

Understanding Dialogue: Language Use and Social Interaction

Martin J. Pickering and Simon Garrod

(University of Edinburgh & University of Glasgow)

Cambridge, England: Cambridge University Press, 2021, xvii+283 pp; hardback, ISBN 978-1-108-47361-3; paperback, ISBN 978-1-108-46193-1

Reviewed by

Rachel Bawden

Inria, France

Understanding Dialogue: Language Use and Social Interaction represents a departure from classic theories in psycholinguistics and cognitive sciences; instead of taking as a starting point the isolated speech of an individual that can be extended to accommodate dialogue, a primary focus is put on developing a model adapted to dialogue itself, bearing in mind important aspects of dialogue as an activity with a heavily cooperative component. As a researcher of natural language processing with a background in linguistics, I find highly intriguing the possibilities provided by the dialogue model presented. Although the book does not itself touch upon the potential for automated dialogue, I am inevitably writing this review from the point of view of a computational linguist with these aspects in mind.

Building on numerous previous works, including many of the authors' own studies and theories, *Understanding Dialogue* presents the **shared workspace framework**, a framework for understanding not just dialogue but cooperative activities in general, of which dialogue is viewed as a subtype. Based on Bratman's (1992) concept of shared cooperative activity, the framework provides a joint environment with which interlocutors can interact, both by contributing to the space (with actions or utterances for example), and by perceiving and processing their own or the other participants' productions. The authors do not limit their work to linguistic communication: Many of their examples, particularly at the beginning of the book, are non-linguistic (e.g., hand shaking, dancing a tango, playing singles tennis); others are primarily physical, but will most likely also involve linguistic communication (such as jointly constructing flat-pack furniture); and others are purely linguistic (e.g., suggesting which restaurant to go to for lunch).

The notion of *alignment* is highly important to this framework both from a linguistic and non-linguistic perspective, and is one of the main inspirations of the book, having previously been presented in *Toward a Mechanistic Theory of Dialogue* by the same authors. As individuals interact via the joint space, alignment concerns the equivalence in their representations at a conceptual level, with respect to their goals and relevant props in the shared environment (*dialogue model alignment*) and linguistic representations shared in the workspace (*linguistic alignment*). Roughly speaking, in this second (linguistic) case, this may for instance correspond to whether or not the individuals have the same representation of the utterance in terms of phonetics (were the sounds perceived correctly?) or in terms of lexical semantics (do they understand the same reference by the word uttered?). From here can be explained a number of different dialogue

<https://doi.org/10.1162/coli.r.00411>

© 2021 Association for Computational Linguistics

Published under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license

behaviors linked to the quest for alignment and the resolution of misalignment should it occur.

The book is structured in four main parts, preceded by an Introduction presenting the challenges of dialogue and the main ideas behind the framework. The focus of the book is clearly stated from the beginning as being dialogue first, in a rejection of models that seek to study language primarily from a monologic point of view. As the authors point out, the notion of alignment underpinning the framework involves by its very nature multiple participants and therefore dialogic interactions must be studied in their own right. I shall provide only a brief summary of the four parts here, highlighting some components that in my view are key to the model, without however covering all themes, which would require a far more extensive description.

Part I introduces the basis of the *shared workspace framework* as applied to activities with a cooperative component and then specifically to dialogue. The basic *sender-receiver framework* is quickly rejected, as it lacks the ability to represent certain key elements of cooperative activities, such as allowing for feedback and representing an environment that is common to the participants. The shared workspace framework is then introduced, along with the four important characteristics of cooperative joint aspect systems that can be successfully illustrated with it: alignment (mentioned above), simulation (the representation of an activity without actually going through with it), prediction (the anticipation of participants' behaviors), and synchrony (concerning the timing of behaviors in a joint activity), elements that are first studied in the context of joint activities in general (Chapter 3), before being reviewed specifically for dialogue (Chapter 4).

Part II is dedicated to the aforementioned concept of alignment, fundamental to the framework of cooperative activity. The chapters in this section look at the distinction between the different levels at which alignment can occur, the processes involved, and the consequences of alignment, such as participants uttering similar linguistic productions. Another important notion introduced in this part is that of the meta-representation of alignment, which represents the participants' belief about how aligned they are, which has inevitable consequences on how they then plan and implement their contributions.

Part III continues with the theme of alignment but turns to aspects involving the efficiency of communication: succinctness of formulation (Chapter 8) and how we time our contributions (Chapter 9). Particularly interesting is the role of commentaries, which are contributions providing some sort of feedback on the alignment of participants and which can therefore affect the participants' meta-representation of alignment. There is an important distinction between positive and negative commentaries, positive commentaries (such as "uh huh" in English) providing feedback that the speaker is aligned, therefore enabling the participants to meta-represent alignment, and negative ones (such as "huh?") indicating a misalignment, but then enabling the participants to recover from that it. These commentaries contribute to the succinctness of dialogue and to maximizing the efficiency of joint participation by indicating meta-alignment. Finally, Chapter 9 discusses the notion of "speaking in good time," related to the necessarily sequential nature of dialogue and the importance of timing, including the effects of different speech rates and the natural adaptation that occurs between interlocutors.

Part IV looks beyond the main theme of dialogue to other forms of conversation, including multiparty conversations and collectives, exploring the possible roles of the different participants, and how this relates back to alignment and their contribution to the shared workspace. Also mentioned is monologue and the challenges that it poses with respect to the primary and more natural form of language communication that is dialogue. The final chapter introduces how the shared workspace can be augmented

by adding props, illustrations, and recordings and by using alternative communicative tools, such as text messages and social media, which come with their own constraints with respect to the access they allow to the shared workspace.

The description of the framework is thorough and well exemplified, with a continuity in the use of examples throughout the book. A repetition and embellishment of schemas helps to keep track of how the new additions from each chapter fit into the framework. I found some of the descriptions a little wordy, particularly because of the reiteration of definitions and motivations, and in the minutely detailed illustration of examples. However, from the point of view of pedagogy, this could be seen as adding clarity, particularly for the reader who decides to focus on particular chapters rather than reading the book from cover to cover. In my opinion, the book will be highly accessible to all readers, even those who have limited background on the topic, and the authors take care to make it clear how their framework and definitions agree with or differ from previous works.

For me, there remain two main areas that could have been worthy of further exploration within the scope of this book. The first is the effect of cultural and linguistic differences. The authors do address the topic in Chapter 11, but in comparison with the detail afforded to the description of the framework, this subject remains rather lacking, with only a short section touching on it, under the title of *Social Norms and Joint Planning*. The authors cite an interesting study by Fujii (2012) on the differences between American and Japanese speakers in terms of their use of language to foster alignment. However, this teaser does not lead on to a deeper discussion about cross-cultural differences as explained in terms of the concepts used in this framework. The second topic is the link to sign languages, which would appear to link more than perfectly with the shared workspace framework and yet is not mentioned by the authors.

There is little doubt that the framework is an important step in modeling dialogue from a psycholinguistic perspective. As a researcher in natural language processing, I would be excited to see the possibilities for this framework in a computational setting for automated dialogue, something that the authors mention in their conclusion. They evoke the failure of current chatbots such as Siri and Alexa to effectively dialogue due to their inability to provide commentary (e.g., in the context of an ambiguous question) and to meta-represent alignment (i.e., have an opinion on whether the representations of the dialogue participants are the same). They suggest that this framework could help provide the solution to the current disruptions in communication we meet when interacting with these systems. I therefore look forward to seeing what progress can be made from this point of view.

References

Bratman, M. E. 1992. Shared cooperative activity. *Philosophical Review*, 101(2):327–340.

Fujii, Y. 2012. Differences of situating Self in the place/ba of interaction between the Japanese and American English speakers. *Journal of Pragmatics*, 44:636–662.

Rachel Bawden is a researcher in natural language processing and a specialist in Machine Translation at Inria, Paris, France. She previously received a PhD in the Machine Translation of dialogue from the LIMSI laboratory in Orsay, France, and spent two years as a postdoctoral researcher at the University of Edinburgh working on the Machine Translation of low-resource languages. Her interest in dialogue stems from her PhD work but also from her bachelor's degree in linguistics at the University of Oxford. Her e-mail address is rachel.bawden@inria.fr.

