

## Understanding Dialogue

Linguistic interaction between two people is the fundamental form of communication, yet almost all research in language use focuses on isolated speakers and listeners. In this innovative work, Pickering and Garrod extend the scope of psycholinguistics beyond individuals by introducing communication as a social activity. Drawing on psychological, linguistic, philosophical and sociological research, they expand their theory that alignment across individuals is the basis of communication, through the model of a ‘shared workspace account’. In this workspace, interlocutors are actors who jointly manipulate and control the interaction and develop similar representations of both language and social context, in order to achieve communicative success. The book also explores dialogue within groups, technologies and the role of culture more generally. Providing a new understanding of cognitive representation, this trailblazing work will be highly influential in the fields of linguistics, psychology and cognitive science.

MARTIN PICKERING is Professor of the Psychology of Language and Communication at the University of Edinburgh. He served as editor of the *Journal of Memory and Language*, was recipient of the Experimental Psychology Society mid-career award, and is a Fellow of the Royal Society of Edinburgh.

SIMON GARROD is Professor of Cognitive Psychology at the University of Glasgow. He was awarded the Distinguished Scientific Contribution Award by the Society for Text and Discourse. Between 1989 and 1999 he was also Deputy Director of the ESRC Human Communication Research Centre. He is a Fellow of the Royal Society of Edinburgh.

Cambridge University Press  
978-1-108-47361-3 — Understanding Dialogue  
Martin J. Pickering , Simon Garrod  
Frontmatter  
[More Information](#)

---

Cambridge University Press  
978-1-108-47361-3 — Understanding Dialogue  
Martin J. Pickering, Simon Garrod  
Frontmatter  
[More Information](#)

---

# Understanding Dialogue

*Language Use and Social Interaction*

---

Martin J. Pickering

*University of Edinburgh*

Simon Garrod

*University of Glasgow*



**CAMBRIDGE**  
UNIVERSITY PRESS

Cambridge University Press  
978-1-108-47361-3 — Understanding Dialogue  
Martin J. Pickering, Simon Garrod  
Frontmatter  
[More Information](#)

## CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom  
One Liberty Plaza, 20th Floor, New York, NY 10006, USA  
477 Williamstown Road, Port Melbourne, VIC 3207, Australia  
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,  
New Delhi – 110025, India  
79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

[www.cambridge.org](http://www.cambridge.org)  
Information on this title: [www.cambridge.org/9781108473613](http://www.cambridge.org/9781108473613)  
DOI: 10.1017/9781108610728

© Martin J. Pickering and Simon Garrod 2021

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2021

*A catalogue record for this publication is available from the British Library.*

ISBN 978-1-108-47361-3 Hardback  
ISBN 978-1-108-46193-1 Paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

## Contents

---

<i>List of Figures</i>	<i>page</i> vii
<i>List of Tables</i>	ix
<i>Glossary</i>	x
<i>Preface</i>	xv
1 The Challenge of Dialogue	1
<b>Part I The Shared Workspace Framework</b>	
2 A Systems Analysis of Joint Activity	11
3 Executing, Understanding and Controlling Joint Activity	44
4 Dialogue as a Joint Activity	67
5 Producing, Controlling and Understanding Dialogue	89
<b>Part II Alignment during Dialogue</b>	
6 Interactive Alignment and Linguistic Representations	109
7 Alignment of Dialogue Models	134
<b>Part III Using the Shared Workspace Efficiently</b>	
8 Saying Just Enough	169
9 Speaking in Good Time	190

vi Contents

**Part IV Extending the Shared Workspace Framework**

10	Communication beyond the Minimal Dyad	211
11	Culture and Language Use	239
12	Conclusion	261
	<i>References</i>	266
	<i>Index</i>	279

## Figures

---

1.1 The shared workspace framework for cooperative joint activities.	<i>page 5</i>
2.1 Two people constructing a piece of flat-pack furniture.	12
2.2 A simple model of a closed dyadic system for cognitive science.	20
2.3 Layered multi-robot architecture.	27
2.4 The shared workspace model of cooperative joint activity and distributed control.	29
3.1 The shared workspace framework for joint action.	45
3.2 A model of action control.	47
3.3 A model of the system for action prediction by simulation.	54
3.4 A model of one individual predicting own and other's action.	58
3.5 A model of the joint action control system.	61
4.1 The shared workspace framework for dialogue.	78
5.1 The shared workspace framework for dialogue.	90
5.2 A model for controlling language production.	92
5.3 A model of the system for utterance prediction via online simulation.	98
5.4 A model of one interlocutor predicting dialogue.	101
5.5 A model of the dialogue control system.	103
6.1 Two dimensions of alignment.	110
6.2 The shared workspace framework for dialogue.	111
6.3 The dialogue planner, containing the situation model and the game model.	113
6.4 Linguistic representations used by the implementer.	121
6.5 Channels of alignment between interlocutors' representations.	125
7.1 The linguistic representation underlying dialogue routines for MARY37 in the event described by <i>John drove Mary to London</i> .	143
7.2 The "ice skater" tangram.	145
7.i The duck-rabbit	150
7.3 Referential alignment and the workspace.	153

viii	List of Figures	
7.4	Referential misalignment and the workspace.	154
7.5	Misheard speech and the workspace.	155
7.6	The director's and addressee's views of the opposite sides of the lattice when the director describes the bottom right candle (from the director's point of view) to the addressee who has a privileged view of an even smaller candle.	157
7.7	An illustration of a critical trial from Keysar et al. (2000) in terms of the workspace.	158
8.1	A model of the dialogue system.	170
8.2	A sequence of descriptions of Figure 7.2.	182
9.1	A model of the dialogue system.	195
9.2	A model of the dialogue system that incorporates timing.	203
9.3	Neural mechanisms of speech rate convergence.	206
10.1	The shared workspace framework for triadic dialogue about the maze game.	212
10.2	The shared workspace framework for a dialogue between an individual ( <i>A</i> ) and a collective of two individuals ( <i>B</i> and <i>C</i> ).	223
10.3	Conversational roles in multi-party dialogue.	229
10.4	The shared workspace framework for monologue (with <i>A</i> as designer and <i>B</i> as audience), when <i>A</i> is giving a radio news broadcast and <i>B</i> is listening.	232
11.1	A state transition network for conversation.	246
11.2	Bruegel the Elder's <i>Netherlandish Proverbs</i> .	250
11.3	A photograph of an arctic tern (with focus on its subject).	251
11.4	The shared workspace augmented by an illustration.	252



## Tables

---

2.1 Examples of cooperative joint activities.	<i>page</i> 13
2.2 Properties of different cooperative joint activities.	15
4.1 An example conversation.	70
6.1 Examples of dialogue games from Levin and Moore.	116

## Glossary

---

**Alignment.** Two individuals are aligned at a given level of representation to the extent that they have the same representation as each other. Alignment has two dimensions (Figure 6.1). First, linguistic alignment involves alignment of linguistic representations, and dialogue model alignment involves alignment of situation models and dialogue game models. Second, focal alignment is short-term and relates to individual steps in the process of alignment, and global alignment is long-term and takes place over the dialogue as a whole (see Section 6.1).

**Collective.** In a multi-party dialogue, a collective is a group of individuals who play a single conversational role (see Figure 10.2).

**Commentary.** A commentary is a contribution that has a commentary function. The commentary function is to indicate meta-representations of alignment or misalignment in order to manage subsequent contributions.

A **positive commentary** indicates high confidence in alignment – that is, that the addressee meta-represents alignment. So when the speaker utters *I met John yesterday*, the commentary *uh-huh* indicates that the addressee believes he has the same representation as the speaker, for example with respect to the reference of *John*. The speaker therefore can continue as planned.

A **negative commentary** indicates low confidence in alignment – that is, that the addressee meta-represents misalignment. The commentary *eh?* indicates that the addressee believes that he does not have the same representation as the speaker. In this case, the speaker has to identify the likely misalignment and attempt to repair.

**Comparator.** A comparator compares a prediction of an action or associated internal state with an actual action or associated internal state and feeds back the comparison to the system that can produce a subsequent action. In our terminology, a monitor is a comparator for language.

**Control.** There are several types of control of action (including language) relevant to our account.

**Closed-loop control** is when feedback from the outcome of an action is used to correct that action.

**Distributed control** is when all agents in a cooperative joint activity exert some control over the activity.

**Individual control** is when a single agent exerts control over the agent's own actions.

**Pseudo-closed-loop control** is when the discrepancy between the predicted and actual action is used to modify the action accordingly. It occurs when an agent uses a forward model in monitoring and repair.

**Conversational roles.** These are the roles that can be distinguished in multi-party dialogue (see Figure 10.3). Dialogue games involve two **players**: the **initiator** (who performs the first move) and the **respondent** (who can complete the game or become the initiator for an embedded game). The **active audience** can play subsequent games but cannot complete the current game. The players and the active audience are the **contributors** – they contribute to the success of the dialogue. The **passive audience** cannot contribute to the dialogue but is taken into account by the players. **Overhearers** also cannot contribute and are not taken into account by the players.

**Cooperative joint activity.** A kind of joint activity that involves cooperation. See Section 2.1 for extensive discussion.

**Depiction** is using iconic signs in communication. Such signs may be gestures or sounds that mimic actions and other events.

**Designer.** The producer of a monologue (such as the writer). The **audience** of a monologue plays an equivalent role to the **passive audience** in a dialogue (see conversational roles).

**Dialogue game** is how interlocutors interact to achieve their joint goal in the dialogue. For example, an information-seeking game is when one interlocutor interacts with her partner to elicit information and the partner responds accordingly. One interlocutor carries out the initiation move in the game and the other carries out the response move. Initiation moves in dialogue games sometimes correspond to individual speech acts (e.g. the initiation move in an information-seeking game will often correspond to an interrogative speech act).

**Dialogue game model** is the part of the dialogue planner which enables an interlocutor to keep track of where both interlocutors are in a dialogue game.

**Dialogue routine** is when interlocutors use the same linguistic expression to refer to the same thing in a particular dialogue (though sometimes the routine will of course persist). See Section 7.1.4.

**Forward modelling.**

**Forward models** are mappings from an action command to its predicted outcomes. They have two forms: (1) the forward action model that maps the

action command to the representations that lead to the action, and (2) the forward perception model that maps from the representations that lead to the action to the perceptual representation of the action.

**Inverse models** are the converse mappings, from the perceptual representation to the representations that lead to the action, and from those representations to the action command. In motor control systems, forward and inverse models are typically paired.

**An efferece copy** is a copy of the action command that is used for modelling the action. It provides the input to the forward models.

**Implementer.** The mechanisms that produces actions and interprets them.

**The joint action implementer** does so as part of cooperative joint activities.

**The dialogue implementer** does so as part of dialogue. It makes use of the mechanisms involved in processing language and refers to representations of semantics, syntax and phonology.

**Individual activity.** Any activity performed by one individual.

**Joint activity.** Any activity performed by two (or more) individuals.

**Joint affordance.** A joint affordance is the behaviour in a cooperative joint activity that is enabled by an object or event in that activity. For example, the instrumentation in a large plane jointly affords flying by the pilot and co-pilot, and the sign *John* jointly affords interpretation as part of a dialogue (e.g. the interlocutors treat it as a reference to a particular John).

**Joint attention.** *A* and *B* jointly attend to an entity if they both attend to it (i.e. it is in parallel attention) and the entity is manifest.

**Joint intention** is when both actors of a cooperative joint activity intend to do something together even when each does not intend to do it alone.

**Joint manipulation** occurs when both actors manipulate an object as part of a cooperative joint activity. *A* and *B* jointly manipulate a buffet if they help each other to sample it; they do not do so if they simply select their own dishes.

**Manifest.** In a cooperative joint activity involving *A* and *B*, an entity is manifest if *A* is confident that both *A* and *B* attend to it and *B* is confident that *A* and *B* attend to it. Entities in the shared workspace are typically (but not necessarily) manifest.

**Meta-representation of alignment.** An interlocutor meta-represents alignment of *X* if she is confident that she has the same representation of *X* as her partner. When this is the case, she **m-tags** her representation of *X*.

**Minimal dyadic conversation.** Two interlocutors having a face-to-face, informal conversation.

**Monadic cognitive science.** Cognitive science in which the constructs are defined with respect to an individual person's mind.

**Monitor.** A monitor compares a prediction of an utterance or associated internal state (e.g. its semantics, syntax, or phonology) with an actual utterance or associated internal state and feeds back the comparison to the language production system.

**Self-monitoring** occurs when an individual monitors her own utterance or associated internal state. (It can use either closed-loop or pseudo-closed-loop control.)

**Other-monitoring** occurs when an individual monitors her partner's utterance. (It uses closed-loop control.)

**Joint monitoring** occurs when an individual monitors a joint utterance – that is, a combination of her own and her partner's contributions to a dialogue. (It uses closed-loop control.)

**Monologue** is a form of communication in which the communicator cannot receive any response from the audience. It involves two roles: the **designer** and the **audience**.

**Parallel attention.** *A* and *B* attend to an entity in parallel if they both attend to it.

**Planner.** The mechanisms that plans actions and interprets those plans.

**The joint action planner** does so as part of cooperative joint activities.

**The dialogue planner** does so as part of dialogue. It makes use of the mechanisms involved in processing dialogue games and situation models.

**Prediction-by-association** is prediction based on associations between events. For example, if *A* has previously observed that *X* precedes *Y*, then when *A* sees *X*, *A* predicts *Y*.

**Prediction-by-simulation** is prediction based on simulating another's action. For example, when *A* is attempting to shake hands with *B*, *A* can predict *B*'s upcoming hand position by imagining preparing her own hand movements and treating them as *B*'s movements (though potentially adjusting for differences between *A* and *B*). When *A* is listening to *B*, *A* can predict *B*'s utterance by imagining preparing her own utterance and treating it as *B*'s utterance.

**Production and action command.** An **action command** emanates from an action planner and initiates two processes. One process produces the behaviour and the other process produces the predicted perception of the behaviour. See Section 3.1. A **production command** is the linguistic equivalent of an action command. The production implementer converts the production command into a series of linguistic representations (semantics, syntax and phonology). The

production command also instigates processes leading to a predicted percept of the utterance being produced. See Section 5.2.

**Props** are entities that are involved in a cooperative joint activity. Props need not be ‘static’ objects, but can also be changing aspects of the environment. For example, the band music constitutes a prop for tango dancers. See also Section 11.1.

**Reference** is when a linguistic expression is used to correspond to a particular entity. The entity can exist in the world, but it also can be represented in a person’s mind.

**Representational parity** is equivalence between the representations involved in (for example) language production and language comprehension.

**Salience** is the state or quality of an item that makes it stand out from its neighbours.

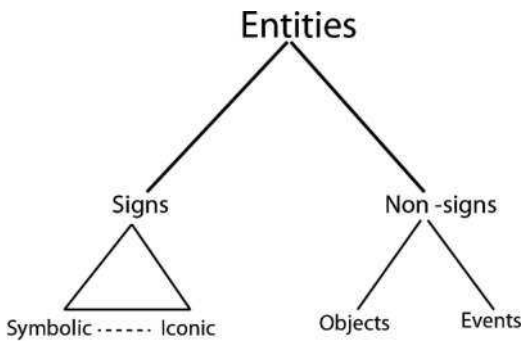
**Shared workspace.** In a cooperative joint activity involving *A* and *B*, the shared workspace contains entities that are in *A* and *B*’s parallel attention.

**Simulation** is the process by which a perceiver represents another entity’s performance of an action in the same way that she would represent her own performance of that action. Typically, this is assumed to involve the conversion of a perceptual representation into an action-based representation.

**Situation model** is a bounded (i.e. limited) representation that captures key elements of a situation under discussion. See Section 6.1.1.

**Steps.** The dialogue planner makes use of steps – that is, sequential components of the plan. See Section 9.1.

**Synchrony.** Two individuals are in synchrony when they construct and modify representations in time with each other. This can lead to synchronized activity as well as synchronized mental states.



Status of entities in the shared workspace

## Preface to *Understanding Dialogue: Language Use and Social Interaction* (Pickering & Garrod)

---

Individual people read books and prepare speeches on their own, but they can do so only because they have learned what they need to do and have practised extensively. In contrast, pairs of people find dialogue quite natural and largely straightforward, at least once they have mastered the language itself. We have always tried to understand how people use language, but rather more recently we realized that it is not enough to understand how they comprehend and produce monologue. We therefore had to address the challenge of dialogue.

There is of course a vast amount of research on how language is used in interaction – in all types of dialogue from casual conversation to formal exchanges in political or institutional contexts. Such research is conducted by linguists, sociologists, anthropologists, historians and literary theorists, and of course by some psychologists. But the great majority of this work is not concerned with what goes on in people's minds. In contrast, cognitive psychologists and neuroscientists do focus on the mind, but specifically on the individual mind engaged in monologue. There is little research on the cognitive mechanisms underlying interactive language.

To study such mechanisms, we need to do three things. First, we need to situate interlocutors within a system. It is not enough to study what each interlocutor does. We also need to understand how each interlocutor's contributions affect the other interlocutor and influence their subsequent contributions – that is, how the interlocutors jointly control the progress of the dialogue. Second, we must appreciate what is different about using language in dialogue – for example, a speaker has to be able to prepare an utterance and simultaneously comprehend an acknowledgement or query by their addressee. And third, we need to develop a theory of moment-by-moment processing. Dialogue is extremely fast and exquisitely timed, and we need to understand mechanisms that work in the order of tens or hundreds of milliseconds. This book addresses the first two concerns and takes a more preliminary approach to the third – so that an extensive programme of experimental work can use the largely chronometric methods of the cognitive sciences to understand the mechanisms of dialogue.

In the past, we have both focused extensively on isolated language processing, like almost all researchers trained as psycholinguists. And we stress that we remain interested in monologue, partly because it is an extremely important form of language use, but primarily because we eventually aim to understand language use in all its diversity. To some extent, we initially regarded dialogue as a new challenge. Simon Garrod was concerned with reference resolution in text comprehension and decided to ask how interlocutors manage to co-refer in the ‘maze game’, whereas Martin Pickering applied the method of priming from language production to see how interlocutors repeat each other’s syntactic structures. We both realized that alignment was key to dialogue – as a result of interaction, interlocutors construct similar representations to each other and therefore understand each other. We developed alignment in our paper, ‘Toward a Mechanistic Theory of Dialogue’, which is the theoretical starting point that led to this book.

However, alignment cannot fully explain how dialogue works – for example, why interlocutors can time their contributions so well in relation to each other. We therefore turned to theories of joint action and in particular to the use of prediction. In recent psycholinguistics, prediction has a central role, but the focus is on predicting the next word in monologue, primarily to allow slightly faster processing. In dialogue, prediction is far more fundamental – interlocutors could not respond to their partners (or indeed interrupt them) almost instantaneously without predicting what they were going to say and when. We wrote a paper, ‘An Integrated Theory of Language Production and Comprehension’, which argued how comprehenders use their production systems to make predictions, as part of a general theory of joint action. It developed our approach but did not focus on dialogue itself.

This book began as an attempt to integrate the theory of prediction and joint action with the theory of alignment. But in doing so, we realized that a new framework was needed, one that regarded dialogue as a cooperative joint activity within a system that is controlled by both interlocutors. We argue that the interlocutors ‘post’ their contributions to what we call a shared workspace – an aspect of reality that contains the words and objects that the interlocutors use as part of the cooperative joint activity of dialogue. They perceive, predict and monitor the contents of this workspace and are engaged in a system involving the interlocutors and the workspace itself, with the dialogue involving constant loops between these components. In this way, dialogue can lead to communicative success or failure. It should be possible to understand how dialogue takes place and what mechanisms interlocutors use to support it.

Our proposals are primarily concerned with face-to-face dyadic interaction – which we regard as the basic form of language use. But ultimately, we wish to understand more complex interactions (as well as monologue). Towards the end of the book, we consider multi-party dialogue, dialogue in which the



interlocutors are remote from each other and dialogue that is modified by technology. The book was completed before coronavirus struck, but we are writing the preface at a time when everyone's lives have been greatly affected by it. Like most people, we rapidly started using videoconferencing technologies, and it has become clear that our framework is very well suited to understanding their effects – for example, a shared document is a prop that typically enhances the shared workspace. Clearly, exploring the effects of such technologies in terms of the shared workspace framework is a potential application of our work.

In one sense, this book has been long in the making – the ideas that we developed go back about twenty years, when we began to construct our approach to dialogue. Over this long period, we have presented aspects of our work to many audiences and wish to thank a large number of people who have helped us to develop our ideas. But we have written the book over the last few years, and from during that time, we particularly thank Dale Barr, Larry Barsalou, Holly Branigan, Ruth Corps, Nic Fay, Chiara Gambi, Oliver Garrod, Rob Hartsuiker, Lauren Hadley, Peter Pickering, Natalie Sebanz, and members of the Edinburgh Psycholinguistics Coffee Group and the Glasgow CSCAN seminar group.

We dedicate this book to Martin's family Elizabeth, Clara, Harry and Amelia, and to Simon's wife Irene and grandchildren Albi, Archie, Faye, Flynn and Joey.

21 July 2020

Cambridge University Press  
978-1-108-47361-3 — Understanding Dialogue  
Martin J. Pickering , Simon Garrod  
Frontmatter  
[More Information](#)

---