A special interest group on developing theories of language use in interaction with conversational user interfaces

Citation for published version:
https://doi.org/10.1145/3544549.3583179

Digital Object Identifier (DOI):
10.1145/3544549.3583179

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
CHI 2023 - Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
A Special Interest Group on Developing Theories of Language Use in Interaction with Conversational User Interfaces

Paola Raquel Peña Huerta
paola.pena@ucdconnect.ie
University College Dublin
Dublin 4, Ireland

Philip R. Doyle
philip.doyle1@ucdconnect.ie
University College Dublin
Dublin 4, Ireland

Emily Y.J. Ip
ipem@tcd.ie
Trinity College Dublin
Dublin, Ireland

Giovanni M. Di Liberto
gdiliber@tcd.ie
Trinity College Dublin
Dublin, Ireland

Darragh Higgins
higgind3@tcd.ie
Trinity College Dublin
Dublin, Ireland

Rachel McDonnell
RAMCDONN@tcd.ie
Trinity College Dublin
Dublin, Ireland

Holly Branigan
Holly.Branigan@ed.ac.uk
University of Edinburgh
Edinburgh, United Kingdom

Joakim Gustafson
jkgu@kth.se
KTH Royal Institute of Technology in Stockholm
Stockholm, Sweden

Donald McMillan
donald.mcmillan@dsv.su.se
Stockholm University
Stockholm, Sweden

Robert J. Moore
rjmoore@us.ibm.com
IBM Research-Almaden Lab
San Jose, United States of America

Benjamin R. Cowan
bejamin.cowan@ucd.ie
University College Dublin
Dublin 4, Ireland

ABSTRACT

CCS CONCEPTS
• Human-centered computing → HCI theory, concepts and models; Interaction design theory, concepts and paradigms;
• Applied computing → Psychology.

KEYWORDS
conversational user interfaces, human-machine dialogue, speech agents, psycholinguistic models

ACM Reference Format:

1 INTRODUCTION

Through the increased growth of speech agents, text based chatbots and social robots, language interactions with machine dialogue partners are now commonplace. Discovering what drives the way we converse with machines is fundamental to understanding our interaction with such automated dialogue partners. However, understanding of what governs user language choices in such human-machine dialogues (HMD) is sparse [14]. This Special Interest Group aims to be a catalyst for discussing and building fundamental theories of how people produce language when engaged in conversation with conversational user interfaces (CUIs). The main objective is to bring together researchers across CHI and related communities (e.g. HRI, CUI, cognitive science, linguistics and speech technology) to map the grand challenges required to be addressed to generate evidence-based theories to explain what impacts our linguistic interactions with CUIs.

2 NEED FOR A SIG

Theory work on language production in CUI interaction is scarce [15]. Current approaches rely on applying existing theoretical accounts of human-human dialogue (HHD) and language use from disciplines such as psycholinguistics and sociolinguistics to help understand mechanisms that lead users to adopt specific language choices and behaviours. Using these accounts as a foundation is useful as it gives us initial methods and concepts that can be applied to CUI interaction. Recent efforts to understand user language production in CUI have taken this approach, directly applying methods and concepts from such disciplines [7, 18, 24]. Yet such methods and concepts also need to be tailored so as to be sensitive to the fundamental differences between human and machine dialogue interaction in terms of partner capabilities, the nature and aims of the dialogue [15, 45].

This SIG will outline 1) the existing approaches and theoretical accounts currently adopted from social science disciplines as well as devise 2) a roadmap of work that needs to be undertaken to build CUI centered theories, sensitive to the nuances of HMD and CUI
interaction. In the following sections, we outline some example theoretical and methodological frameworks employed in HHD that have been influential in recent CUI interaction research. These are by no means exhaustive. We propose to focus on these frameworks as a basis for our discussion, mapping out how these approaches (and other identified during the SIG) can be applied to generate more cohesive theoretical frameworks focused on linguistic interactions with CUIs more specifically.

3 EXAMPLES OF RELEVANT THEORETICAL AND METHODOLOGICAL FRAMEWORKS

3.1 Audience Design, Common Ground and Partner modelling

Research in HHD proposes that people engage in audience design [4] when speakers adapt their utterances based on perceptions of their addressees’ knowledge, dialogue capabilities and beliefs [11, 12]. The concept of common ground is critical to facilitating audience design, in that interlocutors use contextual information, information or descriptions co-constructed during dialogue along with presumed knowledge of their partner to plan utterances so as to optimally design them for their audience [10]. Other HHD accounts such as the monitor and adjust model [3] state that, rather than audience design being the sole driver of language choices, speakers also rely on their own knowledge for utterance construction, only using their partner’s perspective when absolutely necessary [26, 27, 47]. Both of these have been applied in CUI research. A significant amount of work has demonstrated that audience design (also termed recipient design [2]) occurs in HMD, whereby language choices are adaptive and considered based on preconceptions of a machine partner’s abilities and knowledge [5, 8, 29, 32, 34]. Machines tend to be seen as “at risk” listeners [42] in conversation, which leads people to adapt language choices to ensure that communication is successful [7] echoing HHD audience design accounts [4, 10, 13]. Indeed CUI design is also likely important in this regard, with aspects such as agent accent or perceived nationality [18, 28, 46] may act as a cue for informing people’s perceptions of partner capabilities (recently termed partner models [24] in HMD research) that guide language use. Yet, recent studies have questioned the ubiquity of such an account, either finding no audience design effects [16–18] or observing more egocentric production [23].

3.2 Computers are Social Actors

The Computers Are Social Actors (CASA) paradigm asserts that people mindlessly apply social heuristics from human-human interaction in human-computer interaction, leading them to respond to computer systems as they respond to other people. This assertion is founded on studies that showed users applying gender stereotypes to male- and female-voiced computers [39], using social categorisation to inform behaviours and judgements of computers in collaborative tasks [36], being polite to a computer when asked to evaluate it [38], being more helpful to computers that were perceived as more helpful [25], and applying expertise judgements to televisions [37]. The earliest experiments were on computer systems with voice output [40], but this was later generalised to other types of interfaces and devices. This account has been highly influential in the field of HCI, and gives an important potential mechanism which may govern the construction of initial partner models of CUIs as dialogue partners. That is, users may use cues from the interaction to inform the use of social heuristics when in dialogue with CUIs.

3.3 Neurophysiological approaches of CUI Interaction

The study of human communication in neurophysiology entails probing underlying brain mechanisms for metrics that can validate existing psycholinguistic models. Although initially focusing on simple speech listening tasks far from dialogue scenarios (e.g., listening to isolated syllables), the field has undergone a rapid development of naturalistic methodologies in research [20, 22, 31], propelled by the discovery that neural signals track acoustic and linguistic features of a speech input [19]. This robust relationship between specific input features and corresponding neural signal [41] is thought to open up new potential to study the processing of various speech constituents simultaneously in ecologically-valid settings. Such naturalistic paradigms support predictive processing, a long-standing theoretical framework [9, 33] whereby sensory information coupled with the brain’s predictions of upcoming sensory events highlights the “active” nature of perception. These neural architectures can elucidate how humans contextualise speech sounds in dialogue scenarios, informing both HMD and HHD. Indeed new hyperscanning technology has brought us closer to researching more realistic dialogue scenarios enabling the simultaneous recording of EEG signals from multiple participants. This line of work has already demonstrated the possibility of measuring the impact of social cues (e.g., reflecting attention) on neural synchrony across participants [6, 21]. Although dialogue presents challenges for neurophysiological study, extracting objective indices of brain activity during dialogue in interactive scenarios [44] may constitute promising foundations for new insights into CUI based dialogue interaction.

3.4 Naturalistic observation of CUI Language Use

Along with controlled experiment approaches to understanding language production in CUI interaction (e.g. [7, 18] work on understanding language production also uses more naturalistic approaches through the statistical and conversation analysis of real world human-machine dialogues [30, 45]. Conversation analysis in particular has gained significant traction as a practice in CUI research [1, 35, 43, 45], being used to identify and interpret people’s language choices in CUI interaction. Rather than focusing on the frequency of language-based phenomena, as is more common in statistical and controlled experiment approaches, CA gives a rich, in-depth exploration of linguistic effects and behaviours in dialogue. Such approaches and findings, in tandem with controlled and statistical approaches are likely to be integral to the development of fundamental theories of language production in CUI contexts.
4 OPEN QUESTIONS

The SIG will aim to use the outlined theories and methodological approaches as a starting point to discuss how as a community we can develop models and accounts that interpret and explain the mechanisms within language production in CUI interaction. The questions below have been identified as key ones to address as part of the SIG:

- What theoretical lenses are the most appropriate to anchor attempts to devise theoretical insight to language use in CUI interaction?
- What are the key methodological approaches and challenges in ensuring the viability and robustness of theoretical frameworks for language production in CUI interaction and how can these work together?
- What concepts are critical to devising theories in this area?
- What aspects about CUI interaction need to be considered when adapting theories and concepts from other disciplines?
- What are some of the key unknowns around these concepts that need significant work?

REFERENCES


Received 15 December 2022