, it is posited that the accountable nature of interaction with the VUI ostensibly allows for other members of the setting to involve themselves by making requests to the device, such as in cases when the prior attempt 'failed' for whatever reason. In prior observations of touchscreen-based interactions, such accounting work was done by the users of the devices through additional actions, however, in the case of the VUI device this was accomplished in and through the use of the device. Finally, the use of the VUI was shown to occur alongside and during other activities in the multi-activity home, such that use of the device did not preclude either the initial request interlocutor or others from engaging in additional conversations or activities, such as consuming food.

6.4.1 *Methodical accomplishments*

This final empirical chapter has presented three fragments of VUI interaction being accomplished in and through conversation in the home. The use of the VUI was shown to be occasioned as a conversation topic itself, to establish and test the functionality, to play background music at a party, and to play a game with collocated others. Members addressed the VUI, first with the hotword, and the using questions to request it to complete an action. Further repeated utterances of these questions were made as a result of the VUI not responding. Members' addresses to the VUI were held accountable to the normative moral order in which the utterance took place. Talk to the VUI and the responses from it were used by members to practic-

ally reason about why prior requests 'failed', and to attempt revised requests with different prosody or words in a further address to the VUI. Members would cut off the VUI if it was proceeding to take action that was not desired by users. The use of the VUI in the home would stop if the interactional project was successful or abandoned by members as a result of their ostensible inability to complete the occasioned task.

Part III

SYNOPSIS

7

DISCUSSION

This discussion chapter will first turn to answering the research questions in this thesis by revealing: (1) an understanding of what constitutes a causal-social setting, (2) the nature of conversations that unfold within such a setting and the interactional projects that involve the use of devices, and (3) how this device use was interactionally organised in and through the ongoing conversation.

Through this reflection, this discussion will synthesise and illuminate key aspects of the studies, including developing an understanding of the conduct in casual-social settings, reflecting upon the approach taken to the studies in this thesis, and raising insights for future design work. Through reflection on these three points, this discussion will establish the thesis':

1. *Substantive* contribution through the presentation and discussion of members' conduct in casual-social settings and of how device use was involved in this interaction, relating to the existing literature introduced in this thesis (see Chapter 2),

2. *Methodological* contribution in terms of practical approach (i.e. technology and setting) adopted in each of the three studies, and how the methodological validity is maintained, and

3. *Conceptual* contribution through the discussion on how collaboration unfolded, establishing the case for CSCW to study casual-social settings, and for HCI to examine design with VUIs for collaborative action further.

7.1 CONVERSATION IN CASUAL-SOCIAL SETTINGS

This section brings together the empirical work in this thesis and reflects upon the findings in the scope of its overall research questions and objectives. This thesis proposed the following research questions:

RQ_A What is accomplished in and through the use of devices in casual-social settings?

and

RQ_B How is this device use interactionally and accountably organised?

As described in the introduction, both of these questions serve as constituent parts of the overall objective that underpins the empirical work in this thesis, namely to identify:

RQ How are devices used within a multi-party conversation in a casual-social setting?

This section discusses the various key tenets of this thesis: the notion of casual-social settings, the conversations that ensued within those settings, and the purposes for which device use was occasioned in and through the conversation and how this device use was sustained as part of this conversation. This section forms a core component of this thesis' substantive contributions, by synthesising the notion of a casual-social setting and members' conduct within these settings, and how this relates to existing literature.

7.1.1 *Casual-social settings*

This thesis took a pragmatic approach to understanding members' actions in casual-social settings by observing and examining groups of people socialising and interacting together. The introduction of this thesis set out the notion of a casual-social setting (see Chapter 1), and this was progressively developed throughout the empirical chapters

(see 4.2.1, 5.2.1, and 6.2.1), with the work of members in those settings presented through the unpacking of empirical data.

The literature review in this thesis also discussed how the origins of designing for and studying collocated interaction in HCI and CSCW was within meeting rooms and control rooms. As these fields *turned to the social* (see 2.3.1), designers and researchers embarked upon studying a range of "centers of coordination" (L. Suchman, 1997) and later other settings such as public places (Weilenmann and Larsson, 2002), museums and cultural visiting locations (Ciolfi and Bannon, 2003; Fosh et al., 2013), and the home (Rooksby et al., 2015; Ferdous et al., 2016). This literature showed how the use of portable technologies has become ostensibly ubiquitous, with their use observed in all manner of settings as part of various activities. Given the plethora of places in which device use has been identified as a recurrent activity, this thesis was not concerned with identifying further such places but in identifying the members' practices.

Therefore, the setting under study, not a specific location but rather a setting defined by the activity within, consists of a place in which people can gather to socialise, relax, and otherwise engage in an informal conversation. In this sense, the notion of a casual-social setting expanded upon work by others who have explored places such as social places and third places (Oldenburg, 1989). Such spaces were defined as pubs (Fox, 2004, pp. 88–108) and cafés (Laurier, 2008a), and non-home or non-work places (Oldenburg, 1989). In this thesis, the casual-social setting is, then, an amalgam of these definitions, with the underlying requisite for the selection of a place being that it be a place where members socialise together and relax as a group.

Interaction in three settings was empirically studied, in line with the three types of technology studied—that of touchscreen interaction with a portable device, VUI interaction with a portable device, and VUI interaction with a screenless smartspeaker. The first study took place in a pub, a setting in which recent literature in HCI had proclaimed an intrusion by the mobile phone (Su and Wang, 2015), disrupting

the activities within. The second, a café, had also seen such claims, with authors such as Turkle (2011) raising concerns about lost emotion and experience of places due to the use of devices. In one case, for example, she highlights what she deemed a problematic situation in which she was with her daughter in a café in Paris, while her daughter used her phone to communicate with friends back home and thus was unable to experience Paris in the way she previously had (ibid., p. 156). The third setting, the home, had also seen criticism (for example, technology used during mealtimes being perceived as problematic by other family members (Rimer, 2009)). Each of these settings was chosen as they were perspicuous (Garfinkel, 2002a, p. 181) to the study of conversation amongst groups of people where device use was known to unfold in such a gathering. Indeed, as discussed, some literature has examined portable phone use in a variety of settings (see 2.3), yet few have attended to explicating the purposes for which device use is occasioned and how this use is practically accomplished during face-to-face encounters[1], beyond mere glossing of interaction as unfolding or extracting perceptions of it through *a posteriori* methods; in other words, empirical data was nascent.

The first two empirical chapters, then, reveal how interactions in these three settings unfolded when the use of the device was done through touch-based interaction and through touch and voice-based interaction respectively. The third empirical chapter examined interaction that included the members' use of a third technology—a VUI smartspeaker. This technology was new at the time of the study and this thesis represents the first academic effort to study interactions with such a device from an ethnographic in-the-wild approach. Smartspeakers are explicitly designed to be installed in a location, requiring mains electricity and a Wi-Fi connection for operation, and are pictured in marketing materials as being positioned in places such as

1 The work of Brown et al. (2013), Brown et al. (2014) and Pizza et al. (2016) being some of the few exceptions, as discussed in the literature review in 2.3. None of these studies, however, approach the topic of specific social gatherings such as those examined in this thesis.

a desk in an office, or in the home. Therefore, this study also took place in a perspicuous setting for this technology—a communal area in the home.

While this thesis is not the first piece of work to use the term 'casual-social setting', with it applied to places for smokers (Schane et al., 2009), hotel suites (Pigram, 1996), college classrooms (Yamada, 1981), and places where "people can openly meet and interact with one another" (St. Lawrence et al., 1983, p. 42), it is seemingly the first attempt at explicating interaction in such a setting thus defined. The empirical chapters in this thesis reveal how interaction in this sort of setting is replete with complex tasks undertaken to address members' needs, as part of the unfolding casual social interaction. Indeed, this thesis delves deeper than existing literature by its focus on minutiae of interaction amongst groups to uncover members' interactional projects for which devices are used, and thus addresses the gap in relation to empirical data of device use. The next section discusses the findings from the study of social interaction in such settings, using the data presented in this thesis.

7.1.2 *Conversations in casual-social settings*

This thesis was motivated by the plethora of literature in both the popular press and across multiple academic disciplines on social interaction and the influence of technology being present or used during face-to-face encounters. The literature review (see Chapter 2) identified a lack of empirical data on *what is actually done* when people gather to socialise, and this is the first consideration of this discussion, identifying the members' interactional projects in conversation that occasioned the device use.

For the first two studies, which took place in the semi-public settings of a pub and a café, participants were recruited as groups of friends to gather and socialise together. In the third study, families were recruited as households to take part in the study, with the smart-

speaker positioned in a communal place within the home. With regard to the practices observed in the studies themselves, however, the device interactions that took place were occasioned by the members, without guidance or prescription of what should be done as part of the research. In this sense, what was studied was unscripted and naturally occurring interaction within the setting. In all three situations, the focus of the study was on the conversation that unfolded, and of the methodical ways in which device use was occasioned, interleaved, and oriented to within the interaction. In other words, the focus was on the social interaction around the use of the devices, and not the use of the device itself.

Indeed, in each setting, conversations were shown to vacillate between topics, with new topics pivoted to in talk without resolution of prior discussions, e.g. in the resumption of conversation following device use in the pub (see 4.3.3) or without clear antecedent in talk, as Susan proposes that the family play a game while eating a meal together (see 6.3.3). Conversations were also shown to go on in parallel to other discussions at the table, as separate conversational floors, e.g. as the friends have two separate but related conversations about animals with accents (see 5.3.2), or as two conversational topics interleave with each other, as members alternate between the topics, e.g. as the family make a joke about Liam eating his food amongst attempts to get the VUI to start a game (see 6.3.3). Members were observed entering and leaving conversations, e.g. to get drinks (prior to the start of 4.3.2), food, and to visit the toilet.

Additionally, in the observed studies members frequently made their device use naturally accountable by making the specifics of their device use observable and reportable[2], e.g. by making a joke about something they read on their phone (see 4.3.2), by rotating their phone around (see 4.3.2), by making it visible to others (see 5.3.1), or by explaining their reasoning for using the device as they recall read-

2 As introduced previously in 3.2.2, the *natural accountability* of action is an accomplishment of members in the setting (Garfinkel, 1967, pp. 1–34)

ing a news story recently (e.g. as Lily responds and confirms that she had read a story about animal accents (see 5.3.2). This reinforces the perspicuity of device use in these settings: device use was not ostensibly treated as out of place in such a conversation, but rather as part of it, interleaved in—or occurrent alongside—conversation. In other words, device use was treated by the members as part of, or rather *embedded*, within the interaction in the casual-social setting. Furthermore, through unpacking the methodical accomplishments of members, this embedded*ness* is shown to be an interactional accomplishment *as a result of* the practical work by members.

The use of devices was observed to be oriented as a matter to resolve the members' problems for which the device use was occasioned, e.g. through answering the question asked of the VUI (see 5.3.2). In this regard, device use was also shown to be accountably occasioned through talk so as to introduce new information to the conversation (see 4.3.1 and 5.3.1) and to contest an argument (see 4.3.3). Therefore, what this thesis has presented is empirical evidence to show that device use unfolds with members making their device use naturally accountable to co-present others. The selection of these *types* of settings was based upon literature (see Section 2.1) claiming that device use was occurrent during interactions in them. The empirical data reveals how, as part of interaction in a casual-social setting, device use is certainly perspicuous, was made naturally accountable and, perhaps even, an *acceptable*, practice by members. Rather than glossing interaction as problematic, this thesis shows how members undertake interactional work to bring device use into interaction. Device use was brought into the conversations in the setting, either as a topic or point of reference, or as a means of contributing to the conversation (e.g. as in the cases where the device use was used to find new information, see 4.3.1) or to the device use.

In the second and third studies specifically, interaction with VUIs was occasioned by users as they establish and test the capability of the device as part of conversations about the device (see 5.3.3 and

6.3.1). VUI smartspeakers, as per their design by manufactures, were also used for long running tasks, such as playing music in the background at a party (see 6.3.2), or to play games together while having a meal at the dinner table (see 6.3.3). In the data presented in this thesis, members' projects involving the use of a VUI device were shown to be interleaved with the ongoing talk amongst collocated others in the setting. Furthermore, talk to the VUI was made accountable to the members of the setting in and through the conversation, and indeed held to the same normative moral order as talk between interlocutors in the setting (see 6.3.2). In this regard, talk to VUIs is crucially shown to be part of the multi-activity home (Rooksby et al., 2015), and unfolds alongside or interleaved within ongoing activities as an embedded activity in the home (Porcheron et al., 2018).

Of course, not discussed in this thesis are moments where the device use unfolded for purposes of responding to notifications or checking the time, or solitary use of devices without others present. However, this thesis did not set out to document all purposes of using technology in such a gathering, but rather, sought to explicate the interactional projects that occasion device use within a casual-social setting, and illuminate the naturally accountable ways in which this practice unfolded.

Summarily, interaction with all three forms of technologies turned upon both matters raised in the ongoing conversation in the setting and of those ostensibly unrelated to it. Members occasioned device use as a topic in and of itself, or as a resource to address members' interactional problems in talk, such as information deficits or to answer questions. The exhibits of data presented in this thesis demonstrably show how members treat device use within conversation as part of the activity of socialising as a group, rather than as a distinct activity. The next section delves further into how this device use was occasioned and used in conversation in a casual-social setting.

7.1.3 *Device use in conversation*

The above section details the nature of what devices were occasioned for in conversation, in order to attend to members' interactional projects. The final consideration here is of the how device use was brought into and used throughout conversation, and to make sense of how this was interactionally and accountably organised. Each empirical chapter has presented the methodical accomplishments of how and why members used their devices (see 4.4.1, 5.4.1, and 6.4.1 respectively). This section constructs an assemblage of these three chapters to reveal the practice of occasioning device use in conversation and how this device use is done during conversation. This is not to demarcate or even regard these as distinct *stages*, but to illuminate the ways in which device use is brought into conversation and accountably organised through it.

7.1.3.1 *Occasioning device use in conversation*

In all three studies, members routinely occasioned device use in conversation by *self-selecting to use their device*. This was done to contest arguments or answer questions posed in talk, or in the case of studying interaction around the use of touchscreen smartphones in a pub, also for remaining in touch with non-present others. Moreover, in the studies in the pub and café, members had conversations, such as about their favourite dog breeds (see 4.3.1) or whether animals have accents (see 5.3.2), both of which led to a device being used to resolve the occasioned interactional project. This echoes remarks by Brown et al. (2015) on the use of collaborative mobile search being a grossly observable feature of everyday interaction. Even in the home, participants made use of VUI smartspeakers for projects such as playing a game together while collocated around the dining table for a meal (see 6.3.3), raising how technology ostensibly gets used during other ongoing activities in the home.

Of course, members also *selected others to use their device* as part of the interaction in a casual-social setting. With touchscreen device use, this was typically done as another member had a device accessible, and a member oriented to this by asking them to verify a fact using the mobile Internet (see 4.3.1), an already identified use of portable devices in literature (Church and Oliver, 2011), although this thesis presents an empirical account of such an action. However, with voice-based interactions (i.e. both studies two and three), members also turned to asking others for assistance with VUI input if they ostensibly suspected that the device would not be able to 'understand' them.

In situations where the device use was occasioned in an ostensibly unrelated matter to the conversation, this was due to members *attending to an interruption or notification from their device* (e.g. a message arriving or an alarm sounding that may have been configured during a prior interaction with the device). It is this form of occasioning that has attracted some interest in literature that addresses the impacts of technology use in society, with many studies examining how systems might be designed to manage device interruptions and to select the most opportune moments to deliver notifications (Fischer et al., 2011, see 2.1.3). This thesis offers little contribution to this space when these occasioned interactions were not brought into conversation as this was out-of-scope of this thesis' objective in studying conversation in which device use was interleaved.

7.1.3.2 *Using a device in conversation*

People *interleave their device use within the ongoing conversation*, and accountably accomplish this through various means, including shifting their gaze between their device's screen or putting their device down or picking it up (in the first two studies), or by interleaving utterances directed at the VUI amongst talk to others in the setting (in the second and third studies). The conversation that is interleaved amongst this device use included questions in relation to how to complete the task for which the device use was occasioned, e.g. clarification of

the search terms to use (see 4.3.1), to ask another person to help with the device interaction (see 5.3.2), or in relation to other matters of the setting, such as parenting (see 6.3.3). As Brown et al. (2013) and others acknowledge, portable device use is part of the multi-activity of everyday life. Just as in the study of device use and television watching, in which Rooksby et al. (2015) remarked that "attention is accountably managed and organized in the course of watching [television] together" (ibid., p. 13), the findings in this thesis show that members' attention and orientation to device use is managed within a multi-party conversation in other casual-social settings too. In other words, the use of the devices was, in the same regard, identified as being accountably interleaved amongst the ongoing everyday activities by members as a part of those activities.

ACCOUNTABLE DEVICE USE

The notion of natural accountability is such that the members of this setting can both observe and provide a report on the action of others, and that those other members would recognise that report (Garfinkel, 1967, pp. 1–34, see 3.2.2). As discussed above, in the first study, members make their device use accountable through methods such as rotating their device, sharing the visibility of the devices' screen, or by verbally reporting the specifics for the use, reinforcing findings others have elucidated in relation to sharing activities using mobile phones (Raclaw et al., 2016). Although members can observe device use happening, and they can report for what ends the device use is happening as a result of its use in conversation, given the 'private' nature of touchscreen-based device interactions (due to the small screen size), members cannot observe or report on the *specifics* of what device use is unfolding (i.e. what the user is *actually* doing). In this, members rely upon the device user to make these specifics of device use accountable.

In the second study, members drew upon the same methods to account for the specifics of their device use, with members sharing the

visibility of the screen by holding it between themselves and another person, or by verbally reporting on the outcome of requests. As well as demonstrating the perspicuity of device use in these settings (discussed above in 7.1.2), this also reveals the way in which members methodically and accountably organise this interaction. Figure 7.1 below includes three examples from the first two studies of this practice of making the device use accountable through making the screen visible to another person:

(a) Leaning in to read the screen, from Section 4.3.1

(b) Rotating the screen to show an email, from 4.3.2

(c) Sharing screen visibility to show failure of device to respond to voice, from 5.3.3

Figure 7.1: Examples from the first two studies of individuals making device use accountable by sharing visibility of the device screen.

Furthermore, the specifics of input to the VUI were hearable to those in earshot, as a result of interaction unfolding through talk to the device. These requests to the VUI device were occasioned in and through the conversation, and as a result of the hearable nature of the request and context within which it unfolds, members' requests are made naturally accountable (i.e. the specifics of the interaction with the VUI are reportable as well as hearable). For example, in the cases presented in 5 there were conversations about animal accents, the time of sunset, and the capability of the VUIs, and each of these conversations established the context within which the VUI request was made. In other words, members' request to the VUI was hearable as a result of it being delivered through talk, and reportable as a result of it being occasioned in and through conversation. However, the

responses from the VUI typically defaulted to displaying information on the touchscreen of the device. Members reported these responses either by providing a verbal or visual account of how the VUI responded (e.g. showing the screen or verbally explaining the response), or through further VUI interaction (i.e. if the device did not respond they would perform further address to the device, through which it is established the prior request 'failed').

In the third study, interaction with a VUI smartspeaker was done entirely through voice and thus the specifics of their interaction with the device were hearable and made reportable, given that the use of the VUI unfolded as part of the ongoing interactional project that occasioned its use (see 3.2.2). The talk to the VUI was hearable through its utterance, and made reportable through its perspicuity to the ongoing activities in the setting (e.g. testing a new device in the home, playing music at a party, or playing a game together while eating dinner). However, members did provide preparatory accounts in the home in instances where, for example, others were expected to play a game together using the VUI (see 6.3.3). In the three fragments of data presented, other members in the setting (i.e. those who were present for the VUI use), were ostensibly able to recognise and practically reason about the interaction with the VUI device through the requests to and responses from the VUI. This point is exemplified most clearly in instances where the user's request 'failed', for example, and members collaborate on addressing the technical troubles with the VUI without invitation or detail on the problem by the 'original' VUI user. In this sense, the data demonstrates how members accountably recognise the initial request to the VUI and the VUI's failure to adequately respond to that request, and practically reason and respond to it by taking further action, either by repeating or rephrasing the request (and, for example, often with different prosody). In this, members did not explicitly report a failure of their request, as this was done through the use of the device, establishing its use as, perhaps, more 'public' in nature.

A final note on this issue is to remark that, of course, members accounting for their device use is directly tied to the cohort and setting within which the device use occurs. All the studies in this thesis focus on device use when people are with friends and family, and in a casual-social setting. This links with the prior discussion above with regards to the rules of such a setting by establishing what is acceptable (Laurier et al., 2001), and furthermore, it establishes the context in which the expectations for how to deal with situations where technology use occurs are managed:

> [T]here are many expectations about appropriate engagement with various technologies, including mobile phones—and more specifically texting on mobile phones—that are to do with how those particular cohorts organize their everyday affairs
>
> — Tolmie et al. (2008, p. 262)

In this, the practices explicated in this thesis represent and speak to the members of these settings, and this context is imperative to making sense of these findings.

COLLABORATING ON DEVICE USE

Members also included others in their interactional projects as a collaborative activity. This either relied on *assistance from others* such as guidance on what to search for, spelling (primarily study one), or verbal reports of the specificity of what was being done with the device. It also relied on *sharing control* of the task by multiple members being involved with the task either on the same device, or different devices (Brown et al., 2015).

In the case of the first study, members always retained control of the use of their own smartphone by holding it in their hand, although, they did ask others for help with search terms (see 4.3.1) or for performing requests to the VUI if their device did not 'understand' their pronunciation (see 5.3.2). This suggests that the nature of *personal* portable devices such as smartphones supports collaborative prac-

tices *through invitation*. In the second study, another co-present member attempted to complete the task on their own device as someone was struggling to use their VUI, *without invitation*. Moreover, in the third study, featuring only shared VUI devices, members self-selected to assist in completing members' interactional projects, again by 'taking turns' to use the VUI without invitation.

In this regard, interaction through voice, i.e. talk, remains the most obvious way in which members' conduct and use of devices is made recognisable and accountable to others, e.g. as part of the conversation. With the accountability of device use, co-present members are able to *recognise* and ostensibly self-select, without invitation, to involve themselves in the interactional project. This led to members answering questions directed at VUIs (see 5.3.2), assisting with early termination of the VUI response (see 6.3.2), or assisting with the initial device user to complete their interactional project (see 6.3.3).

SUMMARY

This ethnographic study has shown how members were able to interleave their use of devices in conversation to complete their interactional projects. This rubs up against critiques others have made that suggest the use of—or even the mere presence of—devices in conversation have an isolating effect (Turkle, 2011). This thesis, through the adoption of an analytical lens that is 1) agnostic to the morals of actions and unaccountable factors, and 2) is used to reveal the accountable situated action of members within the setting, and shows how this interactional work was co-accomplished by members in the setting as *part-and-parcel* of the face-to-face conversation. In other words, *members accountably attend to interleaving device use in conversation*, and that with the three technologies studied, this use was shown to be collaborative at times. This collaboration turned upon the specificity

of the device use being made observable and reportable[3], and with touchscreen-based interactions this was done through visual or verbal reports. In the cases of talk to VUIs, the interaction itself was hearable and through its utterance as a situated action, and made the specifics of the user's device interaction accountable to the setting as part of the interaction in the setting. In turn, and any troubles with the technology that users experienced were attended to as matters concerning the occasioned interactional project.

Through the presentation of empirical data, members' interactions with devices were interleaved amongst the conversation in each setting. Members did this through occasioning the device in talk or through the conversation occasioning the device use. They accounted for interactions with the device, engaged others in their device use through questions and requests for assistance, and ostensibly treated interactions with the device as part of the normative moral order within the setting. The next section turns to discussing a crucial factor in relation to the observations that unfolded: the methodological considerations of this research.

7.2 METHODOLOGICAL CONSIDERATIONS

The objective of these observational studies was to understand how members practically attend to the matters of using a device in a multi-party conversation. Accordingly, the analytic orientation of ethnomethodology was adopted. Crucially, ethnomethodology provides the analytic lens to explicating the members' practical action and practical reasoning (Crabtree et al., 2012, p. 27) to make sense of and reveal their methodical accomplishments as situated action (see 3.2.2).

As the literature review established, this thesis is far from being the first piece of work to adopt an ethnomethodological orientation

3 It was always recognisable to members of the setting that a person was *using* a device, but only the use of a VUI intrinsically makes the *specifics* of that use hearable to those in earshot in—and naturally accountable through—its use in conversation.

to studies of everyday life (see 2.3), nor is it the first piece of ethnomethodological work to 'create' the situation in which the study was to take place (i.e. participants were recruited to go to a setting, rather than the researcher going to a setting in which participants are already assembled). For example, L. Suchman (1985)'s work, Plans and Situated Actions, which has profoundly influenced HCI and CSCW-based studies of interaction, adopted an "uncontrolled experimentation" (ibid., p. 114) approach to studying the use of an agent-based photocopier (see 2.3.1). In her work, although the basis for on participants were using the photocopier was a research study, participants' interactions were unscripted and unguided, or rather, "uncontrolled". This is the overall approach taken with this thesis, insomuch that although participants were recruited to take part in a research study, there was no specific activity or task for participants to do, other than socialising together 'as they normally would'. The participants can assumedly be considered competent for this as they were recruited as families or groups of friends.

Although the analytic perspective adopted in this thesis was uniform across the three studies, the first two consisted of a video-recorded observation for ninety minutes, whereas the third consisted of contextual audio recording in the home over one month. The study design decisions were based on the appropriateness, or perspicuity, of the device interaction that was of concern to the setting. This section examines two key aspects that are relevant in terms of the methodological contributions of this thesis, namely the application of this approach in relation to the selection of settings perspicuous to the studies, and of the validity of the methodological choices taken.

7.2.1 *Choosing a setting*

The first two studies in this thesis were video-based ethnographic studies. Participants were recruited for the purpose of socialising as groups of friends who would usually socialise together in either a

pub or a café. This thesis follows in the tradition popularised in the CSCW literature of studying the work of a specific setting, such as traffic control rooms (Bentley et al., 1992).

In reviewing the progression of the development of HCI studies, Grudin (1990) remarks that:

> [There is] increasing preparation for the next outward step of the interface, into the social or work setting [...] since most work occurs in a social context, computers will support it more successfully if they implicitly or explicitly incorporate social and organizational knowledge.
>
> — Grudin (ibid., p. 264)

In this, Grudin (ibid.) provides the rationale for engaging in ethnographic studies of social settings in which technology use would eventually unfold, with ethnomethodology especially suited to this cause given its attention to the members' methodical accomplishments. Furthermore, Heath et al. (1994) cautiously summarise what they called the "lack of success of CSCW systems":

> [T]he lack of success of CSCW systems derives not so much from their technological limitations, but more from their insensitivity to the organisation of work and communication in real work environments.
>
> — Heath et al. (ibid., p. 155)

This thesis, of course, studies non-work settings, but given the now widespread use of devices in casual-social settings (see 2.1), there is an established case to undertake studies to reveal the details of the social organisation of interaction in everyday life settings, to support the design of technologies that are used within. The rise of ubiquitous technologies ratifies the need to study everyday life in which these technologies are used—this is part of the *turn to the social* (see 2.3.1).

7.2.1.1 *Choosing a public casual-social setting for portable device use*

Ethnographic studies in CSCW transgressed on to examining the particulars of everyday life, such as watching television (Rooksby et al., 2015) or families eating together at the dinner table (Ferdous et al., 2016). This is the work that this thesis methodologically follows. A crucial factor that unfolds in all of the prior studies discussed, and that is fundamental to any ethnomethodological study (see 3.2.4), is the perspicuity of the interaction to that setting. For example, concerning studies of how couples watch television together and use a mobile phone, it deductively follows to capture data in the main room of the home in which television watching occurs. Likewise, it follows to examine the use of technology at mealtimes at family dinner tables. To summarise, the site of the study naturally follows from the activity to be studied.

With this thesis, however, greater justification is given to the selection of settings because such a choice of setting is ostensibly less deductive. With the two video-supported observational studies, two public settings were chosen for the research to take place. The type of setting was initially conceptually identified as a place where people gather and socialise together. The settings of a pub and café were then selected for a variety of reasons (see 4.2.1 for justification of a pub and 5.2.1 for justification of a café). Summarily, however, there was already literature that revealed the social and relaxed nature of interaction in these settings, and of the device use in these settings. Moreover, personal experiences of technology being used in these situations further influenced this decision.

The selection of these two settings required pre-negotiating access with business owners to ensure studies were able to take place and to ensure procedures were in place for dealing with inadvertent data collection of members of the public (e.g. people passing through the background). However, while Rooksby (2013) argues for such studies taking place in a lab-based setting, it was decided that as there were not factors that needed to be controlled, and as there was no equip-

ment or task other than socialising needed, there were no overriding benefits to running a pseudo-realistic lab-based study over running studies in an 'actual' setting. Furthermore, obviously, others raise issues with this, remarking that laboratory studies are "hardly the stuff of ethnomethodology" (Dourish and Button, 1998, p. 8), underscoring a need to get as close as possible to the phenomena of interest rather than reliance upon creating a simulated setting.

7.2.1.2 *Choosing a home setting for VUI smartspeaker use*

Given the underlying emphasis to study conversation around the naturalistic use of technology, and to ensure consistency with the first two studies, the third study took place in a setting perspicuous to the use of the device—the home. Of course, VUI smartspeakers were designed for places such as the home, and thus this outcome was straight-forward. The challenges of studying interaction with new technology are, however, that users may not have competence in operating it. In some experimental studies of new voice interfaces, for example, researchers have provided training (e.g. Molnar and Kletke (1996) and Schaffer et al. (2015)) to ensure users' competency before an experiment with a VUI. The goal in this thesis, however, was to examine the interaction that unfolded around a device, not just with it, or as part of an initial encounter with the technology, or by following training or guidance on how to use the technology. The goal was to understand how these technologies are used as part of routine interaction in the home. Therefore, to get closer to the phenomena of using these home-destined technologies, a more longitudinal approach to the study was necessitated.

As discussed previously within the empirical chapter concerning conversation around the use of VUI smartspeakers (see 6.2.2), this study relied upon audio collection only, and upon selective recording rather than continuous recording. Given the longitudinal nature and the setting in which the study took place, these decisions necessitated careful consideration of how to capture data ethically, sensitively, and

practically. Other approaches to studying technology in the home include repeated interview visits (e.g. Fuentes et al. (2019a)) and diary studies (e.g. Forlizzi (2007) and Jokela et al. (2015)), yet here there was an intent to actually understand the situated action of members that cannot be explicated through such methods. Furthermore, although some studies of voice interfaces can rely upon self-reported logs of devices (e.g. Ammari et al. (2019)), given the need to examine interaction around the use of the devices this was not a practical method to adopt. Others adopted methods such as relying on participants to start or stop recording devices before or after the interaction in a space or with a device, as previously done with portable device use during television watching (Rooksby et al., 2015). This would be problematic in this study as the use of VUIs can be started and finished in under a minute without much preparation. Watching a television programme may consist of being within a specific space for thirty minutes or more, for example. On the other hand, using a VUI may take a matter of seconds with a user simply 'passing through'.

Therefore, there was not a practical approach within the existing literature on how to accomplish the data collection for this study respecting the constraints outlined above. To achieve the goal of this thesis, a specific recording device was designed and created for this purpose (and has since been released as open source software[4].) to selectively record interactions triggered by nearby users uttering a word. This recording device (known as the CVR) allows a longitudinal study to take place, in which participants learn (or not) how to use the VUI within the home without guidance from researchers, in line with the prior two studies' approach of not guiding interaction with devices. The CVR is always 'listening'—much like the VUI smartspeakers—and retain the last minute of audio in memory. When the programmed word is spoken, the device saves this prior minute and records for one further minute (extending this recording if the interaction with the device continues).

4 See https://github.com/MixedRealityLab/conditional-voice-recorder

Furthermore, it allows for fewer data to be collected, to be done so ethically without capturing all matters of home life, and to not rely upon participants to manage their involvement in the study (i.e. members can focus their efforts on their normal mundane activities in the home as opposed to concerning themselves with the data collection).

The audio collected in this study provides a rich insight into the interactions in the home, in much (although not entirely) the same way as video data:

> [While video data] can form an archive, a corpus of data that can be subject to a range of analytic interests and theoretical commitments, providing flexible resources for future research and collaboration.
>
> — Heath et al. (2010, p. 2)

However, this brings with it a set of limitations, as Crabtree et al. (2012) elaborate:

> [Y]ou cannot see what people are doing alongside of the talk and there are circumstances where this may matter. [...] Always be prepared to elaborate with notes the surrounding action that envelops the sequence of talk you are recording.
>
> — Crabtree et al. (ibid., p. 82)

The approach in this study meant that fieldnotes also could not be taken as data collection was to take place over an extended period without a researcher present. These two factors mean that this thesis presents only partial records of interaction in the home; however, given the parameters outlined above, and the focus being primarily interaction around the use of the VUI device, this was seen as an adequate compromise. What this situation means is that the data presented in this thesis comes with caveats, such as the inability to comment on conversations relating to the VUI device away from the device, or matters that influence the VUI use which unfold over a minute before or after interaction with the device. These caveats limit the drawable

conclusions this these can make with regards to commenting on the specific families' use of, and conversations about, different technologies in the home. However, they do not preclude the examination of how their specific interactions with and around the device unfold *in vivo* where it was recorded.

7.2.2 *Methodological validity*

The approach that was taken in this thesis is not laboratory-based given the clear emphasis on selecting settings in which device use unfolded. In the first two studies, the settings were pre-selected as part of the study design, the participants were all recruited as groups of friends to take part in the study. In the third study, households were recruited as a family to take part in the study together. The purpose of each study was described as one in which interactions with and around technology were to be observed. Such an approach precludes conclusions of matters relating to why device interaction occurred as reasons of motivation, or other non-accountable factors. These issues were disregarded in any case, given ethnomethodology's orientation to the naturally accountable activities of members only. Therefore, this thesis' approach to studying the interactional minutiae of members' accomplishment reveals how and for what purpose *in conversation* device use unfolded.

Moreover, in line with existing work in ethnomethodology on the notion of validity, this thesis does not pretend to demonstrate *all the ways* in which *all interactions with and around devices* might unfold. This is a key tenet of ethnomethodological studies in that, through the presentation of the ethnographic record:

> [M]embers can recognise the work of a setting and also, as they are known and used in common, the machineries of interaction that they employ to accomplish and organise that work too.
>
> — Crabtree et al. (2012, pp. 155–169)

Thus, it is established that the work of producing the ethnographic report of members' actions is such that members can read and recognise the methods that are explicated. This is because this record is constructed using the recognisable accountable practices of the members in the setting, rather than theorising about members' actions. In this sense, this thesis presents the members' practical action and practical reasoning, rather than the analyst's theoretical reasoning. The reliance upon only examining the accountable actions precludes the production of generalisable statements, but also underwrites the validity of the findings of this thesis.

In summary, this thesis selected settings which were perspicuous to the activity under investigation and recruited friends and families to take part in a research study. Such an approach precluded discussing motivations for the device use that occurred, however, the analytic orientation of this thesis also precludes such a stance (given its emphasis on revealing the naturally accountable activities of members, rather than assembling a theoretical understanding of their actions). It is from this regard of producing a record of accountable actions recognisable by members that establishes the validity of the approach taken in this thesis.

7.3 INSIGHTS FOR DESIGN WORK

This chapter's last reflection is upon how interaction unfolded across the three studies with an insight to supporting future design work. Crucially, this chapter builds upon this thesis' methodological contributions to further examine members' conduct, to identify the collaborative efforts, and how HCI and CSCW might respond to these efforts.

This section will reflect upon the multi-party device interactions that unfold in each study, how this turns upon the accountability of device use, and how members collaborate as part of their interactional projects. This will return to the case that the design and use of VUIs is made naturally accountable such that users can involve themselves

in interactions collaboratively without invitation, and that this is of relevance to existing literature in HCI to design collaborative systems for collocated interaction.

7.3.1 *Collaborative device use*

Previously, this thesis introduced work in *mobile collocated interactions* and HCI which focused on the notion that device ownership will, in the future, occur with *shared* devices, and that these devices will support *multi-user* interactions that are *collaborative* (Lucero et al., 2010b, see 2.2.2). Indeed, the findings of the empirical chapters in this thesis show that, in each of the studies and with each modality of interaction with a device, members engaged in collaborative device use. For example, with touchscreen interaction, members were observed collaborating, with one member providing the query terms for a mobile search to be completed by another as part of the interactional project occasioned by their conversation (see 4.3.1). In another case, one member proposed a rephrased request to the user of a VUI, given that the members practically reason that the device 'misinterpreted' the prior request (see 5.3.1). In a third case, the members of a setting took turns trying to start a game by issuing new requests to the VUI, varying the prosody and words used in response to the failures of prior requests (see 6.3.3). This first case, for example, augments the existing literature that identified collaborative mobile search as an everyday task. This examination by Brown et al. (2015) identified that there was

> [...] considerable attention, effort and thought given to co-conversationalists while using a mobile device. Rather than shutting off conversationalists from each other, the devices become a site of investigation and discussion.
>
> — Brown et al. (ibid., p. 516)

This thesis extends this finding to encompass mobile device use for other purposes too, and how such practices unfold with both touch-

screen device use and VUI use, and that members bring this use into conversation in a casual-social setting.

Portable devices such as smartphones are inherently personal in their design and use (Lucero et al., 2010a), engendering 'private' use whereby even if co-present others are aware of—and can observe—device use unfolding, they often cannot observe or report on the *specifics* of that use (e.g. they might not have a line of sight). This is due, in part, to the relatively small size of device screens which inhibit greater visibility of the screen by those who are present beside the user. Attempts to disrupt this private nature of device use have included adding large screens to settings to encourage users to share content from devices (Lucero et al., 2012). However, the findings from the first study show how devices are used as part of collaborative efforts by members, by the user asking others for assistance in completing their interactional projects, irrespective of the small size of the screen, or by making the screen visible to others. As they did this, they made the device interaction accountable to others the setting, revealing the specifics of what was being done with the device, and thus transformed the 'private' device use to one where the specifics of that device use were observable to *some* others (i.e. this was not 'public' use, however, as the screen may have been made visible only one other member).

Following the reflection of the findings from the first study, this thesis posed the question of how the practices of device use could unfold when the interaction mode was augmented with voice input (see 4.4), given that in such cases the talk to the VUI would be hearable to those in earshot, making the specifics of users' actions hearable. In particular, this outlook asked how members would use a VUI as part of a gathering in a casual-social setting (see 4.4.2). As Crabtree et al. (2012) remark:

> [T]alk is the most obvious and pervasive way in which members conduct their work and make whatever it is that

> they are doing into an intersubjectively recognisable and naturally accountable activity.
>
> — Crabtree et al. (2012, p. 44)

The second study addressed this matter and revealed how members' talk to devices, as interleaved within conversation, made the specifics of their actions reportable, with the performance of the request organised within the organisation of the conversation in which the device use was occasioned. Additionally, this study brought to the fore that by switching the interaction mode of the device to voice as well as the use of the touchscreen, members ostensibly *self-selected* to involve themselves with ongoing device tasks. For example, they chose to perform requests on their own device if another member was struggling (see 5.3.1), or they offered assistance to the member to help diagnose problems (see 5.3.2). In this sense, device users did not necessarily account for the specifics of their device use, because their specific input to the device was hearable and made accountable through coherence with the ongoing conversation in the setting. The action of uttering a request to the VUI, in turn, was shown to occasion other members' self-selecting to respond to the device user's request (in other words, talk to the device was responded to by other people who were not the recipient of the utterance). However, members provided accounts, especially to provide the details of the response from the VUI. At times these accounts were implicit made by members (i.e. when a member uttered a subsequent identical request, through which they establish the failure of the previous request). At other moments, it included members explicitly confirming the success or failure of the VUI to respond to the request. In this, through the utterance of the request, the device interaction with the technology becomes 'semi-public', insomuch that members' requests to the VUI were made naturally accountable through their use as part of the social interaction in the setting, yet members were relied upon by others for verbal or visual reports of the VUI's responses given that these responses were displayed on the device touchscreen.

The third study moved this examination into the realm of interaction through voice only, and where the technology under study became a shared device in the home. As such, although the observations in the study of VUIs on portable devices explicated the hearable nature of the request made to the VUI, and its coherence to the conversation establishing its natural accountability, in the study of voice-only interactions, the findings reveal how all members of the setting could orient to and attend to matters of the device interaction. The hearable nature of requests was again revealed, with members' requests being explicitly held to the normative moral order by those in the setting (see 6.3.2).

Through the careful reflection on members' conduct and how they accomplished their collaborative efforts to complete the occasioned interactional projects, this section has identified collaborative action amongst members in such settings, and this establishes the case for further studies in CSCW to critically examine the nature of studying interaction in such settings. Some literature has examined this (as discussed above, see 7.1), but crucially, what this thesis shows is how there is further cause to examine such settings to reveal the collaborative efforts within, to support design work. Although the definition of a casual-social setting was broad, as a concept it builds upon the work of others who have called for ethnographic studies of social settings (Grudin, 1990). The need to study the organisation of interaction of settings in which technology is to be used collaboratively is well rehearsed (e.g. Crabtree et al. (2009) and Heath et al. (1994)), yet the work in this thesis suggests that such technology was shown to be deficient in meeting members' needs, e.g. multiple interactional projects were left unresolved. Nevertheless, as this thesis has argued, members undertake interactional work to account for and accomplish this collaborative effort and the 'need' to 'successfully' complete an interactional project ostensibly does not exist in a casual-social interaction, therefore CSCW should take the opportunity to identify ways to ameliorate these challenges through further ethnographic work.

The concept of a *casual-social setting* is established around the notion of members' conduct, and this thesis has unpacked three such studies of that. Members' conduct is, of course, cohort dependant (see 7.1.3.2) and thus, further studies are needed to reveal more about the interactional projects, and ways in which members' problems are addressed with technology. The next section further reflects upon these responses from the VUI and how the collaboration that unfolded turns upon them.

7.3.2 *VUI responses as supporting further action*

This final discussion section examines an area of concern for HCI in the challenge of designing interfaces to support collaborative action amongst members of the setting. Collaboration, as discussed above, unfolded as an interactional accomplishment in each setting, as members worked together to complete their occasioned projects. With the latter two studies, however, this collaboration ostensibly turned upon the hearable nature of the interaction with the device. This section critically examines how this interaction could be further examined, linking to existing publications in HCI, to design interactions that meet members' interactional needs.

In reflecting upon the study of VUI smartspeakers in the home, VUI interaction is shown to consist of the form *request to VUI-response from VUI*. At times, and in all the cases of VUI device in the home presented in this thesis, a response from the VUI is followed by successive requests by the members of the setting. Through the analysis, the VUI responses themselves are analysed by members for the 'account' of sorts they provide on the state of the VUI device, and its processing of the previously made request, i.e. the response (or not) of a failed request occasioned members to practically reason and perform further actions to accomplish the interactional project. The data from both studies two and three suggest a level of inadequacy of some responses *as resources* to furnish this analysis to proceed with the in-

teraction. Consider how members in the third study fail to start the Beat the Intro Skill due to the repeated non-obvious source of technical trouble (see 6.3.3) before the device ultimately provides a response to a request that includes a partial transcription of the user's spoken words, and the VUI offers a next candidate action. This action provides a mechanism to practically reason about the processing of the voice-only device. Of course, with touchscreen-based devices, this mechanism for users to reason about VUI troubles was provided through messages displayed on the touchscreen, but these were only known to others if the device user 'made them available' by reporting them (verbally or visually).

This thesis argues that the accountable nature of a 'verbal' response from a VUI smartspeaker is used as a resource by the original VUI user, and other co-present members, to make successive requests. It is through the natural accountability of members' successive requests to the VUI, and how these turn upon the response from the VUI, that reveals the practical reasoning of multiple members, collaborating to complete the interactional project. Members demonstrably show reasoning about failure through a discussion in conversation (see 5.3.1), or through repeated utterances to the device by different members with different prosody (see 6.3.3). In both cases, it was this natural accountability of the response, established through the ongoing work of the setting, that provided the resource for further action. Porcheron et al. (2018, p. 10) term this characterisation of responses as "resources for further interaction". In examining interaction in a multi-party setting, this resource is established as supporting further interaction by members of that setting to complete the interactional project.

Designers might, then, consider how members attending to technical troubles with the VUI are also attempting to practically reason about the source of trouble, be it a system problem or a transcription problem. The response (or no-response) is treated as an indicator of what this trouble is. With the GUI on smartphones, the VUI displays text to 'explain' the processing of the user's request, and with

screenless devices, this is done through the verbal response (or lack thereof). At times, the VUI provided a reference to the next candidate action of the device, or the transcription it generated of the user's spoken words, and what provisionally might or might not happen next (see 6.3.2 or 6.3.3). Porcheron et al. (2018) remark how, adjusting, or 'designing' these responses to display this processing, would provide users with resources that can support and occasion further interaction with the VUI device.

In summary, with VUIs, the specific request was shown to be made an accountable accomplishment through the conversation in which it was interleaved (Porcheron et al., 2017b), and with VUI smartspeakers, the hearable response from the VUI was made accountable through both conversation and the request which triggered it. In returning to consider this development in terms of the mobile collocated interactions literature, this thesis remarks upon how the use of the VUI as occasioned in conversation accounts for specific features of device use grossly observable and reportable through talk to all those in earshot. Much of the design-focused literature in mobile collocated interactions has focused on creating technologies for collaborative use between collocated people (e.g. Lucero et al. (2010b) use mobile devices for brainstorming) using touchscreen technologies. One such opportunity might be the inclusion of a voice-interface in the design of such systems to transfer interactions with devices to a more 'public' sphere. The notion that voice-based interfaces could be designed to support collaborative interactions between co-present others has been used in some design efforts in CSCW (Jones et al., 2012). However, through the presentation of empirical data, this thesis reinforces and validates the notion that the adoption of a voice-based interface has the propensity to allow members to collaboratively complete interactional projects occasioned in conversation. However, through the analytic orientation of ethnomethodology, this thesis also reveals how it is *not enough* to produce a VUI to ensure collaborative activity between members—collaboration in these settings turns upon both the practical work of

members to make activities accountable, and of the resources the VUI provides to users, supporting those who are co-present to take further action in completing their interactional projects. This thesis stops short of offering implications for what constitutes 'ideal' interactions with a VUI, but the notion of designing VUIs for collaborative action should now be perused as an avenue in future HCI work, extending prior efforts on collaborative mobile systems (see 2.2.2).

7.4 SUMMARY

This discussion brought together the main strands of this thesis in order to answer the aims and research questions posed. It discussed the conceptual nature of casual-social settings, how conversations unfold in such settings, and how device use is accountably and interactionally organised in and through conversation (see 7.1). By bringing together the methodical accomplishments of members in using devices in interaction, this chapter demonstrates how devices are used as part of socialising together in a group and for the purposes of addressing the members' problems occasioned in conversation. Members also used devices for other purposes, and naturally accounted for this as part of the conversation (e.g. to make jokes, or play games together, and so on).

The chapter then progressed onto making a number of methodological considerations (see 7.2), by reflecting upon the ethnomethodological perspective and practical approach taken in this thesis to the three studies. Each study in this thesis selected a setting that was established as appropriate for the specific technology under examination. This, combined with the careful approach to explicating members' naturally accountable actions, underscored the validity of the approach taken in this thesis. This thesis produced thick descriptions of members' actions, and through the recognisability of these methodical accomplishments, this validity is established.

Finally, this chapter turned to discus elements of the collaborative interaction that unfolded within the data (see 7.3). These instances were identified as turning upon the accountability of device use, with the requests to (and from) the VUIs making the specifics of the interaction hearable and reportable through the work of the setting. The collaborative efforts of members were also identified specifically as an accomplishment of members' efforts in the setting through their use of the device. By considering the existing literature on designing collaborative software for use with portable devices, this thesis posits that the use of a VUI has the potential to support collaboration on members' interactional projects in such settings. Crucially, however, this thesis shows how this requires careful design work through considering the specifics of these requests: it is not enough to just introduce a VUI to a setting, one must undertake work—which is beyond the scope of this thesis—to consider how requests to and responses from the VUI may be used to accomplish members' interactional projects in a casual-social setting.

8

CONCLUSIONS

This chapter summarises and brings together the conclusions of the machinery of members' conduct in using a device in and around conversation in a casual-social setting, the collaborative nature of this device use, and the methodological contributions of this thesis. Finally, this chapter concludes with a summary of the potential future work that could follow on from this thesis.

8.1 SUMMARY

This thesis made a case for studying device use in conversation (see 1.1), adopting a practical approach that eschewed judgements on the values of device use and instead focused on the practical accomplishment of how devices are used during multi-party gatherings in casual-social settings. This case was motivated by a paucity of data on how and for what purpose devices are used in such settings, against a backdrop of existing literature that characterised such use as problematic. The research in this thesis was fundamentally interdisciplinary in its approach and drew upon existing work in Mobile Human-Computer Interaction (HCI), ethnomethodology, and Computer-Supported Cooperative Work (CSCW).

Through the analytic lens of ethnomethodology, this thesis explicated both the problems that occasioned device use in and through the ongoing multi-party conversation, and how members brought their devices into a conversation to address these problems. These problems form the foundation for the interactional projects—ranging from information deficits through to desires to play games—occasioning the use of the device as an activity that unfolds within the setting.

This thesis consisted of three studies, each with a focus on the social organisation of groups of people socialising together and making use of a particular technology. The first study examined conversation around the use of touchscreen portable devices such as smartphones, the second studied conversation around the use of the Voice User Interface (VUI) on touchscreen portable devices, and the third studied conversation around the use of screenless VUI devices, marketed as smartspeakers. Each set of studies took place in a specific setting: the first was in a pub, the second in a café, and the third in participants' home, all of which were considered to be 'casual-social'.

Through explicating the members' methods of using a device in conversation in these settings, this thesis shows how members accountably organise this use and take part in collaborative activities to complete these interactional projects. This collaboration was identified as turning upon the natural accountability of device use, which members accomplished through making the specifics of their use visible and reporting upon it in interaction. When device use was done using a VUI, the use itself was made naturally accountable through the talk to the device interleaved amongst the conversation that occasioned it. In the case of VUIs that feature no Graphical User Interface (GUI) (i.e. the smartspeakers), the responses from the device were oriented in the conversations, with members demonstrably attending to—and substantiating the natural accountability of—these responses through the coherence of the initial request and ongoing conversation. Through the examination and reflection of the data presented in this thesis, the discussed has illuminated how members ostensibly treat the use of devices perspicuously. Furthermore, it was identified how such device use, rather from detracting from the ongoing conversation, became embedded within the activity of socialising together. In this sense, the device use did not unfold instead of—or 'isolated' from—the social activity, but was performed as an activity within the work of socialising together as a group in a casual-social setting.

These studies inform and further encourage HCI work to examine ways of creating technologies to support collaboration on portable and readily-available technologies while people are socialising in groups. Crucially, by demonstrating the intricate ways in which members accomplish collaborative action to complete interactional projects, drawing upon the resources provided by the VUI, this thesis identifies a topic for further examination within both HCI and CSCW. Secondly, by exploring members' collaborative actions, this thesis remarks upon moments of when interaction was undertaken with a VUI smartspeaker, it was the VUI's responses that reported upon the success—or not—of requests made to the device. This identified a key concept to consider for future work, reinforcing findings published as a result of the research in this thesis (see Porcheron et al. (2018)), that classify the responses from VUIs as *resources* for further action as members' problems are addressed.

8.2 CONTRIBUTIONS AND KEY CONCLUSIONS

This thesis makes three key contributions, corresponding to the research questions posed in 1.3:

1. Explication of the *members' methodical accomplishments* of using a device in casual-social gatherings, detailing how they bring devices into an everyday multi-party conversation, and offering an insight into the differences of how device use is used to address the members' problems that arise in such settings,

2. Development of the *methodological approach* in this thesis, both in terms of the application of ethnomethodology and of the nature in which *these technologies* were studied *in these settings*, and

3. *Conceptual insights* of the nature of studies in casual-social settings and 'device talk' to VUIs.

These contributions are summarised in the following sections.

8.2.1 *Members' methodical accomplishments*

The methodological approach taken with this thesis included capturing 'real-world' empirical data of groups of friends or family conversing in a casual-social setting. This thesis defined the conceptual choice of setting as one in which friends and family gather to socialise and relax. Through the analysis of the captured data, the empirical chapters in this thesis progressively explicated the accountable practices of the members of each setting as they interleaved their device use in and through conversation. The orientation to accountable actions of members in the setting allowed for the formation of thick descriptions that detailed the interactional methods people employed as they attended to the interactional projects for which device use was occasioned.

Each study identified members' methodical accomplishments that revealed exactly *how* members of the setting brought a device into the ongoing conversation. These were brought together in the discussion to reveal how the overall accomplishment of using the three different technologies in conversation is done (see 7.1.3). Crucially, the findings reveal how, through using a device in conversation, the natural accountability of members' actions is accomplished as an outcome. With touchscreen devices, the interactional work to account for device use was done through methods such as articulating what was being done with the device, or making the device screen visible to others. In cases where the interaction was with a VUI, the interaction with the VUI (i.e. the *talk to the device*) typically made the specific nature of a member's device use both observable and reportable within the context of conversation within which it was occasioned. Nevertheless, members accounted for this use by providing further information on what occasioned their use through talk (e.g. they recalled a recent news story), or by providing preparatory accounts for what the interactional project they were about to undertake concerned. This accountability provided the resources for other collocated members to ostensibly self-select to become involved in

the interactional projects, and in turn, at situations where members had technical trouble, collaborate with the device interaction. In the case of VUIs on portable devices such as smartphones, in contrast to standalone smartspeakers, the response from the VUI was typically returned via the built-in touchscreen. Through using a device in these settings, members made this response accountable by sharing visibility of the screen, providing verbal reports of responses from the VUI, or through further interactions with the VUI (e.g. repeating a request), which recognisably establish the case for members reasoning that the device failed to respond to a prior request.

Through these three studies, the notion that device use is a mundane phenomenon in the settings is established. Members' problems, established through the conversation of the setting, occasioned the use of technology—ranging from instances of using a device to find new information in the conversation through to playing a game together—such that the use of the device was *part of* the multi-activity in the setting. Conversation topics were also brought about as a result of device use, with members introducing an email to the conversation to make a joke, through to projects such as testing the functionality of the VUI. In these instances, again, device use was occurrent alongside—or interleaved with talk—and was brought into the activity of socialising as an ostensibly mundane activity within the setting. Therefore, this thesis concludes that device use is fundamentally *embedded* in the activity of socialising in a casual-social setting.

Furthermore, as a result of the naturally accountable nature of interacting with a device, dealing with technical troubles and issues in completing device tasks were shown to unfold as collaborative activities. In this regard, members collaborated on mobile search tasks using touchscreen devices, attempted to find new information with VUIs on smartphones, and attempted to start games using VUI smartspeakers in the home. In both of the studies involving VUIs, this collaborative practice unfolded whereby co-present others involved themselves in others' device use *without invitation* or a request for assistance. This

was revealed to turn upon the natural accountability of the use of the VUI, which made the specifics of the interaction hearable and reportable within the context of which the use was occasioned.

Crucially, this thesis challenges existing notions in socio-technical studies that the use of portable devices in such settings is a distraction from ongoing socialising (see 2.1.2). Specifically, this thesis' findings implicate a call for reconsidering how device use is articulated in publications: rather than treat it as incongruous or distinct from social interaction, this thesis emphasises a call for analytic orientations which regard it as accomplished *in and through* members' efforts to socialise. Even more pointedly, the findings suggest that device use is part of the *social order* in these settings (see 3.2.1). This does not invalidate prior findings by others, as such findings were perhaps indicative of the time in which portable devices were more nascent, compared with the ubiquity of device ownership at the time of these studies (see discussion of device ownership in 1.2). In retrospect, for example, Turkle's more recent calls to 'reclaim conversation' (Turkle, 2015, see p. 18) implicates as much: the work of conversing in these sorts of settings *has changed* as a result of technology use, which now ostensibly permeates it. However, to regard this change as negative or positive—which is a categorisation based upon morals—is something this thesis intentionally does not speak to.

Secondly, this thesis' work on voice interfaces—both mobile and in the form of smartspeakers—took place as the technologies were coming to mass-market. Literature that examines their use *in vivo* was nascent and thus this thesis contributes the first empirical account of making such technologies at home. However, studies of other activities in the making of technology at home bare analogous methods to the use of the smartspeaker. For example, in the same regard that *multi-screening* became embedded in the work of *leisure time* through sitting on a sofa in front of a television while making use of mobile devices to augment television watching (Rooksby et al., 2015) (see 2.1.1), smartspeaker devices were placed in *activity centres* in the home—

such as the living room or a kitchen/diner—and members made the use of the smartspeaker naturally accountable as part of the interactions in those settings (see p. 123).

Finally, the technologies under study in this thesis could still be considered to be single-user—even smartspeakers, which operate on a one-at-a-time person-agnostic operation. However, research in *mobile collocated interactions* has attempted to address this by exploring ideas for how to embody notions of groupware from CSCW in portable devices to support collaboration (see p. 26). This thesis specifically contributes the notion of how users' interactions with voice interfaces are transformed into multi-user experiences *by users* through interaction in spite of a lack of specific design cues. Members does this by accounting for device use and occasioning the device use in conversation. This was demonstrable in this thesis in many cases, but of note are also situations in which technical troubles with devices were collaboratively dealt with through conversation, especially with VUIs. For example, the resources made available through interaction with a smartspeaker—the hearable input and output—facilitate the collaborative actions of collocated users. The ambitions of the research agenda to design portable technologies for collaborative action could be extended by further exploring the use of voice interfaces, with the findings of this thesis demonstrating the potential of such technologies.

8.2.2 *Methodological approach*

This thesis adopted the methodological perspective of ethnomethodology (see Chapter 3) to unpack members' actions in the settings. The first two studies took place in the semi-public settings of a pub and café, with participants recruited for the purpose of socialising together, and to be recorded for their interactions. Participants were informed that the focus of the study was on the behaviours around the use of devices, but that there was no requirement to use any device during the study. In the case of the second study, participants were

asked to use the VUI on their smartphone instead of typing where possible, but were again told that this was not a requirement and that the use of the device itself was not required for the study. These settings were selected as they were perspicuous in regards to the technology being studied, i.e. existing literature and personal experience had already established that technology use unfolds within pubs and café among groups of people socialising.

These two studies were, in some regards, an "uncontrolled experiment" (L. Suchman, 1985, p. 114) of sorts, in which people were in the setting for the purposes of being participants in a research study, but that the focus of that study was not on the completion of a specific task or following a protocol. The researcher was present and socialised with the participants for all the gatherings in these two studies and followed a participant-observer approach throughout the gatherings. The gatherings took between 60 and 90 minutes, with the study ending at an agreed time with the participants—the uncontrolled nature of the study was such that the study ended where it was ostensibly deemed appropriate by the researcher. In these two studies, data were collected by video recording the study using fixed wide-angle cameras on tripods, audio recording using a voice recorder, and writing of fieldnotes after the study.

The third study took place in participants' homes and was longitudinal in approach, taking place over one month. Homes were selected as the site of study because these were the sorts of places that VUI smartspeakers were designed for—they are typically non-portable devices designed for use in homes or offices. A longitudinal approach was also adopted in this study given the recentness of the technology being introduced to market (i.e. less than one month on the UK), with the focus on the study being how the technology was used in everyday multi-party gatherings. In this regard, each participating household installed the smartspeaker in a communal area of the home. For data collection, audio capture was selected given the ethical and technical concerns of collecting data in the home (i.e.

participant uneasiness about recording video in the home, and the vast amounts of video data that would be generated). However, a further issue was that it would still remain problematic to continuously record audio in the home (e.g. necessitating an analysis that would be an insurmountable challenge in terms of the length of captured data). Therefore, an audio recording device was designed and built for this study that captured a minute before and after interactions with the VUI only, allowing researchers to make sense of the context within which the device was used (this was called the Conditional Voice Recorder (CVR), see 6.2.2.1).

Across all three studies, the same analytic orientation was adopted in unpacking the collected data. Each study adopted an iterative approach, in line with documented practices in ethnomethodology (e.g. Crabtree et al. (2012) and Heath et al. (2010)), to explicate members' interactional projects, and how they practically interleaved device use within conversation. Ethnomethodology's focus on the naturally accountable practices of members allows this analysis to present 'what is done in the doing', and consequentially support this thesis' underlying goal to reveal how devices are used in multi-party conversation in a casual-social setting.

This thesis' contribution in the form of the CVR provisions a data capture tool to longitudinally study interaction with a voice-driven technology in the home (see 6.2.2.1 for the design of the recorder). Approaches previously adopted to studying technologies in the home include auto-ethnographies, ethnographies, diary studies, log analysis, and interview studies (see p. 118), with each providing researchers with different levels of insight into the situated use of technology. This thesis' work in the home is congruent with existing ethnographic approaches and ideas from CSCW of placing research technologies in the home (Tolmie and Crabtree, 2008). As elaborated on by Crabtree et al. (2003), the challenge for designing technologies for the home is "how people live in the home, what they do when they are at home, and the potential role of technologies within the milieu of domestic

activities" (Crabtree et al., 2003); through its selective capture of interaction and the preceding/succeeding use, the CVR allows researchers to elicit such an insight with relative ease.

By building and deploying a technology to automatically selectively capture interaction with a smartspeaker over an extended period of time that incorporated ethical considerations of long-term data capture in a personal space, this thesis supports an ethnomethodological analysis by providing rich data that includes elements of the context of interaction. Other approaches, such as having a fieldworker 'on site', can become impractical when studying interaction for extended periods of time, with proposed solutions including asking participants to record their interaction (e.g. Rooksby et al. (2015)). Alternatives include continuous video capture, as done by Heath and Luff (1991) in their study in workplace collaboration, but such work was undertaken among colleagues and under a vastly different regulatory environment. This thesis took an automated approach that allowed for the collection of data with some semblance of context through the inclusion of data around device interaction.

The richness of the resulting data included in this thesis validates the approach for further studies of ubiquitous computing in the home, and perhaps other sensitive locations such as the workplace. Such approaches need not be restricted to voice interfaces, but could include other technologies too, detected through properties such as increased electricity consumption. As ubiquitous computing research continues to examine the Internet of Things (IoT)[1], approaches such as the automated data capture in this thesis are likely to become of even greater benefit. IoT technologies are increasingly incorporating elements of autonomy and portability (Fuentes et al., 2019b; Porcheron, 2015), raising challenges for the study of their use: understanding just how people deal with this autonomy in the home will be critical to ensuring systems meet the needs of users (Crabtree et al., 2003).

1 An umbrella term for the routine integration of Internet connectivity to everyday technologies, including sensors and home fittings.

8.2.3 *Conceptual insights*

The final point to address is conceptual insights arising from this thesis. The insights contribute to existing ideas and work in both HCI and CSCW. There are two key concepts this thesis focuses on: the concept of the activity-based settings, and the concept of 'conversation' with a VUI.

8.2.3.1 *Casual-social settings*

This thesis set out to study people socialising together in groups in what was categorised as a 'casual-social setting'. This sort of setting is referentially mentioned in a range of existing academic literature (see p. 2), as a place for people to gather and socialise together, but rarely is it designated as a site for empirical investigation. In seeking to clarify this type of setting, this thesis incorporated notions of other 'sorts' of settings, such as third places (Oldenburg, 1989). This provided the backdrop that established the concept of a casual-social setting as one in which people gather to relax and socialise, and that may be public or private. This definition encompassed each of the three study locations in this thesis: a pub, a café, and a communal area in the home; each of which were *perspicuous* to the device use under study (Garfinkel and Wieder, 1992b). The analysis in this thesis showed how interaction in each of these settings is replete with articulation work to naturally account for and interleave the use of devices within conversation, and that device use is a recurrent activity as part of socialising in these settings.

As far back as the early 1990s, there have been calls for CSCW to examine social settings to inform technology design (Grudin, 1990). What this thesis does, through this study in these three different settings, is reinforce and renew such a call for studying technology use in places where people are collocated: technology use is replete in such settings and there is a ripe opportunity for CSCW and HCI to understand and design technologies to meet members' needs. This

should crucially take place outside of workplaces and homes as distinct settings, and instead be situated in a range of settings defined by the interactional phenomena that is of interest.

Ellis et al. (1991) call the use of technology in such settings "same-time/same-place", distinguishing it from interaction that is asynchronous or consisting of remote communication. Research of such technologies has diminished in CSCW since the early 1990s, although Fischer et al. (2016) attempt to rekindle this, making the call for studies of ubiquitous technology use in new places. In this spirit, this thesis proffers the concept of the casual-social setting for further research, as one with which there is nascent understanding of newer technologies and their use within the work of socialising, and one which should be unpacked in future work.

Finally, this thesis identified how conversation in each setting was occasioned for related interactional projects using the same methods as part of the work of socialising together. The use of the devices was done so as part of the already *established social order* of the setting and regulated as such by those members who were co-present through "whatever organisation" (Sacks, 1992a, pp. 548–549) the world in which the devices inhabited. People were shown to make use of their devices—smartphones and smartspeakers—in and through the existing organisation of their lives; as Sacks (ibid.) argue, the devices "[were] made at home with the rest of world [in so much that the introduction of] each new [device] becomes the occasion for seeing again what we can see anywhere" (ibid., pp. 548–549). In other words, this thesis demonstrates how the use of these devices is brought into the already organised work of the casual-social setting as part of the already established order of socialising.

8.2.3.2 *Conversation with VUIs*

The second conceptual contribution offered by this thesis relates to VUI use and the question as to whether interaction with a VUI is indeed "conversation". Above in 8.2.3.1 the notion of the device use

being made at home was established through members bringing their devices into the work of the highly organised world. With the use of VUIs in particular, interaction was shown to consist of the phenomena of members 'talking to devices' through an ongoing conversation with others (see Chapters 5 and 6). The devices superficially 'talk' back to the user and thus at a cursory glance present the illusion that there is indeed an exchange or 'conversation' between human and machine. Adding to the narrative that such use is a conversation is the plethora of literature describing such interfaces as "conversational" (e.g. McTear et al. (2016)'s work on 'The Conversational Interface'). To examine this claim, this section brings together this thesis' findings to add a perspective grounded in empirical data to the debate, first unpacking the notion of 'talk to devices', and then that of 'talk by the devices' to reflect upon the concept of 'having a conversion with a VUI'.

On the use of the telephone, Sacks (1992a) explicates the characteristics of the opening of a telephone call and in doing so reveals that despite the introduction of a new technology, human interlocutors employed the existing methods, routines, and established social order to converse. Indeed, talk to VUIs has been shown to consist of specific characteristics in this thesis. Whereas, conversation typically unfolds with a minimisation of overlapping talk and gaps between speakers' turns (Sacks et al., 1974, pp. 704–706), the design of a VUI—that it listens from the utterance of a wake word through to a pause in talk—necessitates the production of gaps to delineate the completion of requests from 'other talk'[2]. In the study of a simulated VUI, Wooffitt (1994) similarly remarked upon the "comparatively lengthy silences between system turns" (ibid., p. 104). These system designs are predicated upon one-at-a-time interaction and a typical voice transcrip-

2 This echoes L. A. Suchman (2006)'s ideas of "shared understanding" between device and human, with both exhibiting different "respective views of the interaction" (ibid., pp. 123–124). In this case, a VUI has a different respective view, constrained by only 'understanding' that a request is completed by a drop in the relative ambient volume.

tion system cannot distinguish between multiple concurrent voices. As a result of this limitation, talk to VUIs consists of little overlap, yet is trailed by a gap which signifies the end of the request to the device[3]. In addition to this is the notable lapse in conversation both while a request is made and following a request until a response is produced by the VUI, or it is established by members that no response is forthcoming. Such actions are, in essence, demonstrable of others allowing for the user to 'get the device to work'. In this, the use of the VUI is shown to be a methodical accomplishment, done through the user talking to the device with the adoption of certain characteristics to get *a* desired output from the device

Now, consider the notion of how VUI interaction idealistically proceeds, i.e. through the input of a user speaking some request, and a synthesised voice delivers a response as output. Such an ideal easily affords the notion of referring to such interaction as conversational. There are two parts to this interaction with the VUI: the input and the output, and in combination they ostensibly map to a subset of formalised abstractions of talk (e.g. of a *command/action* or a *question/answer*). However, merely exhibiting a similarity to conversation in structure does not make it a conversation (Button and Sharrock, 1995).

With regard to the input to the VUI, the system itself has no comprehension of what a complete utterance is beyond a break in the stream of words and designates all that came before this break as a *request*. Equally, the device has no notion of social order or grounding of the context of the action that occasioned its use, instead merely operating on a series of words punctuated by a drop in the volume. As Button and Sharrock (ibid.) remark, the VUI would have to "have capacity to conduct the social actions constituent of particular activities" to understand the conversation, however the VUI reduces the situated action to that of a recording of an audio stream. The device then transcribes this captured audio to a series of words and processes it

3 Conversely, however, such a design feature was also shown to be exploitable to disrupt another's device use by talking over the request (see 6.3.3.1).

according to a series of pre-programmed rules[4], in turn abstracting away the situatedness of the action with which the input was produced. In this sense, the pipeline with which VUIs operate reduces the social action of producing input to a mere textual representation of that action devoid of its context[5]. This processing is based upon ideas of talk being formalizable to a series of preconfigured rules, whereas the basis of conversation analysis attests to its unformalizability (Sacks et al., 1974). With this in mind, it could be argued that for various reasons beyond the scope of this thesis, the design of the VUIs are *reductionist* in their treatment of input.

In terms of the response delivered by the VUI, there too are particulars that fracture the notion of the 'conversational' interface. It is, for benefit of the user in the 'here and now', an automated outcome of their actions (which were, as above, shown to be embedded among the established social order). There is no claim that a lamp engages in conversation when the switch is flicked by a person, for example. The light may come on, or it may not, and this outcome may depend upon any manner of technical or social reasons (e.g. the wrong switch may have been flicked). Yet, the seeming complexity, variability, and form through which the response from the VUI is made proffers the idea that it is a 'conversational' exchange. A rebuttal against this simplistic argument would be to say that response from a VUI differs from request to request for no 'obvious' reason (e.g. technical error, variation, software upgrades), it is mostly non-deterministic unlike the lamp and switch, and that it may be imbued with some context (e.g. information about current affairs, sports results, the state

4 This is evident in the way user manuals present to users the possible ways in which a VUI can respond. The VUI cannot deal with what it has not been programmed to deal with.

5 You could argue that a human can parse the context of a letter or an SMS and thus words do carry meaning even in written form. VUIs (at least of today) do not possess the capability to do this, perhaps for a combination of ethical, technical, and legal reasons.

of smart home technologies). Crucially, however, the response from a VUI is indiscriminate to the *context within which the response is delivered*.

Therefore, the non-determinism of the VUI's response comes not from locally-produced interaction with which the device finds itself, but rather from pre-designed features. It is, in this regard, ultimately not conversational as the response from the device adorns none of the qualities one would expect in conversation: the device does not manage its response among the social order, and even more so is 'ignorant' of the social action which occasioned the response in the first place. The work of 'attending' to this response is done through which the routine work of the members of the setting in which the device finds itself. As with the lamp analogy above, the person who turns on the light makes it naturally accountable, and with the VUI so too does the person who makes the VUI request. The device does not converse, but rather audibly simulates talk as a direct consequence of human action alone, it takes no account of how that simulation unfolds in the setting—it does not make its actions naturally accountable, nor does it bare such responsibility as part of the social order.

As such, it can be said that the audible response from the device provides a veneer of "conversation", but that such a notion is merely a simulation of conversation. The user talks to the device, but the device cannot converse. As Button and Sharrock (1995) remark:

> [Despite] the fact that one may be able to reproduce, on the computer, many sequences of conversation that formally resemble the sequences of conversation, which may indeed be formally indistinguishable from them, [it] does not demonstrate that one has thereby enabled a computer to converse in the way that human beings do.
>
> — Button and Sharrock (ibid., p. 111)

In this, what becomes evident is that it matters not if the device seemingly produces a response to a question, or follows an instruction with an action, but that the device itself is not conversing because it rests on the false basis of formalised pre-configured conversation,

rather than the notion of conversation as locally-produced situated action by the interlocutors.

In reflecting on this state, the lack of the VUI's 'competence' in engaging in conversation can be seen as what leads to members of each setting making the device 'at home'. Furthermore, through exploring the concept of 'VUI use as a conversation', the parallels to L. Suchman (1985)'s work on the use of an agent-based photocopier interface (see p. 46) become even more evident. Just as with the photocopier, VUIs possess only a limited sensitivity of the interaction that unfolds, and it is through this that the common sense basis upon which people approach interaction with the interface is revealed (i.e. the common sense basis in this case knowing how to do talk). In this regard, there is a mismatch between the very idea of a 'conversational interface' and the mundane competences implicated in doing conversation. The incompetence of systems to attend to the matters of conversation as humans do implicates the user of the VUI into resolving these misaligned competences. Through this, the user makes the VUI 'at home' through the methods they make any technology at home.

8.3 CRITICAL REFLECTION

The research questions posed in the thesis were developed out of personal intrigue into the sources of the consternation amongst popular press and academic literature of devices being used in face-to-face gatherings. This thesis drew upon three distinct fields of work: 1) Mobile HCI literature on creating and studying technologies for 'mobile collocated interactions' (see 2.2.2), 2) close studies of interaction from CSCW (see 2.3), and 3) the ethnomethodological approach to ethnography (see 3.2). This combination of fields provided the backdrop to the studies in this thesis, with the aim of each study being to examine and explicate the interactional projects of members using a device in a casual-social setting.

The three studies in this thesis were undertaken as independent pieces of work but successive pieces of work, with the second and third motivated by the previous one. In this regard, these studies are examinations of different technologies but with shared characteristics between them, i.e. the first two studies involve the use of portable devices such as smartphones, and the latter two use VUIs on a smartphone and a smartspeaker respectively. Perhaps unsurprisingly, the first two studies both feature the device being used for the same type of interactional project, i.e. to introduce new information to the conversation, and the latter two studies both identified members establishing the capability of the VUI through its use.

Although the critique of technology use in such settings provided the motivational backdrop for this thesis, the work itself speaks little to many of the arguments raised. In part, this is because much of this criticism relies upon *a posteriori* methods, and includes people's reflections upon device use, rather than an examination of their device use *in vivo*. This thesis does little to challenge specific critiques of device use (e.g. of isolation (Turkle, 2011)) because this thesis can only speak to the naturally accountable methods upon which members use a device in the setting and not members' perspectives or feelings regarding device use (unless, of course, they reported these as part of interaction in the studies). Nevertheless, through the approach adopted, what this thesis does accomplish is to show that people account for device use in casual-social settings, and embed it as a constituent activity of socialising together as a group. Device use was shown to be *regulated as an activity in socialising together*, as an ostensibly mundane feature of that interaction, and not in spite of it, and at least rubbing up against the critiques of isolation.

8.3.1 *Limitations of this work*

The interdisciplinary nature of this thesis has resulted in a number of challenges as well as some limitations with the work in this thesis. This section discusses two key limitations of this work:

1. The *limitations of a small n* (i.e. participant/'sample' size) in the empirical studies, and
2. The *discussion and linking of findings to design* are untested implications based on reflection of the empirical data.

LIMITATIONS OF A SMALL *n*

The study of people socialising and using devices was based on numerous small-scale studies of specific groups of participants. This thesis' adoption of an ethnomethodological approach, focused on the naturally accountable actions of members, produced thick descriptions of members' actions that would be recognisable to anyone with a vulgar competency in the setting's work (see 3.2.4). In this sense, this thesis did not rely on *interpretation* to construct 'scenic' descriptions of people interactions, but a praxeological account of member's interactional work. However, the unavoidable caveat is that these findings are not objective (nor could they be), and are not quantifiably generalisable to all situations—they are based upon ethnographic accounts of the studied groups of participants socialising in a given context and setting. Producing such a larger record of many cohorts or different settings would be an insurmountable task for a single thesis, and to do so would likely result in the dilution of the richness in which context is established as a factor in shaping interaction. It is the attention to the minutiae that furnishes the analysis with rich insight into the actions of people, yet also serves as a limitation of this work.

DISCUSSION AND LINKING OF FINDINGS TO DESIGN

This thesis explicated the collaborative efforts of members in using a device in conversation, and of how this turned upon the natural accountability of action, yet this thesis did not generate exhaustive implications for design. Furthermore, this thesis proposed that work in HCI to create collaborative experiences using mobile technologies could explore the use of VUIs in their design. Although others have remarked upon how it is a key activity for ethnographic studies to yield implications for design (e.g. Crabtree et al. (2012)), the focus on this work was addressing the literature gap of empirical data of people socialising and using technology. Nevertheless, implications arising from this underlying research in the empirical chapters have been published in the fields of HCI and CSCW, although these publications differ in parts from the analysis presented in this thesis[6]. However, neither the notion of collaboration with and around VUIs, or the resulting implications in the publications, are established or verified through an experiment. Instead, these notions were derived out reflection of the explicated machinery of interaction, identified through members' naturally accountable actions. Such a limitation is, of course, an avenue for future research and design to expand upon these findings, rather than a devaluation of the work in this thesis.

8.4 FUTURE WORK

This chapter has brought together and summarised this thesis' contributions and limitations. Looking forward, this thesis now ultimately concludes with how future research might proceed that builds upon these conclusions and addresses these limitations.

The naturalistic participant-observer approach used in the first two studies, and the automated selective data capture with the third, allowed for an analysis that oriented to understanding conversation

6 Each empirical chapter corresponds to a single publication: Porcheron et al. (2016a), Porcheron et al. (2017b) and Porcheron et al. (2018) respectively.

among members in the setting. However, this analysis was limited in that further work could be done on the data to orient to different matters of interaction, such as the specific nature of how talk to devices is modulated and how this varies over time, including factors such as 'recipient design' (i.e. how users adapt their voice to get the device to work (Clark, 1996)). In other words, the corpus of data, especially in relation to that of VUIs in the home, is rich and ripe for further analysis drawing upon methods such as Conversation Analysis. Additionally, future studies could orient to different matters of how talk to the devices is constructed, drawing on disciplines such as linguistics to provide greater insight into the language used, as some have already begun to argue for with regards to VUIs (e.g. Sutton et al. (2019)). These approaches would potentially generate additional findings in relation to what this thesis has offered in support of design tweaks, especially in the case of VUI design.

Another limitation is, as discussed, that the conduct observed is of specific cohorts of people, but that further configurations of cohorts and/or settings might reveal different findings. For example, friends socialising together in a pub is not the only combination where technology use is identified as problematic, with recent research in HCI examining notions of couples using mobile devices in bed (Salmela et al., 2019), and relating this to literature on using technology while collocated. With this, the foundations of this thesis hopefully reinforce and support future work in HCI to continually examine and pursue the idea of revealing technology in a range of settings and cohorts, all of which provides richer insights for research, and implications for the design of future technologies. The call for further examination of these settings in CSCW serves as a feasibility proposal to identify the ways in which VUI technologies could be embedded within the work of these—and other—settings.

Finally, the critical discussion of the methods employed by members in using a device identified the collaborative practices of members, and how these turn upon the accountable nature of interaction.

Such approaches steered clear of 'obvious' and redundant challenges such as 'better voice recognition' and instead attempted to explicate nuanced and provocative ideas for how to harness or disrupt existing methods, or to ameliorate the difficulty people had in 'getting devices to work'. The thesis proposed that designers and researchers who currently build and study CSCW systems for collaboration using portable devices could consider ways of using VUIs in future CSCW systems. However, further work could build on this idea to examine and determine its applicability to different settings following a *research through design* approach (Zimmerman et al., 2007), generating meaningful conclusions about just how such interactions unfold, and what implications these have for the design of future technologies.

In summary, the work in this thesis is not a comprehensive answer to what all device use entails in all settings, and does not offer a checklist of solutions for how technology could be redesigned to suit members' interactional needs. Nevertheless, it should serve as an empirical primer that can be utilised as a keystone to supporting further studies across disciplines that examine and design for everyday interaction with devices.

8.5 FINAL REMARKS

In conclusion, this thesis has delivered an empirical insight into for what purposes and how device use is interactionally organised as an activity in multi-party casual-social settings. This was to address a fundamental gap in a crowded body of literature on device use in everyday life, much of which focuses on specifically when we are collocated with others. Through the description, discussion and reflection upon members' actions in the settings, device use was shown to unfold through the routine of socialising together as an *embedded* activity, with members making their device use a naturally accountable phenomenon. Members also collaborated on their interactional projects for which device use was occasioned, with this collaboration

ostensibly turning upon the accountability of the use of the device. This thesis makes a number of contributions to the existing literature on the nature of device use in these settings, reflections of the methodological approach of the three studies in this thesis, and of conceptual insights that call for further work in HCI and CSCW.

Part IV

APPENDIX

A

ADDITIONAL INFORMATION ABOUT THE FIRST STUDY

This appendix includes additional material to the first study of mobile device use in a pub (see Chapter 4):

- Appendix A.1 provides the information sheet and consent form given to participants prior to the study,
- Appendix A.2 provides the post-observation interview questions, and
- Appendix A.3 provides the guidance on the study provided for the collaborative data session.

Additional material related to the study is available on the accompanying CD and online repository:

- The descriptive results from the questionnaire are provided in `studyone/questionnaire.pdf`.

A.1 INFORMATION SHEET AND CONSENT FORM

'Behaviour around Devices in Pubs' Study

Participant Information Sheet

Thank you for agreeing to take part in this study. The purpose of this research is to provide a greater insight into behaviours when we are socialising with friends in a relaxed environment.

The study involves the observation of a group of participants in a pub for between 60 and 90 minutes (it may be shorter). As a participant, you are free to do as you please for the duration of the study.

You will be guaranteed complete anonymity, with all recorded data from your participation identified by number only. Video recordings will be used for research purposes and will only be accessible by researchers working on this study, unless you optionally choose to give explicit consent for use in academic publications and presentations.

Compensation for time and involvement in the study will be in the form of a shopping voucher, no other form of reimbursement will be provided. In order to receive the reimbursement, you must remain involved in the study until completion, as directed by the researcher.

Towards the end of the study, a small questionnaire must be completed, and a semi-structured group interview will take place. After this, the data capture will be ended. We may also want to ask you further questions, via telephone, after the study – you are free to opt-out of this. All of the data collected will be stored in accordance with the Data Protection Act 1998 and will only be accessible by those directly involved in the research.

You must be at least 18 years old to participate in the study. If you wish to withdraw your consent during or after the study, you have the right to do so without consequence. Any captured data, and analysis performed on the data, will be destroyed following your withdrawal. Please speak to the researcher in person, or use the contact details below.

Please do not hesitate to ask any questions. You may contact me at any time for information about the research or in relation to your consent, including withdrawing from the study:

Address	Martin Porcheron, c/o Joel Fischer, Mixed Reality Lab, Computer Science, The University of Nottingham, Wollaton Road, Nottingham, NG8 1BB
Email	martin.porcheron@nottingham.ac.uk
Telephone	**[REDACTED FROM THESIS]**

horizon CENTRE FOR DOCTORAL TRAINING

mixed reality lab

The University of Nottingham UNITED KINGDOM · CHINA · MALAYSIA

'Behaviour around Devices in Pubs' Study

Participant Consent Form

If you have read the information sheet and agree to the three statements below, please complete this form.

- I am at least 18 years old and agree to participate in the study, as described in the information sheet; including the observation, recording, questionnaire and interview.
- I understand that during the study I will be recorded and notes will be made, and because of this it is important to act naturally and in a manner I would typically behave.
- I agree to the following types of data being captured (please initial all that apply):
  Video ..... Audio ..... Observations .....
- I understand that I may withdraw consent from the study at any time, without reason, by contacting the researcher using the provided details, and that all my data will be erased once I do so.

Signature ..... Date ..... / ..... / 2015

Full Name .....

Email .....

Telephone .....

Additionally, I give the following permissions (each of these are optional and do not impact your involvement in the study). Please initial on the dotted line if you agree to give the permission.

I give permission for the data collected to be used in subsequent associated research. .....

I give permission for captured images to be used in presentations and publications. .....

I give permission for anonymised quotes to be used in academic publications. .....

I am willing to participate in a follow-up telephone interview .....

Researcher ..... Date ..... / ..... / 2015

horizon CENTRE FOR DOCTORAL TRAINING

mixed reality lab

The University of Nottingham UNITED KINGDOM · CHINA · MALAYSIA

A.2 EXIT INTERVIEW QUESTIONS

- How do you feel about the presence of phones in social settings?
- Would you say the presence of devices has an impact on the conversation?
- Do you recall a time when you have felt ignored by someone using their phone?
- How feel about devices attempting to stop you using them at inopportune moments?
- Do you see any merits in devices attempting to restrict usage?
- Do you feel that people have a responsibility to the group dynamic?
- Would you categorise mobiles as a support tool or a distraction device
- Could you see mobile phones being used in conversation without detracting from it?
- Would you be willing to take part in further studies that required the installation of an app?
- Were you disturbed by presence of cameras or recording equipment?
- Did you feel you acted unnaturally due to nature of study?
- Did you remain aware of the presence of camera?

A.3 DATA SESSION GUIDANCE

SETTING

The recordings for this data session were made in a Nottingham-based pub, near to the university, during normal opening hours. The

majority of the recordings took place in the afternoon when the pub was relatively quiet.

A table that was in the corner of pub was chosen for the groups to sit at, and two GoPro cameras were positioned to capture all those present at the table. One camera was positioned on a tripod and the other was positioned on a ledge within the pub. An audio recorder was placed on the table to capture higher-quality audio.

Groups of 3 or 4 friends were recruited through email and word-of-mouth communication to participate in a study that involved "going to the pub". There were no prerequisites other than that all members of the group should be friends; groups were purposefully told minimal information before the study, other that what was required by the ethics committee.

Upon arrival at the pub, groups were greeted and invited to sort drinks out before taking their seats. Consent forms were completed, and an opportunity for individuals to ask questions was provided. The common question amongst groups was whether any tasks were required and this was answered accordingly.

The researcher sat at the table and engaged with the group where appropriate.

FOCUS

The focus of this research is to discover the interactional methods through which devices are topicalised, then sustained/co-oriented to, and then disengaged from, within social collocated interactions.

B

TRANSCRIPT NOTATION

In general, the orthographic notation for fragments of transcribed data is based upon the system employed by Heath et al. (2010). However, a number of differences and simplifications have been made for brevity, clarity, and conciseness. This notation itself was derived from the notation system originally described by Atkinson and Heritage (1984), devised by Gail Jefferson, and often used in Conversation Analysis-related literature.

Summarily, the notation used for transcribed data in this thesis adopts the following conventions:

- the volume of talk is denoted as `LOUD` or `°quiet°`,
- emphasis is denoted with `underlined text`,
- shifts in intonation are given as arrows, i.e. ↑ for a rising intonation and ↓ for a falling intonation,
- a single dash (`-`) when an utterance is cut off,
- an equals (`=`) at the end of an utterance and at the start of a following utterance to denote contiguous talk (indentation is used to improve clarity and readability in these cases),
- elongation of sounds and words are like `th::is`, where the `th` sound is two-tenths of a second in length,
- pauses between words and utterances are given as `(.)`, where each individual period represents a tenth of a second; or as `(0.4)`, where this represents a pause of 0.4 seconds,
- overlapping talk or action is denoted using opening square brackets (`[`) and closing square brackets (`]`) where possible or applic-

able (sometimes this closing bracket is omitted if two concurrent utterances end simultaneously at an end of a turn),

- indentation is often used with overlapping talk or action to aid readability,
- actions are given in `((double parentheses))`,
- utterances to an electronic device as a query (i.e. as input to the device) are given as **`bold text`**,
- digitally produced spoken words from an electronic device are preceded and succeeded with two forward slashes and typed in `// italics //` , and
- Names are typically denoted using the first three letters from the first name of the participant, e.g. `LIL` for *Lily*; the researcher is identified as `RES`.

C

FRAGMENTS FROM THE FIRST STUDY

This appendix includes the full fragments presented over a number of excerpts in Chapter 4. There are three fragments:

- Appendix C.1 is the transcript for the *Miniature Schnauzers* fragment, which is composed of Data Excerpts 4.1, 4.2, 4.3, and 4.4;
- Appendix C.2 is the transcript for the *Font Size* fragment, which is composed of Data Excerpts 4.5, 4.6, and 4.7; and
- Appendix C.3 is the transcript for the *Shorthand* fragment, which is composed of Data Excerpts 4.8, 4.9, and 4.10.

C.1 MINIATURE SCHNAUZERS

01		CAL	i like miniature schnauzers
02		DAY	°how big are schn-?°
03		←CAL	it's like (.) like (.) they're
04			so: cute
05		←CAL	((briefly looks at her bag to
06			her left before looking back))
07		DAY	i like big dogs
08		←CAL	i know, but google schnauzer,
09			right?
10		DAY	((gets phone out from bag))
11		CAL	((leans towards DAY))
12			the puppies (.) schnauzer

puppies are gorgeous
(...)
CAL so it's miniature schnauzer
DAY how do you?=


←CAL =erm::
DAY (sccchhhh) (tea) (ee) (ar)
CAL oh schnauzer (.)
it's s-c-h-n-a-u-z- n-a-u-
(2.2) schnauzer
DAY oh, schnauzer


←CAL schnauzer, go look at
schnauzer puppies right↑
((continues to look at phone))
DAY °my internet is rubbish so
this may take some time°
DAY ((looks down and unlocks
phone)) ↓oh that: thing


←CAL ((leans towards DAY and
shifts gaze towards her
screen)) yes look at them
oo:::↑
DAY °schanuzer°

C.2 FONT SIZE

←JAY beginning of september they
had their (.) all their
christmas stuff out (.) and I
was °like oh my god nobody
( )°


←LAW °jesus!°
JAY we just booked ours (1.0) we
do me and liam and james and
malcolm do (one every year and
we) just booked it
MAL du bois↑
LAW =sorry (.) have you (.) um (.)
((jovially)) jonathan has sent
round an email (.) this is
great for your study isn't it?

←RES going to have to zoom in for
the camera (.) it's only set
to 720p!
JAY mu:::::a::h
LAW yeah (.) that's (.) that (.)
that's the email!-
(...)
MAL is that him or is that your
phone fitting the line in?

C.3 SHORTHAND

←LAW isn't it mainly phonetic?
JAY it's like:
(3.2)

←JAY there's various versions so
the one she tried to teach me
first so i could start going
is missing out all the vowels
LAW ((briefly looks at JAY while
picking up his phone, he then
(begins to use his phone once
he has it in his hands))
LAW ↓yeah

←JAY and once you get good at that
you just write a lot quicker
(0.7) but then she had one
which was literally like (.)=
LAW ↓yeah
JAY =swiggles and just didn't look
like anything and i don't know
if that's phonetic or::::
MAL ( )

←LAW hang on! ((typing on phone
with thumbs)) schuh::::::::ort
(.....) hand (..)
my mum's regular handwriting
RES i know some people who miss
out vowels (.) like the e

28		←JAY	that's how i do it (.) missing
29			out vowels is very very good
30			but there's a squiggly one i
31			don't understand
32		MAL	this is why i didn't do
33			ethnography (1.8) just get the
34			participants to fill
35			everything out
36		LAW	i didn't do ethnography either
37			either!
38		MAL	yeah↑ you do it- i'm not- i'm
39			not doing=

40		←LAW	=oo↑ that's got (2.0) that's
41			cinnamon in it or something
42			something (..) smells amazing

D

ADDITIONAL INFORMATION ABOUT THE SECOND STUDY

This appendix includes additional material to the second study of Voice User Interface (VUI) use in a café (see Chapter 5):

- Appendix D.1 provides the information sheet and consent form given to participants prior to the study,
- Appendix D.2 provides the post-observation interview questions, and
- Appendix D.3 provides the post-observation questionnaire.

Additional material related to the study is available on the accompanying CD and online repository:

- The descriptive results from the questionnaire are provided in `studytwo/questionnaire.pdf`.

D.1 INFORMATION SHEET AND CONSENT FORM

MOBILE DEVICE USE IN SOCIAL SETTINGS STUDY
Participant Information Sheet

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Thank you for agreeing to take part in this study. The purpose of this research is to provide a greater insight into how we interact with each other and the technology around us. This study consists of an observation of a group of participants for up to 90 minutes (it may be shorter), the main purpose is to observe you engaging in a typical discussion with friends. The only formal requirement is that where possible you make use of automatic speech recognition systems on your mobile device, where possible (e.g. Google Now, Siri, Cortana), and use a dictation system, if present, instead of typing. Towards the end of the study you will be asked to complete a questionnaire and participate in a semi-structured group interview, after which the data capture will be ended. Compensation for time and involvement in the study will be a shopping voucher.

Please note that:

- You must be at least 18 years old to participate in the study.
- If you wish to withdraw your consent during or after the study, you have the right to do so without consequence. Any identifiable data will be destroyed following your withdrawal. Please speak to the researcher in person, or use the contact details below.
- You will be guaranteed complete anonymity, with all recorded data from your participation identified by number or pseudonym only. Video will be used for research purposes and will only be accessible by researchers working on this research, unless you optionally choose to give explicit consent for use in academic publications and presentations.
- All of the data collected will be stored in accordance with the Data Protection Act 1998 (DPA) and will be accessible by those involved in this research. The DPA requires all data to be stored on a secure password-protected drive in a secure facility, and kept only as long as is needed.
- Compiled data will be managed in accordance with both University regulations and funding body (EPSRC) guidelines on research data management. This means that research data must be retained and deposited in an accessible repository.

Please do not hesitate to ask any questions. You may contact the researcher at any time for information about the research or in relation to your consent, including withdrawal from the study:

Researcher Martin Porcheron
Email martin.porcheron@nottingham.ac.uk
Address Computer Science, The University of Nottingham, Wollaton Road, Nottingham, NG8 1BB

MOBILE DEVICE USE IN SOCIAL SETTINGS STUDY

Participant Consent Sheet

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

If you have read the information sheet and agree to participate, please complete this form. You are reminded that:

- You must be at least 18 years old and agree to participate in the study, as described in the information sheet; including the observation, audio/video recording, and interview.
- Collected data will be stored and managed in accordance with the Data Protection Act, University regulations, and funding body guidelines on research data management.
- You may withdraw consent at any time, without reason, by contacting the researcher using the provided details, and that all my data will be erased once I do so.

I agree to the following types of recording of:

Video ..... (initial) **Audio** ..... (initial)

I give permission for captured imagery to be used in presentations and publications ..... (initial)

Please complete the details below to consent to the study:

Signature ..... **Date** .....

Full Name .....

Email .....

Researcher ..... **Date** .....

D.2 EXIT INTERVIEW QUESTIONS

- How do you feel about the presence of phones in social settings?
- Would you say the presence of devices has an impact on the conversation?
- Do you recall a time when you have felt ignored by someone using their phone?
- How feel about devices attempting to stop you using them at inopportune moments?
- Do you see any merits in devices attempting to restrict usage?
- Do you feel that people have a responsibility to the group dynamic?
- Would you categorise mobiles as a support tool or a distraction device?
- Could you see mobile phones being used in conversation without detracting from it?
- Would you be willing to take part in further studies that required the installation of an app?
- Were you disturbed by presence of cameras or recording equipment?
- Did you feel you acted unnaturally due to nature of study?
- Did you remain aware of the presence of cameras?

D.3 QUESTIONNAIRE

MOBILE DEVICE USE IN SOCIAL SETTINGS STUDY
Questionnaire

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Age ..... **Gender** ..... **Ethnicity/Nationality** .....

Please state how much you agree/disagree with the following statements, in relation to the study.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
"I felt comfortable during the study"	☐	☐	☐	☐	☐
"I behaved as normally as I would when socialising with friends"	☐	☐	☐	☐	☐
"I would have acted differently if I wasn't being video recorded"	☐	☐	☐	☐	☐
"I was consciously aware of my behaviour more so than normal"	☐	☐	☐	☐	☐

Please complete the following two statements using your own personal judgement.

	Very Infrequently	Infrequently	Sometimes	Often	Very Often
"I used my mobile device..."	☐	☐	☐	☐	☐
"My friends used their mobile devices..."	☐	☐	☐	☐	☐

These questions are about mobile devices you own and would consider using when socialising with friends.

	Smart phone	Tablet	Smart watch	Fitness band	Music Player
Which mobile devices do you own?	☐	☐	☐	☐	☐
Of the devices you own, which do you have with you now?	☐	☐	☐	☐	☐
Of the devices you own, which did you use during the study?	☐	☐	☐	☐	☐
Of the devices you own, which would you consider using when socialising with friends?	☐	☐	☐	☐	☐

Please answer the following questions by stating how much you agree/disagree with the statement. If a question relates to a device you do not own, choose *N/A*.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	N/A
"I feel comfortable looking at my device for a short period of time when while socialising"	☐	☐	☐	☐	☐	☐
"I would explain to my friends why I am using my device"	☐	☐	☐	☐	☐	☐
"I am OK with a friend using a mobile device whilst socialising with me"	☐	☐	☐	☐	☐	☐
"I use my phone more than the other devices I own"	☐	☐	☐	☐	☐	☐
"I feel that I use my mobile device as much as others when in social situations"	☐	☐	☐	☐	☐	☐

E

FRAGMENTS FROM THE SECOND STUDY

This appendix includes the full fragments presented over a number of excerpts in Chapter 5. There are three fragments:

- Appendix E.1 is the transcript for the *When Does the Sun Go Down?* fragment, which is composed of Data Excerpts 5.1, 5.2, and 5.3;
- Appendix E.2 is the transcript for the *Do Animals Have Accents?* fragment, which is composed of Data Excerpts 5.4, 5.5, 5.6, and 5.7; and
- Appendix E.3 is the transcript for the *Hey Siri! . . . Call My Mother* fragment, which is composed of Data Excerpts 5.8 and 5.9.

E.1 WHEN DOES THE SUN GO DOWN?

```
01  HAR   i'll be fine in like three minutes ((holds hands in front of
02        eyes))
03  RES   keeps coming back as well like
04  SAL   as soon as you change it comes back
05  JUL   yeah yeaha
06      (0.3)
07  RES   there's actually just someone out there with a light!
08  ALL   ((laugh))
        (...)
16  JUL   [ ((removes cover from device but leaves open)) ]
17  SAL   ((laughs))
18  JUL   ((presses button on device))
19  HAR   there we go!
20  JUL   what's the time of sunset?
21      (1.3)
22  ALL   ((gaze at the tablet))
23      (3.0)
```

```
24  JUL   ok! // ((device displays clock)) //
25  ART   ((leans in to look))
26  SAL   that's [ a      ] fucking analogue clock it pisses me off!
27  HAR          [ today? ]
28  HAR   ilunno (0.6) 24 hour=
29  JUL   <no no no!> it misunderstood actually (0.8) understood what's
30        the [ time ]
31  HAR    [ time ] now
32  JUL   so-
33  ART   soaoah yeah↑
34  JUL   shall i ask (1.6) um:=
35  HAR                       =what time will the [ sun set? ]
36  JUL                                           [ ((holds button)) ]
37  JUL   // ((audible chime)) //
38      (4.0)
39  JUL    // ((on screen text: go ahead i'm listening...)) //
40      (0.3)
41  JUL   when does the sun go down?
42      (2.9)
43  JUL   sunset will be at [ seventeen thirty two ]
44  ART                     [ ther:::e you go      ]
```

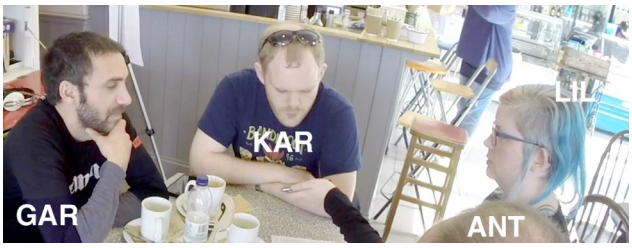
(line 1)

(line 25)

E.2 DO ANIMALS HAVE ACCENTS?

01 KAR do cats acth- (0.5) can you work out whether it's french because
02 because its talking in a- doing a french cat impression
03 LIL i::::: think some animals you can
04 (1.9)
05 LIL ((picks up phone from table))
(...)
40 LIL er:::m: ((holding phone in front of her at chest level))
41 (3.7)
42 LIL ((moves phone up to face)) **do animals have accents?**
43 (2.1)
44 GAR ((shifts gaze to LIL))
45 yes they do actually! i think i've read something
46 LIL i think i have [too↓]
47 GAR [yeas!] [(0.6) cows! i- i]
48 read about cows that they have
49 different accents around the world
50 KAR [you missed mine- my racist joke]
51 LIL **DO: ANIMALS HAVE ACCENTS!**
52 (2.4)
53 LIL °rubbish°=
54 KAR =parrots presumably do=
55 LIL =can you ask it?
56 ((holds phone out in front of KAR's face))
57 RES ((retrieves phone out of pocket))
58 KAR **DO: ANIMALS HAVE ACC::ENTS!**
59 (0.9)
60 LIL no:!
61 RES *// sorry i'm- //*
62 RES ((RES touches screen to stop utterance))
63 RES **do animals have accents?**
64 LIL **do: animals have accents?**
65 RES *// ok i've found this on the web //* (sigh)
66 GAR do [they?]
67 LIL [ah (.)] it's working now!

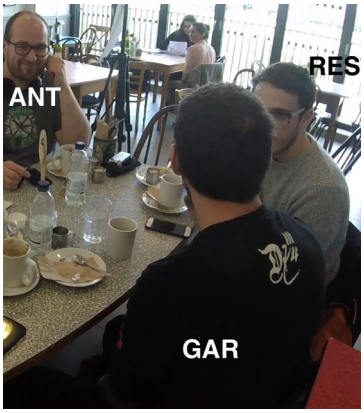
(line 44)



(line 56)

E.3 HEY SIRI! . . . CALL MY MOTHER

01	GAR	i'm curious if I say in
02		romanian (.) to call my mother
03		(0.7)
04	GAR	it will actually find the
05		contact for my mother is (.)
06		mama in romanian (.) if I say
07		call my mum will it actually
08		call my mother which is in a
09		contact as mama (0.7) will it
10		make the connection between
11		mama and mum
12	RES	cos you can also tell people
13		who they (.) like you can say
14		like



GAR **hey siri=**
RES =my mother is this person (0.8)
GAR ((glances down at screen)) ((moves device in front of mouth)) **hey siri**
GAR ((moves device to chest height between the two))
(1.0)
RES i'd press the button
(1.2)
GAR ((moves device in front of mouth)) **hey siri**
GAR ((moves device to chest height between the two))
(2.4)
GAR ((moves device in front of mouth)) **call my mother**
← ((GAR and RES look at screen))

(5.9)
// what is your mother's name? //
RES ((points towards screen)) yeah but then
(0.9)
GAR **my mother is mama**
GAR *// i can't find anyone called mamma //*

F

ADDITIONAL INFORMATION ABOUT THE THIRD STUDY

This appendix includes additional material to the third study of Voice User Interface (VUI) use in a home (see Chapter 6):

- Appendix F.1 provides the information sheet and consent form given to participants prior to the study, and
- Appendix F.2 provides the Amazon Echo Help Guide produced for the study, and left with participant households.

F.1 INFORMATION SHEET AND CONSENT FORM

PERSONAL ASSISTANTS IN THE HOME STUDY
Household Information Sheet

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Thank you for agreeing to take part in this study. This study is interested in how the use of personal assistants, such as Siri or Cortana, could be used in a home. We will give you a self-contained personal assistant to use in your home and help you set it up as well as showing you how it is used. The personal assistant will listen for instructions and respond accordingly by talking back to you.

The personal assistant will store the audio around (i.e. 1 minute before and after) the interaction with the device in order for us to understand what lead to it being used as well as the specific use and your respond to the personal assistant. We will use this information purely to understand how the assistant is used in everyday life. You can request that any particular query and the surrounding captured audio to be deleted at any time. If you have any technical problems with the personal assistant at any point you should contact the researcher.

At the end of the study we will collect the device and ask that you take part in a semi-structured interview to reflect upon interacting with the personal assistant in the home.

Please note that:

- You must be at least 18 years old and a permanent resident of the home to provide the consent to participate in the study. In completing the consent form, you provide consent on behalf of all adults in the household and assent for the study to take place.
- You must be a legal guardian of any children or minors in the home, and in consenting to participant in this study, you are providing consent for them to participate in the study too.
- If you wish to withdraw your consent during or after the study, you have the right to do so without consequence. Any identifiable data will be destroyed following your withdrawal. Please speak to the researcher in person, or use the contact details below.
- All recorded data from your participation identified by number or pseudonym only. Audio will be used for research purposes and will only be accessible by researchers working on this research, unless you optionally choose to give explicit consent for use in academic publications and presentations.
- All of the data collected will be stored in accordance with the Data Protection Act 1998 (DPA) and will be accessible by those involved in this research. The DPA requires all data to be stored on a secure password-protected drive in a secure facility, and kept only as long as is needed.
- Compiled data will be managed in accordance with both University regulations and funding body (EPSRC) guidelines on research data management. This means that research data must be retained and transcripts will be deposited in an accessible repository.

Please do not hesitate to ask any questions. You may contact the researcher at any time for information about the research or in relation to your consent, including withdrawal from the study:

Researcher Martin Porcheron
Email martin.porcheron@nottingham.ac.uk
Address Computer Science, The University of Nottingham, Wollaton Road, Nottingham, NG8 1BB

PERSONAL ASSISTANTS IN THE HOME STUDY
Household Consent Sheet

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

If you have read the information sheet and agree to participate, please complete this form. You are reminded that:

- You are consenting on behalf of all adults in the household as a representative of the household. Furthermore, you are the legal guardian of any children in the house and provide consent on their behalf.
- You must be at least 18 years old and agree to participate in the study, as described in the information sheet; including the audio capture and interview.
- Your data will be stored and managed in accordance with the Data Protection Act 1998, University regulations, and funding body guidelines on research data management.
- You can request exclusion of particular data from analysis. This data will be deleted.
- You may withdraw consent at any time, without reason, by contacting the researcher using the provided details, and that all my data will be erased once I do so.
- You must report any problems with the device to the researcher.

Please complete the details below to consent to the study:

Address .....

.....

..... **Postcode** .....

Please initial below to agree to the following statements:

I give permission for the personal assistant to be installed in my home, and for audio data captured to be used in research. ..... (initial)

I know I am consenting on behalf of all members of the household, including those under 18 years of age. ..... (initial)

I agree to take part in a follow-up interview after the personal assistant has been collected by the researcher. ..... (initial)

Optionally, I give permission for captured audio to be used in presentations and publications. ..... (initial)

Please complete the details below to consent to the study:

Signature ..... **Date** .....

Full Name .....

Email .....

Telephone .....

Researcher ..... **Date** .....

F.2 AMAZON ECHO HELP GUIDE

PERSONAL ASSISTANTS IN THE HOME STUDY
Amazon Echo Help Guide (v2)

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

The Amazon Echo is registered to a single Amazon account, and must be set using an Android or iOS (iPhone or iPad) app. This app can be used to configure an Echo that has been setup, and used to install additional commands (called *Skills*) for you to use. You can also manage your Amazon Echo from alexa.amazon.com by logging in with the Amazon account used to setup the Echo.

EXAMPLE QUERIES

By default, Amazon Echo is triggered using the word "*Alexa*" followed by your query. **For this study, we ask that you don't change 'hotword' Echo listens for**, as this is what the conditional voice recorder (CVR) listens for. The Alexa app provides a history of queries and responses that have been asked to Amazon Echo. Alexa responds to the following queries:

- You can ask questions, such as "why is the sky blue?", or "what's the definition of…?" , or "who is the lead singer of [band]?"
- You can ask for Wikipedia summaries by saying "Wikipedia" followed by the topic
- You can play radio stations by saying "play radio [station name]"
- You can find local businesses and restaurants nearby by asking for them, or even the opening hours for local businesses
- You can ask for news and sports results
- You can set alarms and times

ADDING MULTIPLE AMAZON ACCOUNTS

It is possible to create an "Amazon Household", consisting of two or more Amazon accounts. With multiple profiles added to an Amazon Echo you can listen to another account's content (e.g. music and audiobooks), and manage shared features such as to-do lists. To enable this:

1. Visit alexa.amazon.com on a computer, and login with the account used to setup Amazon Echo.
2. On the left-hand side, select the *Settings* option, and then *Household Profile*
3. Login with the second Amazon account to link the profiles

To switch profiles on Echo, say "switch accounts" — you can also ask which profile is currently active by asking "which profile is this?".

CONFIGURABLE OPTIONS

The Alexa app provides a number of configurable options for Amazon Echo. Note, however, that once set up, you must always use the same instance of the app on the same phone or tablet unless you have configured multiple profiles (see above). The app allows you to configure:

- Accounts for online streaming music (Amazon Prime Music, Spotify and TuneIn Radio) and eBook (Audible) services. If you have Amazon Prime, you can play music from the Amazon Prime Music Library by simply asking for a song, artist, or genre
- To add, edit, and remove items from a to-do list and a shopping list
- To view a list of timers and alarms created by talking to your Amazon Echo
- Configuration options for smart home devices (e.g. Philips Hue light bulbs)

- Settings to configure the Amazon Echo
    - You can set the location of the Amazon Echo, allowing you to ask for local weather information or other location-dependant information (e.g. restaurants)
    - You can choose which news publication provides the "flash briefing" news updates on Amazon Echo
    - You can link Amazon Echo to your Google Calendar to ask questions about your schedule

ADDITIONAL SKILLS

Amazon Echo can be configured to listen and respond to additional commands, called "Skills". A large library of skills is available, and can be explored either through the Alexa app, or through the website alexa.amazon.com.

USING THE CONDITIONAL VOICE RECORDER (CVR)

The CVR listens for when you use the word "Alexa" and then saves the last minute of audio, and starts recording for one further minute. If you say "Alexa" again in that minute, the recording time is extended. When plugged in, the box takes roughly one minute to activate and during this time a feint blue light on the box shows. Once the system is active, the blue light becomes brighter.

KEY INFORMATION ABOUT THE CVR

- **When the blue light is on, the box is listening for the word "Alexa".** It keeps a copy of the last one minute of sound in memory if the blue light is on, but this is never saved to disk unless it believes it hears the word "Alexa**".**
- **When the red and blue lights are on, the CVR is recording**, the red light turns off once recording is completed.
- **To disable the CVR from listening, press the button** — the blue light will go out. If the box is currently recording, this recording is never saved, and the red light goes out too. Pressing the button again turns the CVR back on. The CVR is less effective at detecting voice from a distance than Echo. Therefore, for the study, we ask that you mostly try to use the Echo when close to it.
- **If there are no lights visible**, try pressing the button once and waiting. If lights still do not show, restart the CVR by unplugging the USB cable from the plug, and plugging it back in.
- **If the red and blue lights are flashing**, the system has crashed. Restart the CVR by unplugging the USB cable from the plug, and plugging it back in.
- **If the box does not respond to button presses**, the system may has crashed. Restart the CVR by unplugging the USB cable from the plug, and plugging it back in.

STUDY HELP

For help, concerns, questions, or queries, please contact the researcher, Martin Porcheron, by email at **martin.porcheron@nottingham.ac.uk** or by calling **[REDACTED FROM THESIS]**.

G

FRAGMENTS FROM THE THIRD STUDY

This appendix includes the full fragments presented over a number of excerpts in Chapter 6. There are three fragments:

- Appendix G.1 is the transcript for the *Where is Greece?* fragment, which is composed of Data Excerpts 6.1, 6.2, and 6.3;
- Appendix G.2 is the transcript for the *New Year's Music* fragment, which is composed of Data Excerpts 6.4, 6.5, and 6.6; and
- Appendix G.3 is the transcript for the *Beat the Intro* fragment, which is composed of Data Excerpts 6.7, 6.8, and 6.9.

G.1 WHERE IS GREECE?

```
LEA    alexa (.) where is greece
   (2.0)
ALE    // greece is a un-recognised country in the northern hemisphere
       (.) it shares a border with turkey, albania, bulgaria
       and macedonia= //
ISA                =[ that's it ]
LEA                =[ that's it ]
   (1.1)
JOH    alexa where is amfissa
   (2.0)
ALE    // amfissa is a city in phocidos (..) greece (.) it is 82 miles
       133 kilometres west of athens and 26 miles 42 kilometres south
       of lamia //
   (0.5)
JOH    alexa where is (0.3) delph-ee
   (7.0)
ALE    // delph-i is a village in carroll county indiana (.) indiana (.)
       it is 62 miles 99 kilometres north of Indianapolis and 87 miles
       140 kilometres= //
NIK                  =alexa stop
```

G.2 NEW YEAR'S MUSIC

01 NIK **alexa**
02 (2.6)
03 ISA **play some new year's music**
04 (1.8)
05 ALE *// here's a station for jazz music (.) instrumental jazz //*
06 *((begins playing jazz music))*
07 ISA **alexa this is not what we wanted**
08 [((laughs))]
09 NIK [(1.2) **alexa** (1.1) **shut**] **up!**
10 ISA hey::↑(.) **alexa nikos apologises for being so rude**
11 (0.3)
12 ALE hi there
13 (1.0)
14 [*((resumes playing jazz music))*]
15 NIK [(2.4) **alexa stop**] **stop!**

G.3 BEAT THE INTRO

01 SUS i'd like to play beat the intro in a minute
02 LIA [oh no::]
03 SUS [**alexa**][(1.1)] **beat the in**[**tro**
04 CAR [°yeah°]
05 LIA [°no:::...°
06 (0.6)
07 CAR it's mother's day?
08 (0.4)
09 SUS it's ( ) yep (.) listen (.) you need to keep on eating your
10 orange stuff (.) liam
11 (0.7)
12 CAR and your green stuff
13 SUS **alexa** (1.3) **alexa** (0.5)=
14 CAR =°and your brown stuff°
15 SUS **play beat the intro**
16 EMM °and the yellow stuff?°
17 LIA °and the meat stuff°
18 (0.9)
19 ALE *// resuming the music //*
20 EMM ((laughs))
21 ALE *((music plays))*
22 SUS oh no::!

EMM ((laughs))
CAR **alexa stop:**
ALE *((stops playback))*
LIA no::::::::::::
CAR **ale**[**xa** (1.0)] **bea:t: the (.) intro**
SUS [((laughs))]
SUS it does it for you
(5.0)
EMM nope (.) she didn like tha:::::t
EMM **alexsa** [(1.0)] **play beat the intro::**
CAR [is it called beat the intro?]
(2.1)
ALE *// you want to hear a station for b b intro [(0.5)] right? //*
EMM [°**no:**°]
(1.1)
EMM **no:** (.) **i don't alex:a** (0.5) **no!**
(1.3)
ALE *// alrig↑ht //*
(0.7)
CAR we played it the other ni:ght! the game we played
the [other night ((laughs))]
SUS [yeaherr:: **alexa**] **skills** (.) **beat the intro**
(4.5)
SUS °uh::↓:°
EMM she didn like tha:↓:t
SUS **alechSA::::::**

BIBLIOGRAPHY

Abowd, G. D. (1999). 'Classroom 2000: An experiment with the instrumentation of a living educational environment'. In: *IBM Systems Journal* 38.4, pp. 508–530. ISSN: 0018-8670. DOI: 10.1147/sj.384.0508 (cit. on p. 25).

Adobe Inc. (2018). *ADI – State of Voice Assistants*. URL: https://www.slideshare.net/adobe/adi-state-of-voice-assistants-113779956 (visited on 03/03/2019) (cit. on p. 16).

Ames, M. G. (Feb. 2013). 'Managing Mobile Multitasking: The Culture of iPhones on Stanford Campus'. In: *Proceedings of the 2013 conference on Computer Supported Cooperative Work*. CSCW '13. New York, NY, USA: ACM, p. 1487. ISBN: 9781450313315. DOI: 10.1145/2441776.2441945 (cit. on pp. 18–20, 22).

Ammari, T., Kaye, J., Tsai, J. Y. and Bentley, F. (2019). 'Music, search and IoT: How people (really) use voice assistants'. In: *ACM Transactions on Computer-Human Interaction*. TOCHI 26.1. DOI: 10.1145/3311956 (cit. on pp. 119, 172).

Anderson, B. (1997). 'Work, Ethnography and System Design'. In: *The Encyclopedia of Microcomputers* 20, pp. 159–183 (cit. on p. 41).

Ask Your Target Market (2016). *AI Personal Assistants Survey*. Tech. rep. URL: https://aytm.com/blog/market-pulse-research/ai-personal-assistants-survey/ (cit. on p. 7).

Atkinson, J. M. and Heritage, J. (1984). 'Transcript Notation'. In: *Structures of Social Action: Studies in Conversation Analysis*. Cambridge University Press, pp. ix–xvi. DOI: 10.1017/CBO9780511665868 (cit. on p. 214).

Bannon, L., Bowers, J., Carstensen, P., Hughes, J. A., Kuutii, K., Pycock, J., Rodden, T., Schmidt, K., Shapiro, D., Sharrock, W. and Viller, S. (Jan. 1993). 'Informing CSCW System Requirements'. In: *COMIC Deliverable 2.1*. Lancaster, UK: The COMIC Project (cit. on pp. 31, 126).

Bell, G. and Dourish, P. (2007). 'Yesterday's tomorrows: Notes on ubiquitous computing's dominant vision'. In: *Personal and Ubiquitous Computing* 11.2, pp. 133–143. ISSN: 16174909. DOI: 10.1007/s00779-006-0071-x (cit. on p. 16).

Bellotti, V. and Bly, S. (1996). 'Walking Away from the Desktop Computer: Distributed Collaboration and Mobility in a Product Design Team'. In: *Proceedings of the 1996 ACM conference on Computer supported cooperative work*. CSCW '96. New York, NY, USA: ACM, pp. 209–218. ISBN: 0897917650. DOI: 10.1145/240080.240256 (cit. on p. 25).

Bentley, R., Hughes, J. A., Randall, D., Rodden, T., Sawyer, P., Shapiro, D. and Sommerville, I. (1992). 'Ethnographically-informed systems design for air traffic control'. In: *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*. CSCW '92 November. New York, NY, USA: ACM, pp. 123–129. ISBN: 0897915429. DOI: 10.1145/143457.143470 (cit. on pp. 31, 169).

Berger, P. L. and Luckmann, T. (1966). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. New York: Anchor Books (cit. on p. 47).

Bergqvist, J., Dahlberg, P., Ljungberg, F. and Kristoffersen, S. (1999). 'Moving Out of the Meeting Room'. In: *Proceedings of European conference on Computer Supported Cooperative Work*. ECSCW' 99. Dordrecht: Springer Netherlands, pp. 81–98. DOI: 10.1007/978-94-011-4441-4_5 (cit. on p. 24).

Bergstrom-Lehtovirta, J., Eklund, T., Jylhä, A., Kuikkaniemi, K., An, C. and Jacucci, G. (Jan. 2013). 'BubblesDial: Exploring Large Display Content Graphs on Small Devices'. In: *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia*. MUM '13. New York, NY, USA: ACM, p. 1. ISBN: 9781450326483. DOI: 10.1145/2541831.2541845 (cit. on p. 26).

Bittner, E. (1973). 'Objectivity and realism in sociology.' In: *Phenomenological sociology* 110 (cit. on p. 40).

Brown, B. and Juhlin, O. (2015). *Enjoying Machines*. Cambridge, Massachusetts: MIT Press, p. 219. ISBN: 978-0262028783 (cit. on pp. 21, 27).

Brown, B., McGregor, M. and Laurier, E. (2013). 'iPhone in vivo: Video Analysis of Mobile Device Use'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. New York, NY, USA: ACM, p. 1031. ISBN: 9781450318990. DOI: 10.1145/2470654.2466132 (cit. on pp. 33, 155, 162).

Brown, B., McGregor, M. and McMillan, D. (2014). '100 Days of iPhone Use: Understanding the Details of Mobile Device Use'. In: *Proceedings of the Extended Abstracts of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. CHI EA '14. New York, NY, USA: ACM, pp. 2335–2340. ISBN: 9781450324748. DOI: 10.1145/2559206.2581296 (cit. on pp. 1, 4, 33, 155).

Brown, B., McGregor, M. and McMillan, D. (Jan. 2015). 'Searchable Objects: Search in Everyday Conversation'. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. CSCW '15. New York, NY, USA: ACM, pp. 508–517. ISBN: 9781450329224. DOI: 10.1145/2675133.2675206 (cit. on pp. 4, 26, 34, 160, 165, 176).

Bull, M. (Jan. 2005). 'No Dead Air! The iPod and the Culture of Mobile Listening'. In: *Journal of Leisure Studies* 24.4, pp. 343–355. ISSN: 0261-4367. DOI: 10.1080/0261436052000330447 (cit. on pp. 21, 27).

Button, G. (2000). 'The ethnographic tradition and design'. In: *Design Studies*. ISSN: 0142694X. DOI: 10.1016/S0142-694X(00)00005-3 (cit. on p. 41).

Button, G. (2012). 'What does 'work' mean in 'ethnomethodological studies of work?': Its ubiquitous relevance for systems design to support action and interaction'. In: *Design Studies* 33.6, pp. 673–684. ISSN: 0142694X. DOI: 10.1016/j.destud.2012.06.003 (cit. on p. 44).

Button, G., Coulter, J., Lee, J. R. E. and Sharrock, W. (1995). *Computers, Minds and Conduct*. Cambridge, UK: Polity Press, p. 237. ISBN: 0745615716 (cit. on p. 46).

Button, G., Crabtree, A., Rouncefield, M. and Tolmie, P. (Jan. 2015). *Deconstructing Ethnography: Towards a Social Methodology for Ubiquitous Computing and Interactive Systems Design*. Human–Computer Interaction Series. Cham: Springer International Publishing. ISBN: 978-3-319-21953-0. DOI: 10.1007/978-3-319-21954-7 (cit. on pp. 41, 44, 49).

Button, G. and Sharrock, W. (1995). 'On simulacrums of conversation: Toward a clarification of the relevance of conversation analysis for human-computer interaction'. In: *The Social and Interactional Dimensions of Human-Computer Interface*. Ed. by P. J. Thomas. Cambridge, UK: Press Syndicate of the University of Cambridge. Chap. 6, pp. 107–125 (cit. on pp. 198, 200).

Button, G. and Sharrock, W. (Jan. 2009). 'Studies of Work and the Workplace in HCI: Concepts and Techniques'. In: *Synthesis Lectures on Human-Centered Informatics* 2.1, pp. 1–96. ISSN: 1946-7680. DOI: 10.2200/S00177ED1V01Y200903HCI003 (cit. on p. 50).

Campbell, S. W. (Feb. 2007). 'Perceptions of Mobile Phones in College Classrooms: Ringing, Cheating, and Classroom Policies'. In: *Communication Education* 55.3, pp. 280–294. DOI: 10.1080/03634520600748573 (cit. on p. 18).

Campbell, S. W. and Park, Y. J. (Mar. 2008). 'Social Implications of Mobile Telephony: The Rise of Personal Communication Society'. In: *Sociology Compass* 2.2, pp. 371–387. DOI: 10.1111/j.1751-9020.2007.00080.x (cit. on p. 22).

Cecchinato, M. E., Cox, A. L. and Bird, J. (2017). 'Always On(line)? User Experience of Smartwatches and their Role within Multi-Device Ecologies'. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. New York, NY, USA: ACM, pp. 3557–3568. ISBN: 9781450346559. DOI: 10.1145/3025453.3025538 (cit. on p. 20).

Chattopadhyay, D., Salvadori, F., O'Hara, K. and Rintel, S. (Sept. 2018). 'Beyond Presentation: Shared Slideware Control as a Resource for Collocated Collaboration'. In: *Human–Computer Interaction* 33.5-6, pp. 455–498. ISSN: 0737-0024. DOI: 10.1080/07370024.2017.1388170 (cit. on p. 24).

Church, K., Cousin, A. and Oliver, N. (Sept. 2012). 'I Wanted to Settle a Bet! - Understanding Why and How People Use Mobile Search in Social Settings'. In: *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '12. New York, NY, USA: ACM, p. 393. ISBN: 9781450311052. DOI: 10.1145/2371574.2371635 (cit. on pp. 18, 26).

Church, K. and Oliver, N. (2011). 'Understanding Mobile Web and Mobile Search Use in Today's Dynamic Mobile Landscape'. In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI '11. New York, NY, USA: ACM, p. 67. ISBN: 9781450305419. DOI: 10.1145/2037373.2037385 (cit. on pp. 25, 161).

Church, K. and Smyth, B. (2008). 'Understanding the Intent Behind Mobile Information Needs'. In: *Proceedingsc of the 13th International Conference on Intelligent User Interfaces*. IUI '09. New York, NY, USA: ACM, p. 247. ISBN: 9781605581682. DOI: 10.1145/1502650.1502686 (cit. on p. 30).

Ciolfi, L. and Bannon, L. (2003). 'Learning from museum visits: Shaping design sensitivities'. In: *Proceedings of HCI International 2003* (cit. on pp. 25, 154).

Clark, H. H. (1996). *Using language*. Cambridge University Press (cit. on p. 205).

Clark, H. H. and Brennan, S. E. (1991). 'Grounding in Communication'. In: *Perspectives on Socially Shared Cognition*, pp. 222–233. ISSN: 14726920. DOI: 10.1037/10096-006 (cit. on p. 88).

Clarke, I., Kell, I., Schmidt, R. and Vignali, C. (Oct. 2000). 'Thinking the thoughts they do: Symbolism and meaning in the consumer experience of the "British pub"'. In: *British Food Journal* 102.9, pp. 692–710. DOI: 10.1108/13522759810235197 (cit. on p. 58).

Cole, H. and Stanton, D. (Dec. 2003). 'Designing mobile technologies to support co-present collaboration'. In: *Personal and Ubiquitous Computing* 7.6, pp. 365–371. DOI: 10.1007/s00779-003-0249-4 (cit. on pp. 24, 26).

Counts, S. and Fellheimer, E. (Jan. 2004). 'Supporting Social Presence Through Lightweight Photo Sharing On and Off the Desktop'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '04. New York, NY, USA: ACM, pp. 599–606. ISBN: 1-58113-702-8. DOI: 10.1145/985692.985768 (cit. on p. 25).

Crabtree, A. (2003). *Designing Collaborative Systems: A Practical Guide to Ethnography*. London: Springer-Verlag. ISBN: 1-85233-718-4. DOI: 10.1007/b97516 (cit. on p. 31).

Crabtree, A., Benford, S., Greenhalgh, C., Tennent, P., Chalmers, M. and Brown, B. (Jan. 2006). 'Supporting Ethnographic Studies of Ubiquitous Computing in the Wild'. In: *Proceedings of the 6th ACM Conference on Designing Interactive Systems*. DIS '06. New York, NY, USA: ACM, p. 60. ISBN: 1595933670. DOI: 10.1145/1142405.1142417 (cit. on pp. 44, 83).

Crabtree, A., Rodden, T., Hemmings, T. and Benford, S. (2003). 'Finding a Place for UbiComp in the Home'. In: *UbiComp 2003: Ubiquitous Computing*. Ed. by A. K. Dey, A. Schmidt and J. F. McCarthy. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 208–226. ISBN: 978-3-540-39653-6. DOI: 10.1007/978-3-540-39653-6_17 (cit. on pp. 123, 193, 194).

Crabtree, A., Rodden, T., Tolmie, P. and Button, G. (2009). 'Ethnography Considered Harmful'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09 October 2015. New York, NY, USA: ACM, p. 879. ISBN: 9781605582467. DOI: 10.1145/1518701.1518835 (cit. on pp. 23, 32, 35, 41, 44, 179).

Crabtree, A., Rouncefield, M. and Tolmie, P. (Jan. 2012). *Doing Design Ethnography*. Human-Computer Interaction Series. London: Springer London. ISBN: 978-1-4471-2725-3. DOI: 10.1007/978-1-4471-2726-0 (cit. on pp. 2, 35, 37, 41–47, 49, 56, 61, 84, 93, 127, 167, 173, 174, 177, 178, 193, 204).

Crabtree, A. and Tolmie, P. (2016). 'A Day in the Life of Things in the Home'. In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '16. New York, NY, USA: ACM, pp. 1736–1748. ISBN: 9781450335928. DOI: 10.1145/2818048.2819954 (cit. on p. 25).

Crabtree, A., Tolmie, P. and Rouncefield, M. (2013). '"How Many Bloody Examples Do You Want?" Fieldwork and Generalisation'. In: *Proceedings of the 13th European Conference on Computer Supported Cooperative Work*. ECSCW 2013. London: Springer London, pp. 1–20. ISBN: 978-1-4471-5345-0. DOI: 10.1007/978-1-4471-5346-7_1 (cit. on p. 43).

Cutrell, E. B., Czerwinski, M. and Horvitz, E. (Apr. 2000). 'Effects of Instant Messaging Interruptions on Computing Tasks'. In: *CHI '00 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '00. New York, NY, USA: ACM, pp. 99–100. ISBN: 1-58113-248-4. DOI: 10.1145/633292.633351 (cit. on p. 20).

Cycil, C., Perry, M., Laurier, E. and Taylor, A. S. (2013). ''Eyes free' in-car assistance'. In: *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '13. New York, NY, USA: ACM, p. 332. ISBN: 9781450322737. DOI: 10.1145/2493190.2493207 (cit. on p. 88).

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., Georgila, K., Gratch, J., Hartholt, A., Lhommet, M., Lucas, G., Marsella, S., Morbini, F., Nazarian, A., Scherer, S., Stratou, G., Suri, A., Traum, D., Wood, R., Xu, Y., Rizzo, A. and Morency, L.-p. (2014). 'SimSensei Kiosk: A Virtual Human Interviewer for Healthcare Decision Support'. In: *International Conference on Autonomous Agents and Multi-Agent Systems*. Ed. by L. Alessio, P. Scerri, A. Bazzan and M. Huhns. AAMAS '14 1. International Foundation for Autonomous Agents and Multiagent Systems, pp. 1061–1068. ISBN: 1450327389 (cit. on p. 143).

Dey, A. K. and Häkkilä, J. (2008). 'Context-Awareness and Mobile Devices'. In: *Mobile Computing: Concepts, Methodologies, Tools, and Applications. IGI Global*. IGI Global. Chap. 13, pp. 205–217. ISBN: 9781605660547. DOI: 10.4018/978-1-60566-054-7.ch238 (cit. on p. 28).

DiDomenico, S. M. and Boase, J. (2013). 'Bringing mobiles into the conversation: Applying a conversation analytic approach to the study of mobiles in co-present interaction'. In: *Discourse 2.0: Language and New Media (pp.* Ed. by D. Tannen and A. M. Trester. Vol. 2. Georgetown, DC, USA: Georgetown University Press, pp. 119–132. ISBN: 1589019547 (cit. on p. 33).

Dourish, P. (2014). 'Reading and Interpreting Ethnography'. In: *Ways of Knowing in HCI*. Ed. by J. S. Olson and W. A. Kellog. New York, NY, USA: Springer. Chap. 1, pp. 1–23. ISBN: 9781493903788. DOI: 10.1007/978-1-4939-0378-8_1 (cit. on p. 49).

Dourish, P. and Button, G. (1998). 'On "Technomethodology": Foundational Relationships Between Ethnomethodology and System Design'. In: *Human-Computer Interaction* 13.4, pp. 395–432. ISSN: 0737-0024. DOI: 10.1207/s15327051hci1304_2 (cit. on p. 171).

Dow, S. P., Lee, J., Oezbek, C., MacIntyre, B., Bolter, J. D. and Gandy, M. (2005). 'Wizard of Oz Interfaces for Mixed Reality Applications'. In: *CHI '05 extended abstracts on Human factors in computing systems* -. CHI '05. New York, NY, USA: ACM, p. 1339. ISBN: 1595930027. DOI: 10.1145/1056808.1056911 (cit. on p. 36).

Durrant, A., Rowland, D., Kirk, D. S., Benford, S., Fischer, J. E. and McAuley, D. (2011). 'Automics: Souvenir Generating Photoware for Theme Parks'. In: *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems*. CHI '11. New York, NY, USA: ACM, p. 1767. ISBN: 9781450302289. DOI: 10.1145/1978942.1979199 (cit. on pp. 3, 4, 25).

Edelsky, C. (Dec. 1981). 'Who's got the floor?' In: *Language in Society* 10.03, pp. 383–421. DOI: 10.1017/S004740450000885X (cit. on pp. 68, 102).

Edwards, W. K. and Grinter, R. E. (2001). 'At Home with Ubiquitous Computing: Seven Challenges'. In: *Ubicomp 2001: Ubiquitous Computing*. Berlin, Heidelberg: Springer, pp. 256–272. DOI: 10.1007/3-540-45427-6_22 (cit. on p. 25).

Ellis, C. A., Gibbs, S. J. and Rein, G. (Jan. 1991). 'Groupware: some issues and experiences'. In: *Communications of the ACM* 34.1, pp. 39–58. ISSN: 00010782. DOI: 10.1145/99977.99987 (cit. on pp. 24, 196).

Fatah gen. Schieck, A., Schnädelbach, H., Motta, W., Behrens, M., North, S., Ye, L. and Kostopoulou, E. (2014). 'Screens in the Wild'. In: *Proceedings of The International Symposium on Pervasive Displays*. PerDis'14. New York, NY, USA: ACM, pp. 166–167. ISBN: 9781450329521. DOI: 10.1145/2611009.2617199 (cit. on p. 4).

Ferdous, H. S., Vetere, F., Davis, H., Ploderer, B. and O'Hara, K. (Jan. 2016). 'Technologies At Mealtime: Collocated Interactions In The Family Home'. In: *CHI '16 Workshop on Proxemic Mobile Collocated Interactions* (cit. on pp. 118, 154, 170).

Fischer, J. E. (May 2010). 'Interrupting the Here and Now: Implications and Opportunities'. In: *Mensch & Computer 2010: Interaktive Kulturen*, pp. 159–168. ISBN: 978-3-486-70408-2 (cit. on p. 20).

Fischer, J. E., Greenhalgh, C. and Benford, S. (Aug. 2011). 'Investigating Episodes of Mobile Phone Activity as Indicators of Opportune Moments to Deliver Notifications'. In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI '11. New York, NY, USA: ACM, pp. 181–190. ISBN: 978-1-4503-0541-9. DOI: 10.1145/2037373.2037402 (cit. on pp. 29, 161).

Fischer, J. E., Porcheron, M., Lucero, A., Quigley, A., Scott, S., Ciolfi, L., Rooksby, J. and Memarovic, N. (2016). 'Collocated Interaction: New Challenges in 'Same Time, Same Place' Research'. In: *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*. CSCW '16 Companion. New York, NY, USA: ACM, pp. 465–472. ISBN: 9781450339506. DOI: 10.1145/2818052.2855522 (cit. on pp. 24, 196).

Fischer, J. E., Reeves, S., Moran, S., Greenhalgh, C., Benford, S. and Rennick-Egglestone, S. (July 2013). 'Understanding Mobile Notification Management in Collocated Groups'. In: *Proceedings of the 13th European Conference on Computer Supported Cooperative Work*. ECSCW '13 Chapter 2. London: Springer London, pp. 21–44. ISBN: 978-1-4471-5345-0. DOI: 10.1007/978-1-4471-5346-7_2 (cit. on pp. 20, 33).

Fischer, J. E., Yee, N., Bellotti, V., Good, N., Benford, S. and Greenhalgh, C. (Jan. 2010). 'Effects of Content and Time of Delivery on Receptivity to Mobile Interruptions'. In: *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*. MobileHCI '10. New York, NY, USA: ACM, p. 103. ISBN: 9781605588353. DOI: 10.1145/1851600.1851620 (cit. on p. 54).

Flintham, M. D., Velt, R., Wilson, M. L., Anstead, E. J., Benford, S., Brown, A., Pearce, T., Price, D. and Sprinks, J. (2015). 'Run Spot Run: Capturing and Tagging Footage of a Race by Crowds of Spectators'. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. New York, NY, USA: ACM, pp. 747–756. ISBN: 9781450331456. DOI: 10.1145/2702123.2702463 (cit. on p. 3).

Forlizzi, J. (2007). 'How Robotic Products Become Social Products: An Ethnographic Study of Cleaning in the Home'. In: *Proceeding of the ACM/IEEE international conference on Human-robot interaction*. HRI '07. New York, NY, USA: ACM, p. 129. ISBN: 9781595936172. DOI: 10.1145/1228716.1228734 (cit. on p. 172).

Fosh, L., Benford, S., Reeves, S., Koleva, B. and Brundell, P. (Jan. 2013). ''See Me, Feel Me, Touch Me, Hear Me': Trajectories and Interpretation in a Sculpture Garden'. In: *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*. CSCW '13. New York, NY, USA: ACM, pp. 110–149. ISBN: 9781450318990. DOI: 10.1145/2470654.2470675 (cit. on pp. 3, 154).

Fox, K. (1996). *Passport to the Pub: Tourist's Guide to Pub Etiquette*. The Do-Not Press, p. 80. ISBN: 978-1899344093 (cit. on p. 53).

Fox, K. (2004). *Watching the English: The Hidden Rules of English Behaviour*. London: Hodder & Stroughton. ISBN: 1848940505 (cit. on pp. 58, 154).

Fraser, N. M. and Gilbert, G. N. (Jan. 1991). 'Simulating speech systems'. In: *Computer Speech & Language* 5.1, pp. 81–99. ISSN: 10958363. DOI: 10.1016/0885-2308(91)90019-M (cit. on p. 35).

Fuentes, C., Porcheron, M., Fischer, J. E., Costanza, E., Malik, O. and Ramchurn, S. D. (2019a). 'Tracking the Consumption of Home Essentials'. In: *CHI Conference on Human Factors in Computing Systems Proceedings*. CHI '19. New York, NY, USA: ACM. ISBN: 9781450359702. DOI: 10.1145/3290605.3300869 (cit. on p. 172).

Fuentes, C., Porcheron, M., Fischer, J. E., Costanza, E., Verdezoto, N., Herskovic, V., Zuckerman, O. and Takayama, L. (2019b). 'New Directions for the IoT: Automate, Share, Build, and Care'. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI EA '19. New York, NY, USA: ACM, pp. 1–8. ISBN: 9781450359719. DOI: 10.1145/3290607.3299000. URL: http://dl.acm.org/citation.cfm?doid=3290607.3299000 (cit. on p. 194).

Garfinkel, H. (n.d.). 'Missed Orderliness' (cit. on p. 44).

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, NJ, USA: Prentice-Hall, p. 304. ISBN: 978-0-7456-0005-5 (cit. on pp. 2, 4, 39, 43, 46, 47, 55, 91, 92, 157, 162).

Garfinkel, H. (2002a). *Ethnomethodology's Program: Working Out Durkheim's Aphorism*. Ed. by A. W. Rawls. Rowman & Littlefield Publishers, Inc. ISBN: 0-7425-1642-3 (cit. on pp. 2, 34, 58, 128, 155).

Garfinkel, H. (2002b). *Ethnomethodology's Program: Working Out Durkheim's Aphorism*. Ed. by A. Rawls. Legacies of Social Thought. Rowman & Littlefield Publishers, Inc (cit. on p. 38).

Garfinkel, H. and Sacks, H. (1970). 'On Formal Structures of Practical Actions'. In: *Theoretical Sociology: Perspectives and Developments*, pp. 338–366 (cit. on pp. 4, 47, 48).

Garfinkel, H. and Wieder, D. L. (1992a). 'Two Incommensurable, Asymmetrically Alternate Technologies of Social Analysis'. In: *Text in context: contributions to ethnomethodology*. ISBN: 0803942532 (cit. on p. 50).

Garfinkel, H. and Wieder, D. L. (Jan. 1992b). 'Two Incommensurable, Asymmetrically Alternate Technologies of Social Analysis'. In: *Text in Context Contributions to Ethnomethodology*. Sage, pp. 175–206. URL: http://scholar.google.com/scholar?q=related:%7B%5C_%7DhqaHTplX%7B%5C_%7DQJ:scholar.google.com/%7B%5C&%7Dhl=en%7B%5C&%7Dnum=20%7B%5C&%7Das%7B%5C_%7Dsdt=0,5%20papers3://publication/uuid/83B6EBE6-C947-4565-BCC8-5CB055746109 (cit. on p. 195).

Geertz, C. (1973). 'Thick description: Toward an interpretive theory of culture.' In: *Interpretation of Cultures: Selected Essays by Clifford Geertz*. New York, NY, USA: Basic Books, p. 3 (cit. on p. 43).

Gellersen, H. W., Schmidt, A. and Beigl, M. (2002). 'Multi-Sensor Context-Awareness in Mobile Devices and Smart Artefacts'. In: *Mobile Networks and Applications* 5.2, pp. 341–351. DOI: 10.1023/A:1016587515822 (cit. on p. 28).

Geser, H. (Jan. 2006). 'Is the Cell Phone Undermining the Social Order?: Understanding Mobile Technology from a Sociological Perspective'. In: *Knowledge, Technology & Policy* 19.1, pp. 8–18. DOI: 10.1007/s12130-006-1010-x (cit. on pp. 21, 22).

Gibson, J. J. (1979). 'The Theory of Affordances'. In: *The Ecological Approach to Visual Perception*. Dallas: Houghtom Miffli. ISBN: 0-89859-959-8. DOI: citeulike-article-id:3508530 (cit. on p. 136).

Gjerlufsen, T., Klokmose, C. N., Eagan, J., Pillias, C. and Beaudouin-Lafon, M. (2011). 'Shared substance: developing flexible multi-surface applications'. In: *Proceedings of the 2011 annual conference on Human factors in computing systems*. CHI '11. New York, NY, USA: ACM, p. 3383. ISBN: 9781450302289. DOI: 10.1145/1978942.1979446 (cit. on p. 25).

Goffman, E. (Jan. 1968). *Behavior in Public Places: Notes the Social Organisation of Gatherings*. New York: Simon and Schuster, p. 248. ISBN: 1439108692 (cit. on p. 2).

Goodwin, C. and Goodwin, M. H. (Jan. 1996). 'Seeing as Situated Activity: Formulating Planes'. In: *Cognition and Communication at Work*, pp. 61–95 (cit. on p. 31).

Goodwin, C. and Heritage, J. (Oct. 1990). 'Conversation Analysis'. In: *Annual Review of Anthropology* 19.1, pp. 283–307. DOI: 10.1146/annurev.an.19.100190.001435 (cit. on p. 60).

Grudin, J. (1990). 'The Computer Reaches Out: The Historical Continuity of Interface Design'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '90. New York, NY, USA: ACM, pp. 261–268. ISBN: 0201509326. DOI: 10.1145/97243.97284 (cit. on pp. 169, 179, 195).

Haddington, P., Keisanen, T., Mondada, L. and Nevile, M. (2014). *Multiactivity in social interaction: Beyond multitasking*. John Benjamins Publishing Company (cit. on p. 2).

Harmon, E. and Mazmanian, M. (2013). 'Stories of the Smartphone in Everyday Discourse: Conflict, Tension & Instability'. In: *Human Factors in Computing Systems*, pp. 1051–1060. DOI: 10.1145/2470654.2466134 (cit. on p. 4).

Heath, C., Hindmarsh, J. and Luff, P. (Feb. 2010). *Video in Qualitative Research: Analysing Social Interaction in Everyday Life*. SAGE, p. 184. ISBN: 9781446246696 (cit. on pp. 2, 49, 60, 62, 92, 173, 193, 214).

Heath, C., Jirotka, M., Luff, P. and Hindmarsh, J. (June 1994). 'Unpacking Collaboration: the Interactional Organisation of Trading in a City Dealing Room'. In: *Computer Supported Cooperative Work*. CSCW 3.2, pp. 147–165. ISSN: 0925-9724. DOI: 10.1007/BF00773445 (cit. on pp. 169, 179).

Heath, C. and Luff, P. (Jan. 1991). 'Disembodied conduct: communication through video in a multi-media office environment'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '91. New York, NY, USA: ACM. DOI: 10.1145/108844.108859 (cit. on p. 194).

Heath, C. and Luff, P. (1992). 'Collaboration and control: Crisis management and multimedia technology in London Underground control rooms'. In: *Computer Supported Cooperative Work* 1.1990, pp. 69–94. DOI: 10.1007/BF00752451 (cit. on p. 31).

Hoey, E. M. (2015). 'Lapses: How People Arrive at, and Deal With, Discontinuities in Talk'. In: *Research on Language and Social Interaction* 48.4, pp. 430–453. ISSN: 0835-1813. DOI: 10.1080/08351813.2015.1090116 (cit. on p. 107).

Höysniemi, J., Hämäläinen, P. and Turkki, L. (2004). 'Wizard of Oz Prototyping of Computer Vision Based Action Games for Children'. In: *Proceeding of the 2004 conference on Interaction design and children building a community*. IDC '04. New York, NY, USA: ACM, pp. 27–34. ISBN: 1581137915. DOI: 10.1145/1017833.1017837 (cit. on p. 36).

Hudson, J. M., Christensen, J., Kellogg, W. A. and Erickson, T. (Jan. 2002). '"I'd Be Overwhelmed, But It's Just One More Thing To Do": Availability and Interruption in Research Management'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '02. New York, NY, USA: ACM, pp. 97–98. ISBN: 1581134533. DOI: 10.1145/503376.503394 (cit. on p. 20).

Hughes, E. C. (1958). 'Men and Their Work.' In: *American Sociological Review*. ISSN: 00031224. DOI: 10.2307/2089393 (cit. on p. 41).

Hughes, J., King, V., Rodden, T. and Andersen, H. (1994). 'Moving Out from the Control Room: Ethnography in System Design'. In: *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. CSCW'94. New York, NY, USA: ACM, pp. 429–439. ISBN: 0897916891. DOI: 10.1145/192844.193065 (cit. on p. 31).

Humphreys, L. (Dec. 2005). 'Cellphones in public: social interactions in a wireless era'. In: *New Media & Society* 7.6, pp. 810–833. ISSN: 1461-4448. DOI: 10.1177/1461444805058164 (cit. on p. 21).

Humphreys, L., Von Pape, T. and Karnowski, V. (May 2013). 'Evolving Mobile Media: Uses and Conceptualizations of the Mobile Internet'. In: *Journal of Computer-Mediated Communication* 18.4, pp. 491–507. DOI: 10.1111/jcc4.12019 (cit. on p. 19).

Hurter, C., Lesbordes, R., Letondal, C., Vinot, J.-L. and Conversy, S. (2012). 'Strip'TIC'. In: *Proceedings of the International Working Conference on Advanced Visual Interfaces*. AVI '12. New York, NY, USA: ACM, p. 225. DOI: 10.1145/2254556.2254598 (cit. on p. 25).

Iqbal, S. T. and Horvitz, E. (2010). 'Notifications and awareness: A field study of alert usage and preferences'. In: *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. CSCW '10, pp. 27–30. ISBN: 9781605587950. DOI: 10.1145/1718918.1718926 (cit. on p. 20).

Isaacs, E., Szymanski, M., Yamauchi, Y., Glasnapp, J. and Iwamoto, K. (2012). 'Integrating Local and Remote Worlds Through Channel Blending'. In: *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. CSCW '12. New York, NY, USA: ACM, p. 617. ISBN: 9781450310864. DOI: 10.1145/2145204.2145299 (cit. on p. 34).

Jarusriboonchai, P., Lundgren, S., Olsson, T., Fischer, J., Memarovic, N., Reeves, S., Woźniak, P. and Torgersson, O. (2014). 'Personal or Social? Designing Mobile Interactions for Co-located Interaction'. In: *Proceedings of the 8th Nordic Conference on Human-Computer Interaction*. NordiCHI '14. New York, NY, USA: ACM, pp. 829–832. ISBN: 9781450325424. DOI: 10.1145/2639189.2654840 (cit. on p. 24).

Jokela, T., Ojala, J. and Olsson, T. (2015). 'A Diary Study on Combining Multiple Information Devices in Everyday Activities and Tasks'. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. New York, NY, USA: ACM, pp. 3903–3912. ISBN: 9781450331456. DOI: 10.1145/2702123.2702211 (cit. on pp. 118, 172).

Jones, A., Barthes, J.-P. A., Moulin, C. and Lenne, D. (2014). 'A rich multi-agent architecture for collaborative multi-touch multi-user devices'. In: *2014 IEEE International Conference on Systems, Man, and Cybernetics*. SMC '14, pp. 1107–1112. ISBN: 978-1-4799-3840-7. DOI: 10.1109/SMC.2014.6974062 (cit. on p. 87).

Jones, A., Moulin, C., Barthes, J.-P. A., Lenne, D., Kendira, A. and Gidel, T. (2012). 'Personal Assistant Agents and Multi-agent Middleware for CSCW'. In: *Proceedings of the 2012 IEEE 16th International Conference on Computer Supported Cooperative Work in Design*. CSCWD 2012, pp. 72–79. ISBN: 9781467312127. DOI: 10.1109/CSCWD.2012.6221800 (cit. on pp. 25, 182).

Jordan, B. and Henderson, A. (Jan. 1995). 'Interaction Analysis: Foundations and Practice'. In: *Journal of the Learning Sciences* 4.1, pp. 39–103. ISSN: 1050-8406. DOI: 10.1207/s15327809jls0401_2 (cit. on p. 56).

Kadous, M. W. and Sammut, C. (2004). 'InCa: A Mobile Conversational Agent'. In: *PRICAI 2004: Trends in Artificial Intelligence* 3157, pp. 644–653. ISSN: 03029743. DOI: 10.1007/978-3-540-28633-2_68 (cit. on p. 87).

Kendrick, K. H. and Drew, P. (Jan. 2016). 'Recruitment: Offers, Requests, and the Organization of Assistance in Interaction'. In: *Research on Language and Social Interaction* 49.1, pp. 1–19. ISSN: 0835-1813. DOI: 10.1080/08351813.2016.1126436 (cit. on p. 108).

Kern, N. and Schiele, B. (2003). 'Context-aware notification for wearable computing'. In: *Seventh IEEE International Symposium on Wearable Computers, 2003*. ISWC '03. IEEE, pp. 223–230. ISBN: 0-7695-2034-0. DOI: 10.1109/ISWC.2003.1241415 (cit. on p. 29).

Klemmer, S. R., Sinha, A. K., Chen, J., Landay, J. A., Aboobaker, N. and Wang, A. (2000). 'Suede: A Wizard of Oz Prototyping Tool for Speech User Interfaces'. In: *Proceedings of the 13th annual ACM symposium on User interface software and technolog*. Vol. 2. UIST '00. New York, NY, USA: ACM, pp. 1–10. ISBN: 1581132123. DOI: 10.1145/354401.354406 (cit. on p. 35).

Kopp, S., Gesellensetter, L., Krämer, N. C. and Wachsmuth, I. (2005). 'A Conversational Agent as Museum Guide – Design and Evaluation of a Real-World Application'. In: *Lecture Notes in Computer Science*. Vol. 3661 LNAI, pp. 329–343. ISBN: 3540287388. DOI: 10.1007/11550617_28 (cit. on p. 85).

Krehl, C., Sharples, S. C. and Flintham, M. D. (Sept. 2013). 'Less is More: Classifying Mobile Interactions to Support Context Sensing in Journeys'. In: *Proceedings of the 27th International BCS Human Computer Interaction Conference*. BCS-HCI '13. British Computer Society: British Computer Society, pp. 8–10 (cit. on p. 32).

Kumar, A., Paek, T. and Lee, B. (2012). 'Voice Typing: A New Speech Interaction Model for Dictation on Touchscreen Devices'. In: *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems*. CHI '12. New York, NY, USA: ACM, p. 2277. ISBN: 9781450310154. DOI: 10.1145/2207676.2208386 (cit. on p. 88).

Kushlev, K., Proulx, J. and Dunn, E. W. (2016). '"Silence Your Phones": Smartphone Notifications Increase Inattention and Hyperactivity Symptoms'. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16, pp. 1011–1020. ISBN: 9781450333627. DOI: 10.1145/2858036.2858359 (cit. on p. 54).

Lanigan, J. D. (2009). 'A Sociotechnological Model for Family Research and Intervention: How Information and Communication Technologies Affect Family Life Jane'. In: *Marriage & Family Review* 45.6-8, pp. 587–609. ISSN: 0149-4929. DOI: 10.1080/01494920903224194. arXiv: arXiv:1011.1669v3 (cit. on p. 21).

Laurier, E. (Apr. 2008a). 'Drinking up endings: Conversational resources of the café'. In: *Language & Communication* 28.2, pp. 165–181. DOI: 10.1016/j.langcom.2008.01.011 (cit. on p. 154).

Laurier, E. (2008b). 'How Breakfast Happens in the Café'. In: *Time & Society*. ISSN: 0961463X. DOI: 10.1177/0961463X07086306 (cit. on p. 89).

Laurier, E., Whyte, A. and Buckner, K. (July 2001). 'An ethnography of a neighbourhood café: informality, table arrangements and background noise'. In: *Journal of Mundane Behaviour* 2.2, pp. 195–232 (cit. on pp. 57, 89, 165).

Lee, J., Kwon, J. and Kim, H. (2016). 'Reducing distraction of smartwatch users with deep learning'. In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. MobileHCI '16. New York, NY, USA: ACM, pp. 948–953. ISBN: 9781450344135. DOI: 10.1145/2957265.2962662 (cit. on p. 29).

Licklider, J. C. R. (1960). 'Man-Computer Symbiosis'. In: *IRE Transactions on Human Factors in Electronics* HFE-1.1, pp. 4–11. ISSN: 0099-4561. DOI: 10.1109/THFE2.1960.4503259 (cit. on p. 85).

Livingston, E. (1987). *Making sense of ethnomethodology*. Taylor & Francis. ISBN: 978-0710212610 (cit. on p. 48).

Lopez-Tovar, H., Charalambous, A. and Dowell, J. (Mar. 2015). 'Managing Smartphone Interruptions through Adaptive Modes and Modulation of Notifications'. In: *Proceedings of the 20th International Conference on Intelligent User Interfaces*. IUI '15. New York, NY, USA: ACM, pp. 296–299. ISBN: 978-1-4503-3306-1. DOI: 10.1145/2678025.2701390 (cit. on p. 20).

Lucero, A., Holopainen, J. and Jokela, T. (Jan. 2012). 'MobiComics: Collaborative Use of Mobile Phones and Large Displays for Public Expression'. In: *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '12. New York, NY, USA: ACM, p. 383. ISBN: 9781450311052. DOI: 10.1145/2371574.2371634 (cit. on pp. 3, 25, 27, 177).

Lucero, A., Jones, M., Jokela, T. and Robinson, S. (Mar. 2013). 'Mobile Collocated Interactions: Taking an Offline Break Together'. In: *interactions* 20.2, p. 26. ISSN: 10725520. DOI: 10.1145/2427076.2427083 (cit. on pp. 24, 26).

Lucero, A., Keränen, J. and Jokela, T. (Apr. 2010a). 'Social and Spatial Interactions: Shared Co-Located Mobile Phone Use'. In: *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*. CHI EA '10. New York, NY, USA: ACM, pp. 3223–3228. ISBN: 978-1-60558-930-5. DOI: 10.1145/1753846.1753962 (cit. on p. 177).

Lucero, A., Keränen, J. and Korhonen, H. (Jan. 2010b). 'Collaborative Use of Mobile Phones for Brainstorming'. In: *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI '10. New York, NY, USA: ACM, p. 337. ISBN: 9781605588353. DOI: 10.1145/1851600.1851659 (cit. on pp. 176, 182).

Lucero, A. and Vetek, A. (Nov. 2014). 'NotifEye: Using Interactive Glasses to Deal with Notifications while Walking in Public'. In: *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology*. ACE '14. New York, NY, USA: ACM. ISBN: 9781450329453. DOI: 10.1145/2663806.2663824 (cit. on p. 29).

Luff, P. and Heath, C. (1998). 'Mobility in Collaboration'. In: *Proceedings of the 1998 ACM conference on Computer supported cooperative work*. CSCW '98. New York, NY, USA: ACM, pp. 305–314. ISBN: 1581130090. DOI: 10.1145/289444.289505 (cit. on p. 25).

Luger, E. and Sellen, A. (May 2016). '"Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents'. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. New York, NY, USA: ACM, pp. 5286–5297. ISBN: 9781450333627. DOI: 10.1145/2858036.2858288 (cit. on p. 88).

Lundgren, S., Fischer, J. E., Reeves, S. and Torgersson, O. (Jan. 2015). 'Designing Mobile Experiences for Collocated Interaction'. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. CSCW '15. New York, NY, USA: ACM, pp. 496–507. ISBN: 9781450329224. DOI: 10.1145/2675133.2675171 (cit. on pp. 3, 26).

Lundgren, S. and Torgersson, O. (2013). 'Bursting the Mobile Bubble'. In: *First International Workshop on Designing Mobile Face-to-Face Group Interactions at ECSCW '13*. 2008. URL: http://www.cse.chalmers.se/research/group/idc/ituniv/courses/13/mc/lundgren-mogi2013.pdf (cit. on p. 27).

Lynch, M. (1993). *Scientific Practice and Ordinary Action: Ethnomethodology and Social Studies of Science*. Cambridge University Press. ISBN: 9780511625473. DOI: 10.1017/CBO9780511625473 (cit. on p. 48).

Malinowski, B. (1922). *Argonauts of the Western Pacific: An account of native enterprise and adventure in the Archipelagoes of Melanesian New Guinea*. London: G. Routledge & Sons. ISBN: 8135977840 (cit. on pp. 39, 40).

Mark, G., Voida, S. and Cardello, A. (2012). '"A Pace Not Dictated by Electrons": An Empirical Study of Work Without Email Gloria'. In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*. CHI '12. New York, NY, USA: ACM, p. 555. ISBN: 9781450310154. DOI: 10.1145/2207676.2207754 (cit. on p. 20).

Martelaro, N. and Ju, W. (2017). 'WoZ Way: Enabling Real-Time Remote Interaction Prototyping & Observation in On-Road Vehicles'. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. CSCW '17. New York, NY, USA: ACM, pp. 169–182. ISBN: 9781450343350. DOI: 10.1145/2998181.2998293 (cit. on p. 36).

Mass Observation (1943). *The Pub & the People: A Worktown Study*. London: Victor Gollancz Ltd (cit. on p. 55).

McDonagh, E. C. (1950). 'Television and the Family'. In: *Sociology & Social Research* (cit. on p. 18).

McGregor, M. and Tang, J. (2017). 'More to Meetings: Challenges in Using Speech-Based Technology to Support Meetings'. In: *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '17. New York, NY, USA: ACM. DOI: 10.1145/2998181.2998335 (cit. on p. 118).

McMillan, D., Loriette, A. and Brown, B. (Apr. 2015). 'Repurposing Conversation: Experiments with the Continuous Speech Stream'. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. New York, NY, USA: ACM, pp. 3953–3962. ISBN: 9781450331456. DOI: 10.1145/2702123.2702532 (cit. on p. 29).

McTear, M., Callejas, Z. and Griol, D. (2016). *The Conversational Interface: Talking to Smart Devices*. Springer International Publishing. ISBN: 978-3-319-32965-9. DOI: 10.1007/978-3-319-32967-3 (cit. on pp. 88, 119, 197).

Mehrotra, A., Pejovic, V., Vermeulen, J., Hendley, R. and Musolesi, M. (Jan. 2016). 'My Phone and Me: Understanding People's Receptivity to Mobile Notifications'. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. New York, NY, USA: ACM, pp. 1–12. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858566 (cit. on p. 54).

Memarovic, N., Fatah gen. Schieck, A., Schnädelbach, H., Kostopoulou, E., North, S. and Ye, L. (2016). 'Longitudinal, Cross-site and "In the Wild": a Study of Public Displays User Communities' Situated Snapshots'. In: *Proceedings of the 3rd Conference on Media Architecture Biennale*. MAB '16. New York, NY, USA: ACM, pp. 1–10. ISBN: 9781450347495. DOI: 10.1145/2946803.2946804 (cit. on p. 3).

Molnar, K. K. and Kletke, M. G. (Sept. 1996). 'The impacts on user performance and satisfaction of a voice-based front-end interface for a standard software tool'. In: *International Journal of Human-Computer Studies* 45.3, pp. 287–303. ISSN: 10715819. DOI: 10.1006/ijhc.1996.0053 (cit. on p. 171).

Murtagh, G. M. (2002). 'Seeing the "Rules": Preliminary Observations of Action, Interaction and Mobile Phone Use'. In: *Wireless World: Social and Interactional Aspects of the Mobile Age*. London: Springer-Verlag. Chap. 6, pp. 81–91. ISBN: 978-1-85233-477-2. DOI: 10.1007/978-1-4471-0665-4_6 (cit. on p. 32).

Nakamura, T. (Feb. 2015). 'The action of looking at a mobile phone display as nonverbal behavior/communication: A theoretical perspective'. In: *Computers in Human Behavior* 43, pp. 68–75. DOI: 10.1016/j.chb.2014.10.042 (cit. on pp. 32, 33, 54).

Nguyen, D., Doğruöz, A. S., Rosé, C. P. and Jong, F. de (2016). 'Computational Sociolinguistics: A Survey'. In: *Computational Linguistics* 42.3, pp. 537–593. ISSN: 04194217. DOI: 10.1162/COLI_a_00258. arXiv: 1508.07544 (cit. on p. 120).

Nguyen, H. T. (2008). 'Sequence organization as local and longitudinal achievement'. In: *Text & Talk - An Interdisciplinary Journal of Language, Discourse Communication Studies* 28.4, pp. 501–528. ISSN: 1860-7330. DOI: 10.1515/TEXT.2008.025 (cit. on p. 46).

Norman, D. A. (1988). 'The Psychology of Everyday Things'. In: Basic Books, p. 257. ISBN: 978-0-465-06709-1 (cit. on p. 136).

O'Hara, K., Massimi, M., Harper, R. H. R., Rubens, S. and Morris, J. (2014). 'Everyday Dwelling with WhatsApp'. In: *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. CSCW '14. New York, NY, USA: ACM, pp. 1131–1143. ISBN: 9781450325400. DOI: 10.1145/2531602.2531679 (cit. on p. 18).

O'Hara, K., Mitchell, A. S. and Vorbau, A. (Apr. 2007). 'Consuming Video on Mobile Devices'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '07. New York, NY, USA: ACM, p. 857. ISBN: 9781595935939. DOI: 10.1145/1240624.1240754 (cit. on p. 25).

Ofcom (2015). *The Communications Market 2015*. Tech. rep., p. 1. URL: https://www.ofcom.org.uk/%7B%5C_%7D%7B%5C_%7Ddata/assets/pdf%7B%5C_%7Dfile/0022/20668/cmr%7B%5C_%7Duk%7B%5C_%7D2015.pdf (cit. on p. 7).

Office for National Statistics (2019). *Envelopes out, smart speakers in as the ONS updates the inflation basket*. URL: https://www.ons.gov.uk/news/news/envelopesoutsmartspeakersinastheonsupdatestheinflationbasket (visited on 11/03/2019) (cit. on p. 7).

Okoshi, T., Nozaki, H., Nakazawa, J., Tokuda, H., Ramos, J. and Dey, A. K. (2016). 'Towards attention-aware adaptive notification on smart phones'. In: *Pervasive and Mobile Computing* 26, pp. 17–34. ISSN: 15741192. DOI: 10.1016/j.pmcj.2015.10.004 (cit. on p. 29).

Oksman, V. and Turtiainen, J. (June 2004). 'Mobile Communication as a Social Stage: Meanings of Mobile Communication in Everyday Life among Teenagers in Finland'. In: *New Media & Society* 6.3, pp. 319–339. DOI: 10.1177/1461444804042518 (cit. on p. 21).

Oldenburg, R. (Aug. 1989). *The Great Good Place: Cafes, Coffee Shops, Community Centers, General Stores, Bars, Hangouts, and How They Get You Through the Day*. 1. De Capo Press, p. 384 (cit. on pp. 2, 56, 89, 154, 195).

Olson, J. S., Olson, G. M., Storrøsten, M. and Carter, M. (Oct. 1993). 'Groupwork close up: a comparison of the group design process with and without a simple group editor'. In: *ACM Transactions on Information Systems* 11.4, pp. 321–348. ISSN: 10468188. DOI: 10.1145/159764.159763 (cit. on p. 24).

Oulasvirta, A., Rattenbury, T., Ma, L. and Raita, E. (June 2011). 'Habits make smartphone use more pervasive'. In: *Personal and Ubiquitous Computing* 16.1, pp. 105–114. DOI: 10.1007/s00779-011-0412-2 (cit. on p. 54).

Park, C., Lim, J., Kim, J., Lee, S.-J. and Lee, D. (2017). '"Don't Bother Me. I'm Socializing!": A Breakpoint-Based Smartphone Notification System'. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. CSCW '17. New York, NY, USA: ACM, pp. 541–554. ISBN: 9781450343350. DOI: 10.1145/2998181.2998189 (cit. on p. 54).

Payr, S. (2013). 'Virtual butlers and real people: styles and practices in long-term use of a companion'. In: *Your Virtual Butler: The Making-Of*. Ed. by R. Trappl. Berlin, Heidelberg: Springer-Verlag Berlin, pp. 134–178. DOI: 10.1007/978-3-642-37346-6_11 (cit. on p. 85).

Pelikan, H. R. M. and Broth, M. (2016). 'Why That Nao?: How Humans Adapt to a Conventional Humanoid Robot in Taking Turns-at-Talk'. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. New York, NY, USA: ACM, pp. 4921–4932. ISBN: 9781450333627. DOI: 10.1145/2858036.2858478 (cit. on pp. 36, 87).

Perry, M., O'Hara, K., Sellen, A., Brown, B. and Harper, R. H. R. (2001). 'Dealing with mobility: understanding access anytime, anywhere'. In: *ACM Transactions on Computer-Human Interaction*. TOCHI 8.4, pp. 323–347. ISSN: 10730516. DOI: 10.1145/504704.504707 (cit. on p. 21).

Peters, O. and Allouch, S. ben (Aug. 2005). 'Always connected: a longitudinal field study of mobile communication'. In: *Telematics and Informatics* 22.3, pp. 239–256. DOI: 10.1016/j.tele.2004.11.002 (cit. on pp. 20, 21).

Pielot, M., Dingler, T., San Pedro, J. and Oliver, N. (2015). 'When Attention is not Scarce - Detecting Boredom from Mobile Phone Usage'. In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp '15, pp. 825–836. ISBN: 9781450335744. DOI: 10.1145/2750858.2804252 (cit. on p. 54).

Pielot, M. and Rello, L. (2015). 'The Do Not Disturb Challenge: A Day Without Notifications'. In: *Extended Abstracts of the ACM CHI '15 Conference on Human Factors in Computing Systems*. Vol. 2. CHI EA '15. New York, NY, USA: ACM, pp. 1761–1766. ISBN: 9781450331463. DOI: 10.1145/2702613.2732704 (cit. on p. 20).

Pigram, J. J. (1996). 'Best practice environmental management and the tourism industry'. In: *Progress in Tourism and Hospitality Research* 2.3-4, pp. 261–271. DOI: 10.1002/pth.6070020307 (cit. on pp. 2, 156).

Pizza, S., Brown, B., McMillan, D. and Lampinen, A. (May 2016). 'Smartwatch in vivo'. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. New York, NY, USA: ACM, pp. 5456–5469. ISBN: 9781450333627. DOI: 10.1145/2858036.2858522 (cit. on pp. 34, 119, 155).

Pollner, M. (Mar. 2012). 'The End(s) of Ethnomethodology'. In: *The American Sociologist* 43.1, pp. 7–20. ISSN: 0003-1232. DOI: 10.1007/s12108-011-9144-z (cit. on p. 50).

Porcheron, M. (2015). 'Reflecting on the Study of Mobile Collocated Interactions'. In: *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. MobileHCI '15. New York, NY, USA: ACM, pp. 1163–1167. ISBN: 9781450336536. DOI: 10.1145/2786567.2794345. URL: http://dl.acm.org/citation.cfm?doid=2786567.2794345 (cit. on p. 194).

Porcheron, M., Fischer, J. E., McGregor, M., Brown, B., Luger, E., Candello, H. and O'Hara, K. (2017a). 'Talking with Conversational Agents in Collaborative Action'. In: *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. CSCW '17 Companion, pp. 431–436. ISBN: 9781450346887. DOI: 10.1145/3022198.3022666 (cit. on p. 118).

Porcheron, M., Fischer, J. E., Reeves, S. and Sharples, S. (2018). 'Voice Interfaces in Everyday Life'. In: *Proceedings of the 2018 ACM Conference on Human Factors in Computing Systems*. CHI '18. ACM, New York, NY, USA: ACM. DOI: 10.1145/3173574.3174214 (cit. on pp. 117, 159, 181, 182, 187, 204).

Porcheron, M., Fischer, J. E. and Sharples, S. C. (2016a). 'Using Mobile Phones in Pub Talk'. In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '16. New York, NY, USA: ACM, pp. 1647–1659. ISBN: 9781450335928. DOI: 10.1145/2818048.2820014 (cit. on pp. 53, 204).

Porcheron, M., Fischer, J. E. and Sharples, S. C. (2017b). '"Do Animals Have Accents?": Talking with Agents in Multi-Party Conversation'. In: *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '17. New York, NY, USA: ACM, pp. 207–219. DOI: 10.1145/2998181.2998298 (cit. on pp. 82, 182, 204).

Porcheron, M., Lucero, A. and Fischer, J. E. (2016b). 'Co-Curator: Designing for Mobile Ideation in Groups'. In: *Proceedings of the 20th International Academic Mindtrek Conference*. AcademicMindtrek '16. New York, NY, USA: ACM, pp. 226–234. ISBN: 9781450343671. DOI: 10.1145/2994310.2994350 (cit. on p. 27).

Poushter, J. (2016). *Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies*. Tech. rep. URL: http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/ (cit. on p. 16).

Raclaw, J. shua, Robles, J. S. and DiDomenico, S. M. (2016). 'Providing Epistemic Support for Assessments Through Mobile-Supported Sharing Activities'. In: *Research on Language and Social Interaction* 49.4, pp. 362–379. ISSN: 08351813. DOI: 10.1080/08351813.2016.1199089 (cit. on pp. 34, 162).

Raento, M., Oulasvirta, A., Petit, R. and Toivonen, H. (Apr. 2005). 'ContextPhone: A Prototyping Platform for Context-Aware Mobile Applications'. In: *IEEE Pervasive Computing* 4.2, pp. 51–59. ISSN: 1536-1268. DOI: 10.1109/MPRV.2005.29 (cit. on p. 28).

Redström, J. (Mar. 2006). 'Towards user design? On the shift from object to user as the subject of design'. In: *Design Studies* 27.2, pp. 123–139. DOI: 10.1016/j.destud.2005.06.001 (cit. on p. 27).

Reeves, S. (2011). *Designing Interfaces in Public Settings*. Human-Computer Interaction Series. London: Springer London, p. 409. ISBN: 978-0-85729-264-3. DOI: 10.1007/978-0-85729-265-0 (cit. on pp. 25, 42).

Reis, S., Church, K. and Oliver, N. (2012). 'Rethinking mobile search: Towards casual, shared, social mobile search experiences'. In: *Searching 4 Fun Workshop at the 34th European Conference on Advances in Information Retrieval*. ECIR '12 (cit. on p. 2).

Rigby, J. M., Brumby, D. P., Gould, S. J. J. and Cox, A. L. (2017). 'Media Multitasking at Home'. In: *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video*. TVX '17. New York, NY, USA: ACM, pp. 3–10. ISBN: 9781450345293. DOI: 10.1145/3077548.3077560 (cit. on p. 142).

Rimer, S. (2009). *Play With Your Food, Just Don't Text!* URL: https://www.nytimes.com/2009/05/27/dining/27text.html (visited on 10/03/2019) (cit. on p. 155).

Rogers, E. M. (1995). *Diffusion of Innovation*. 4th ed. The Free Press, p. 536. ISBN: 978-0029266717 (cit. on p. 6).

Rogers, Y. (2012). *HCI Theory: Classical, Modern, and Contemporary*. Vol. 5. 2, pp. 1–129. ISBN: 1946768019467699. DOI: 10.2200/S00418ED1V01Y201205HCI014 (cit. on p. 24).

Rooksby, J. (July 2013). 'Wild in the Laboratory: A Discussion of Plans and Situated Actions'. In: *ACM Transactions on Computer-Human Interaction* 20.3, pp. 1–17. ISSN: 10730516. DOI: 10.1145/2491500.2491507 (cit. on pp. 34, 35, 170).

Rooksby, J., Smith, T. E., Morrison, A., Rost, M. and Chalmers, M. (Jan. 2015). 'Configuring Attention in the Multiscreen Living Room'. In: *Proceedings of the 14th European Conference on Computer Supported Cooperative Work*. ECSCW '15 Chapter 13. Springer International Publishing, pp. 243–261. ISBN: 978-3-319-20498-7. DOI: 10.1007/978-3-319-20499-4_13 (cit. on pp. 4, 33, 34, 118, 141, 142, 154, 159, 162, 170, 172, 190, 194).

Sacks, H. (1992a). 'A single instance of a phone-call opening; Caller—Called, etc.' In: *Lectures on conversation*. Ed. by G. Jefferson. Vol. 2, Pt. 8, Oxford: Blackwell Publishing. Chap. 3, pp. 542–553. ISBN: 978-1-55786-705-6 (cit. on pp. 196, 197).

Sacks, H. (Jan. 1992b). *Harvey Sacks: Lectures on Conversation*. Ed. by G. Jefferson. Oxford: Blackwell Publishing, p. 818. ISBN: 978-1-55786-705-6 (cit. on p. 44).

Sacks, H., Schegloff, E. A. and Jefferson, G. (Jan. 1974). 'A Simplest Systematics for the Organization of Turn-Taking for Conversation'. In: *Language* 50.4, p. 696. ISSN: 00978507. DOI: 10.2307/412243. URL: https://www.jstor.org/stable/412243?origin=crossref (cit. on pp. 91, 197, 199).

Sadler, K., Robertson, T. and Kan, M. (2006). '"It's Always There, It's Always On": Australian Freelancer's Management of Availability Using Mobile Technologies'. In: *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*. MobileHCI '06. New York, NY, USA: ACM, p. 49. ISBN: 1595933905. DOI: 10.1145/1152215.1152226 (cit. on p. 21).

Salmela, T., Colley, A. and Häkki, J. (2019). 'Together in Bed?: Couples' Mobile Technology Use in Bed'. In: *CHI Conference on Human Factors in Computing Systems Proceedings*. CHI '19. New York, NY, USA: ACM. ISBN: 9781450359702. DOI: 10.1145/3290605.3300732 (cit. on p. 205).

Salovaara, A. (2007). 'Appropriation of a MMS-Based Comic Creator: From System Functionalities to Resources for Action'. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. CHI '07. New York, NY, USA: ACM, p. 1117. ISBN: 9781595935939. DOI: 10.1145/1240624.1240794 (cit. on p. 26).

Schaffer, S., Schleicher, R. and Möller, S. (2015). 'Modeling input modality choice in mobile graphical and speech interfaces'. In: *International Journal of Human Computer Studies* 75, pp. 21–34. ISSN: 10959300. DOI: 10.1016/j.ijhcs.2014.11.004 (cit. on p. 171).

Schane, R. E., Glantz, S. A. and Ling, P. M. (Aug. 2009). 'Social Smoking: Implications for Public Health, Clinical Practice, and Intervention Research'. In: *American Journal of Preventive Medicine* 37.2, pp. 124–131. ISSN: 0749-3797. DOI: 10.1016/J.AMEPRE.2009.03.020 (cit. on pp. 2, 156).

Schegloff, E. A. (1987). 'Analyzing Single Episodes of Interaction: An Exercise in Conversation Analysis'. In: *Social Psychology Quarterly* 50.2, pp. 101–114 (cit. on p. 42).

Schegloff, E. A. (2007). *Sequence Organization in Interaction*. Cambridge: Cambridge University Press, p. 318. ISBN: 9780511791208. DOI: 10.1017/CBO9780511791208 (cit. on p. 45).

Schilit, B., Adams, N. and Want, R. (1994). 'Context-aware computing applications'. In: *Workshop on Mobile Computing Systems and Applications*. IEEE Computer Society Press, pp. 85–90. ISBN: 0-8186-6345-6. DOI: 10.1109/MCSA.1994.512740 (cit. on p. 28).

Schirra, S., Sun, H. and Bentley, F. (2014). 'Together Alone: Motivations for Live-Tweeting a Television Series'. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. CHI '14. New York, NY, USA: ACM, pp. 2441–2450. ISBN: 9781450324731. DOI: 10.1145/2556288.2557070 (cit. on p. 118).

Schmidt, K. (Oct. 2011). 'The Concept of 'Work' in CSCW'. In: *Computer Supported Cooperative Work* 20.4-5, pp. 341–401. ISSN: 0925-9724. DOI: 10.1007/s10606-011-9146-y (cit. on p. 43).

Schulze, F. and Groh, G. (2016). 'Conversational context helps improve mobile notification management'. In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '16. New York, NY, USA: ACM, pp. 518–528. ISBN: 9781450344081. DOI: 10.1145/2935334.2935347 (cit. on p. 29).

Seddon, C. (2011). 'Lifestyles and Social Participation'. In: *Social Trends* 41.1, pp. 146–180. ISSN: 2040-1620. DOI: 10.1057/st.2011.7 (cit. on p. 58).

Siewiorek, D. P., Smailagic, A., Furukawa, J., Moraveji, N., Reiger, K. and Shaffer, J. (2003). 'SenSay: A context-aware mobile phone'. In: *Seventh IEEE International Symposium on Wearable Computers, 2003*. ISWC. IEEE, pp. 248–249. ISBN: 0-7695-2034-0. DOI: 10.1109/ISWC.2003.1241422 (cit. on p. 28).

Slack, R. (May 2000). ''Reflexivity or Sociological Practice: A Reply to May''. In: *Sociological Research Online* 5.1, pp. 1–5. ISSN: 1360-7804. DOI: 10.5153/sro.416 (cit. on p. 50).

Sohn, T., Li, K. A., Griswold, W. G. and Hollan, J. D. (Apr. 2008). 'A Diary Study of Mobile Information Needs'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08. New York, NY, USA: ACM, pp. 433–442. ISBN: 9781605580111. DOI: 10.1145/1357054.1357125 (cit. on pp. 4, 30).

St. Lawrence, J. S., Bradlyn, A. S. and Kelly, J. A. (Jan. 1983). 'Interpersonal Adjustment of a Homosexual Adult'. In: *Behavior Modification* 7.1, pp. 41–55. ISSN: 0145-4455. DOI: 10.1177/01454455830071003 (cit. on pp. 2, 156).

Su, N. M. and Wang, L. (Jan. 2015). 'From Third to Surveilled Place: The Mobile in Irish Pubs'. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. New York, NY, USA: ACM, pp. 1659–1668. ISBN: 9781450331456. DOI: 10.1145/2702123.2702574 (cit. on pp. 1, 4, 19, 54, 57, 154).

Suchman, L. (1985). *Plans and Situated Actions: The Problem of Human Machine Communication*. Cambridge: Cambridge University Press, p. 220. ISBN: 978-0521337397 (cit. on pp. 23, 31, 47, 168, 192, 201).

Suchman, L. (1997). 'Centers of Coordination: A Case and Some Themes'. In: *Discourse, Tools and Reasoning: Essays on situated cognition*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 41–62. DOI: 10.1007/978-3-662-03362-3_3 (cit. on pp. 31, 154).

Suchman, L. A. (2006). *Human–Machine Reconfigurations: Plans and Situated Actions*. Cambridge: Cambridge University Press, p. 314. ISBN: 9780511808418. DOI: 10.1017/CBO9780511808418 (cit. on p. 197).

Sutton, S., Foulkes, P., Kirk, D. and Lawson, S. (2019). 'Voice as a Design Material: Sociophonetic Inspired Design Strategies in Human-Computer Interaction'. In: *CHI Conference on Human Factors in Computing Systems Proceedings*. CHI '19. New York, NY, USA: ACM. DOI: 10.1145/3290605.3300833 (cit. on p. 205).

Tolmie, P. and Crabtree, A. (2008). 'Deploying Research Technology in the Home'. In: *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*. CSCW '08. New York, NY, USA: ACM, pp. 639–648. ISBN: 9781605580074 (ISBN). DOI: 10.1145/1460563.1460662 (cit. on pp. 118, 193).

Tolmie, P., Crabtree, A., Rodden, T. and Benford, S. (Nov. 2008). '"Are You Watching This Film or What?" Interruption and the Juggling of Cohorts'. In: *Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work*. CSCW '08. New York, NY, USA: ACM, p. 257. ISBN: 9781605580074. DOI: 10.1145/1460563.1460605 (cit. on pp. 142, 165).

Törrönen, J. and Maunu, A. (Feb. 2005). 'Going out, sociability, and cultural distinctions'. In: *Nordic Studies on Alcohol and Drugs* 22.1_suppl, pp. 25–43. ISSN: 1455-0725. DOI: 10.1177/14550725050220 1S07 (cit. on p. 53).

Turkle, S. (Jan. 2011). *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books. ISBN: 978-0465010219 (cit. on pp. 1, 4, 14, 17–20, 26, 55, 83, 141, 155, 166, 202).

Turkle, S. (2015). *Reclaiming Conversation: The Power of Talk in a Digital Age*. Penguin, p. 448. ISBN: 978-1-59429-555-2 (cit. on pp. 18, 55, 190).

Vanden Abeele, M. M. P., Antheunis, M. L. and Schouten, A. P. (2016). 'The effect of mobile messaging during a conversation on impression formation and interaction quality'. In: *Computers in Human Behavior* 62, pp. 562–569. ISSN: 07475632. DOI: 10.1016/j.chb.2016.04.005 (cit. on p. 54).

Vardoulakis, L. P., Ring, L., Barry, B., Sidner, C. L. and Bickmore, T. (2012). 'Designing Relational Agents as Long Term Social Companions for Older Adults'. In: *Intelligent Virtual Agents*, pp. 289–302. DOI: 10.1007/978-3-642-33197-8_30 (cit. on p. 85).

Watts, J. C., Woods, D. D., Corban, J. M., Patterson, E. S., Kerr, R. L. and Hicks, L. C. (1996). 'Voice loops as cooperative aids in space shuttle mission control'. In: *Proceedings of the 1996 ACM conference on Computer supported cooperative work*. CSCW '96. New York, NY, USA: ACM, pp. 48–56. ISBN: 0897917650. DOI: 10.1145/240080.240188 (cit. on p. 31).

Wei, R. and Leung, L. (Feb. 1999). 'Blurring public and private behaviors in public space: policy challenges in the use and improper use of the cell phone'. In: *Telematics and Informatics* 16.1-2, pp. 11–26. DOI: 10.1016/S0736-5853(99)00016-7 (cit. on p. 22).

Wei, R. and Lo, V.-H. (Feb. 2006). 'Staying connected while on the move: Cell phone use and social connectedness'. In: *New Media & Society* 8.1, pp. 53–72. ISSN: 1461-4448. DOI: 10.1177/1461444806059870 (cit. on pp. 4, 22).

Weilenmann, A. and Larsson, C. (Jan. 2002). 'Local use and sharing of mobile phones'. In: *Wireless World*. Chapter 7. London: Springer, pp. 92–107. ISBN: 978-1-85233-477-2. DOI: 10.1007/978-1-4471-0665-4_7 (cit. on pp. 32, 154).

Weiser, M. (1991). *The Computer for the 21st Century*. DOI: 10.1038/scientificamerican0991-94 (cit. on pp. 16, 27, 28).

Wooffitt, R. (1994). 'Applying Sociology: Conversation Analysis in the Study of Human-(Simulated) Computer Interaction'. In: *Bulletin de Méthodologie Sociologique* 43.1, pp. 7–33. ISSN: 0759-1063. DOI: 10.1177/075910639404300103 (cit. on pp. 36, 197).

Wooffitt, R., Fraser, N. M., Gilbert, G. N. and McGlashan, S. (1997). *Humans, Computers and Wizards*. 1st ed. London, UK: Routledge, p. 207. ISBN: 0-415-06948-3 (cit. on p. 36).

Yamada, M. (1981). 'Invisibility is an unnatural disaster'. In: *This Bridge Called My Back: Writings by Radical Women of Color*. Ed. by C. Moraga and G. E. Anzaldúa. Persephone Press, pp. 30–35. DOI: 978-0-930436-10-0 (cit. on pp. 2, 156).

Zimmerman, J., Forlizzi, J. and Evenson, S. (2007). 'Research Through Design as a Method for Interaction Design Research in HCI Design Research in HCI'. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. CHI '07. New York, NY, USA: ACM, p. 493. ISBN: 9781595935939. DOI: 10.1145/1240624.1240704 (cit. on p. 206).

Zue, V., Seneff, S., Glass, J. R., Polifroni, J., Pao, C., Hazen, T. J. and Hetherington, L. (2000). 'JUPlTER: a telephone-based conversational interface for weather information'. In: *IEEE Transactions on Speech and Audio Processing* 8.1, pp. 85–96. ISSN: 10636676. DOI: 10.1109/89.817460 (cit. on p. 86).

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Specific modifications were made to the template by Martin Porcheron to correctly format fragments of talk and action:

https://github.com/mporcheron/latex-emcafragments

Submitted Version as of 2019-08-09 (ee3e548).


University of
Nottingham
UK | CHINA | MALAYSIA