Alignment in Communication

EDITED BY Ipke Wachsmuth,
Jan de Ruiter, Petra Jaecks
and Stefan Kopp

Alignment in Communication

Towards a new theory of communication

Edited by
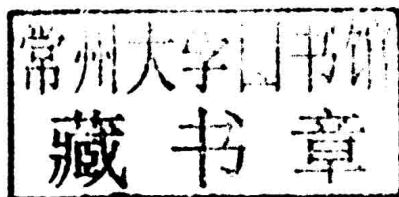
Ipke Wachsmuth
Jan de Ruiter
Petra Jaecks
Stefan Kopp
Bielefeld University

John Benjamins Publishing Company
Amsterdam / Philadelphia

Alignment in Communication

Advances in Interaction Studies (AIS)

Advances in Interaction Studies (AIS) provides a forum for researchers to present excellent scholarly work in a variety of disciplines relevant to the advancement of knowledge in the field of interaction studies. The book series accompanies the journal *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems.*

The book series allows the presentation of research in the forms of monographs or edited collections of peer-reviewed material in English.

For an overview of all books published in this series, please see
http://benjamins.com/catalog/ais

Editors

Kerstin Dautenhahn
The University of Hertfordshire

Angelo Cangelosi
University of Plymouth

Editorial Board

Volume 6

Alignment in Communication. Towards a new theory of communication
Edited by Ipke Wachsmuth, Jan de Ruiter, Petra Jaecks and Stefan Kopp

Dedicated to Gert Rickheit †, initiator and co-founder of the Collaborative Research Center "Alignment in Communication"

Table of contents

Introduction

Why a new theory of communication?

Ipke Wachsmuth, Jan de Ruiter, Petra Jaecks and Stefan Kopp

Alignment in Communication is the theme of a novel direction in communication research, which focuses on interactive adaptation processes assumed to be more or less automatic in humans. Originating from an international workshop organized at Bielefeld's Center for Interdisciplinary Research (ZiF), the book presents cornerstones of a forming new theory of communication, the ultimate purpose of which is to extend our knowledge about cognitive processes in human communication, as well as creating a foundation for natural multimodal dialogue in human-machine interaction. The introduction chapter, in addition to giving a brief overview of the content of the book, provides a definition of alignment in communication, justifies the need for this new perspective, and outlines a general framework for alignment research.

1. Subject and motivation

Cooperation and coordination in communication and action have been major research issues in the cognitive sciences and the study of artificial agents in recent years. Started in July 2006, the Bielefeld Collaborative Research Center SFB 673 "Alignment in Communication"[1] investigates special modes of coordination, called alignment (Rickheit & Wachsmuth, 2008). Alignment refers to those interactive adaptation processes among agents that are assumed to be more or less automatic in humans and therefore do not involve explicit negotiation and control by the participants. This approach thus emphasizes the role of automaticity and routinization in bringing about similarity of mental representations at various levels within and between interlocutors. Its scientific prospects are manifold: It promises the prerequisites for the specification of cognitive mechanisms in intra- and interpersonal coordination, and the generalization to domains like multimodal communication and human-machine interaction.

---

1. http://www.sfb673.org

Our research initiative, which focuses on this 'new side' of communication, aims to contribute to the theoretical development in the humanities and at the same time produce practical advancements in technology. Its ultimate purpose is to extend our knowledge about cognitive processes underlying language production and comprehension in human communication, as well as about natural multimodal dialogue in human-machine interaction. To achieve these goals, we employ an interdisciplinary approach, bringing together linguistics, artificial intelligence, neuroinformatics, neurolinguistics, computational linguistics, and psycholinguistics. In addition, we employ a conjoint methodology that integrates theory, description, and experimentation with simulation and evaluation.

There are many phenomena – for example, utterances that are started by one speaker and subsequently completed by the other speaker – clearly demonstrating that communication is, to a large extent, a matter of joint activity based on fine-tuned 'mechanistic' coordination. Most current theories of human communication, like the joint project approach (Clark, 1996) or dialogue game theory (Mann, 1988), do not provide sufficient explanations for these and similar empirical findings. Moreover, in communication and human information processing, the role of automaticity, tacit conventions, and alignment (rather than explicit negotiation) have been underestimated. Last but not least, given the decisive role that embodiment and multimodality have obtained in communication research (cf. Wachsmuth, Lenzen, & Knoblich, 2008), a new theory of communication should clearly go beyond mere language.

Consequently, we focus our research efforts on these less obvious aspects of communication and build theories of communication around the notion of alignment taking place between interlocutors. By alignment in communication, we mean both the seemingly casual process of adaptation of the participants in a conversation, as well as the state of similarity in (mental) representations ensuing from that. This is the crucial point: Similarity in mental representations, brought about by alignment processes, is what facilitates conversation and exempts agents from constantly and explicitly exchanging information about their respective points of view. Accordingly, we define 'alignment' as an ensemble of verbal and non-verbal means that serve to increase the similarity in the representational structure of two interacting dynamic systems in a largely automatic fashion, that is, without an explicit exchange of information on system states.

Our approach is inspired by the mechanistic account of language processing in dialogue by Pickering, Garrod and collaborators (Pickering & Garrod, 2004). Their notion of interactive alignment presented a novel alternative to established theories of human linguistic communication and emphasized the role of automaticity and routinization in bringing about alignment at various levels of representation both within and between interlocutors. Taking the alignment approach as a starting

point, the Bielefeld research initiative in SFB 673 has two overall goals: First, to investigate the role of alignment as a pioneering explanation of natural language use in conversation. Second, to explore the notion of alignment as a general principle of of interpersonal coordination and communication by testing the interactive alignment approach in situations that go beyond verbal conversation between humans.

Our long term goal is to work on a comprehensive theory of communication in human-human and human-machine environments. Our research agenda is firmly based on extensive empirical observations of natural language use in life-like situations from both experimental and corpus studies, and subject to evaluation by means of implementation in natural language processing systems or in artificial conversational agents. It is also integrated in technical applications in a cognitively motivated way. As a theory of dialogue it accommodates the agents that engage in conversation, verbal and non-verbal processes, and the particulars of the situation. Serving as a theory of multimodal communication, it is meant to shed more light on the role of nonverbal channels such as gesture, gaze, and body posture, which produce subtle semiotic information in dyadic face-to-face communication. Taken as a theory of human-machine communication, it is to provide new approaches for designing visual-verbal interfaces for more natural human-machine cooperation.

2.   Why is a new perspective needed?

Since the seminal work on communication as signal transmission by Shannon (1948), it has become clear that the communication between rational agents is too complex, flexible, and unpredictable to be based on signal processing alone. However, there is at present no alternative theory of communication that is (1) comprehensive enough to cover the multitude of observed communicational phenomena, and (2) concrete enough to provide guidelines and blueprints for implementing communicative systems in artificial agents.

Theorists from many different frameworks have addressed the issue of communication between rational agents in some form or other, with varying assumptions, methodologies, and empirical findings. Each of these frameworks has different strengths and weaknesses. The discourse-analytical approaches, leaning on (but extending) notions developed in theoretical linguistics, are axiomatic, very general and have the advantage of being relatively easy to implement in standard "symbolic" architectures (Poesio & Traum, 1997), i.e. computational architectures where symbols are moved between memory stores and are acted upon by an explicit set of rules. However, these implementations often lack in robustness and can be prohibitively slow. This probably also holds for the elegant but highly

reflexive/recursive approach by Grice and later by neo-Gricean theorists like Levinson (2006). A related approach advocated by Clark and colleagues (e.g. Clark, 1996) is based on tight coordination between interactants and the central notion of "common ground", which is not just shared knowledge, but knowledge that is known by the participants to be shared. The fact that the computations involved in the processing of common ground are both potentially leading to infinite recursion, and at the very least appear to be prohibitively complex, has led to new approaches such as the mentioned one by Pickering and Garrod (2004), in which communication is seen as representational alignment arising from automatic processes (such as 'priming'). Although no full implementations exist to date, this 'mechanistic' approach bears the promise to be much more efficient than the classical symbolic approaches.

The notion of interactive alignment presents a new, innovative alternative to the established theories of human communication. Nevertheless, doubts exist as to whether the alignment approach is sufficiently rich to serve as a comprehensive theory of communication, as it does not address the full range of complexities in communication (including, e.g., audience design, repair, and implicit negotiation of meaning). Further limitations still arise from the fact that the account requires further empirical support and conceptual clarification. Also, the interface between alignment and negotiation of meaning in conversation is still not very well understood. As our research programme has progressed, our knowledge about alignment processes has become more fine-grained and more sophisticated, as has our awareness of limitations of the approach. A case in point, illustrating the shift to a complementary perspective with respect to alignment research, is the discovery that synchronization of timing in human verbal actions is an important prerequisite for a model of communication. Dialogue participants' timing mechanisms do not seem to be controlled and planned to the same degree as, for example, selecting dialogue acts or producing explicit feedback. It appears that the difference between "low-level" sequencing and "higher rationality" mechanisms (like 'intending' in the Gricean manner) already requires alignment. There is ample evidence that alignment phenomena can be detected in experiments, reconstructed in small-scale theories and simulated using various methods such as real or virtual robots or programmes mimicking speakers' interaction, which has also stimulated work addressing causal explanations.

A crucial limitation of Pickering and Garrod's (2004) approach that became apparent during our research so far is that it does not incorporate incremental enrichment of situation models. This calls for a more discourse-related notion of meaning, as known from dynamic semantics (or so-called 'update semantics'). To a large degree, the meaning of an utterance is its potential to change the interlocutor's 'situation model'. The same utterance or nonverbal signal can have a

different meaning in different contexts. Thus a new theory of communication will have to go beyond routinization and look for additional, automatic mechanisms that are responsible for maintaining alignment in natural dialogues. Another limitation pertains to surface vs. structural alignment: Pickering and Garrod (2004) see alignment as producing identical linguistic surface structures by conversational dyads. Especially for aligning situation models, which is essentially a semantic-pragmatic task and also the presumed ultimate goal of dialogue in the alignment model, this turns out to be problematic. It is well known in linguistics that a wide range of variation in surface forms is possible without obstructing content alignment. Acknowledging that not all alignment relies on surface identity, and that superficial form alignment can easily lead to misalignment at other (especially higher) representational levels, we are taking this into account by paying more attention to semantic and pragmatic levels of representation and processing and by specifying what "classical" (e.g., symbolic) processes are needed to supplement the alignment model. This is necessary generally (i.e., even in ordinary communication) and not only in special cases, as Pickering and Garrod (2004) suggest.

3.   Development of the research area

There are many indications that "alignment in communication" is a vital research area in the European research communities and abroad. The Pickering and Garrod paper (2004) has also been an attempt to boost the development of dialogue theories which do not rely massively on the postulation of mental states like belief, intentions and higher order constructs such as the various kinds of mutuality and on very rich logics to model these. As a consequence, there is growing awareness of the fact that one has to capture the automaticity and economy of processing in dialogue.

Moreover, the topic of alignment has increasingly started to move into the focus of human-machine interaction researchers. Currently we see growing work on adaptive dialogue systems that reproduce audience design effects, as well as on embodied systems capable of automatic alignment of nonverbal behaviour. The latter refers to studies on how virtual characters and humanoid robots can create motor resonances in human observers and establish rapport by behavioural mimicry as a form of contingent feedback. Additionally, in human-robot interaction the focus has shifted from modelling dialogue as a means to "steer" the robot, towards more social interactions where non-verbal cues play an important role to convey robot internal states to the user. Here, alignment methods become increasingly relevant in robots that are able to show emotions, and in human-robot scenarios that need more sophisticated communication and social interaction skills.

Midway through our research agenda, an international workshop[2] was organized at Bielefeld's Center for Interdisciplinary Research (ZiF), to discuss our research progress with an interdisciplinary group of experts. One motive was the observation that many communication researchers tend to have implicit assumptions when talking about communication and its modelling and that these should be made explicit. A second motive was to set up a discussion between the different streams prevalent in communication research and dialogue theory, in order to work towards a new theory of communication. Originating from this ZiF workshop, this book presents cornerstones of a forming new theory of communication, the ultimate purpose of which is to extend our knowledge about cognitive processes in human communication, as well as about natural multimodal dialogue in human-machine interaction.

The participants of the workshop agreed that there is converging evidence from experimental and observational studies that alignment of many aspects of both non-linguistic and linguistic behaviour is pervasive in interaction. The present book contains evidence on why such alignment occurs and the role it plays in communication. Examples of alignment phenomena are complemented by a discussion of methodologies and explanatory frameworks from dialogue theory and pragmatics, ranging from human-human interaction to communication with robotic and virtual artificial systems. The book further emphasizes the multi-componential nature of alignment, including low-level cognitive mechanisms unmediated by beliefs (e.g., incremental syntax and lexical specifications, or models of temporal co-ordination), as well as connections between mental and emotional states. To more fully understand these processes, the book attains a perspective that goes beyond the individual, in that individuals' behaviours and cognitive states are related to each other, for instance, how addressees play an active role in shaping speakers' utterances, how interpersonal alignment depends on the usage of speech-accompanying gestures in dialogue, or how alignment between speakers and listeners appears to be based on interweaving between processes of language production and comprehension within each interlocutor.

4.   Outline of contents

The ten chapters that follow focus on diverse aspects of alignment in communication, including those that extend human interaction to multimodal as well as emotional communication and human-machine interaction.

---

2.   http://www.sfb673.org/tantoc

In Chapter 2 "Methodological paradigms in interaction research", Jan de Ruiter surveys a number of commonly used research methods in human interaction and communication research, and elaborates on their relative strengths, weaknesses, and advantages. It is argued that the complexity of human communication not only requires a multidisciplinary approach, but also that it is essential to combine the various qualitative and quantitative methods to gain progress in this multi-faceted field.

In Chapter 3 "A multidimensional activity based approach to communication", Jens Allwood proposes a model that aims to clarify the relations between cognition and communication on the one hand, and between communication and joint activity on the other. Face-to-face communication is seen as continuous multidirectional flow of different types of information and at several levels of processing, intentional control and awareness, resting upon a coactivation of meaning potentials between interaction partners.

Chapter 4 "On making syntax dynamic: the challenge of compound utterances and the architecture of the grammer", by Eleni Gregoromichelaki, Ruth Kempson, Christine Howes, and Arash Eshghi, sets out the broad array of compound utterance data displayed in conversational exchanges and shows how severe a challenge these data pose for common grammar formalisms. It then shows how if syntax and lexical specifications are defined exclusively in terms of mechanisms for the incremental build-up of semantic representations, the relevant findings can be explained more adequately.

In Chapter 5 "Automatic and strategic alignment of co-verbal gestures in dialogue", Stefan Kopp and Kirsten Bergmann provide an overview of their own and others' empirical research on co-verbal gestures. In their opinion, the findings reveal different kinds of partner-specific adaptations in dialogue: Automatic adaptation may occur through sensorimotor resonances and direct perception-action-links; strategic adaptation may result from recipient design processes operating on representations of common ground. A combined model of speech-gesture production and understanding is proposed to explain these different components of interpersonal alignment.

Chapter 6 "Interaction phonology – a temporal co-ordination component enabling representational alignment within a model of communication", by Petra Wagner, Zofia Malisz, Benjamin Inden, and Ipke Wachsmuth, contrasts mechanisms and models of temporal co-ordination with alignment of representations, for instance a convergence of semantic models. The fundamental assumption here is that alignment of representations on any level needs a logistic component explaining coordinative processes in time. This logistic component – or Interaction Phonology – enables interlocutors to guide their attention to relevant details in the speech signal. It is argued that dynamic oscillator models of entrainment may

provide testable formal models for the temporal co-ordination of interlocutors' speech productions.

In Chapter 7 "Communication as moving target tracking: Dynamic Bayesian inference with an action-perception-learning cycle", Byoung-Tak Zhang describes how a dynamic inference cycle model of human communication offers an overarching view to compare and integrate the mathematical tools developed separately in dynamical systems theory, decision theory, information theory, statistical physics, and computational learning. In addition, he discusses how such a model could help develop technologies for multimodal embodied interaction and human-like cognitive agents.

In Chapter 8 "Language variation and mutual adaptation in interactive communication: putting together psycholinguistic and sociolinguistic perspectives", Constanze Vorwerg discusses the role of language variation in interactive communication and its interrelation with mutual adaptation processes, providing evidence from sociolinguistic and psycholinguistic approaches. Amongst other factors, comprehension vs. production and automatic vs. strategic adaptation are discussed with respect to language variation. It is argued that a new theory of communication needs to combine different perspectives – focusing both on socially meaningful, systematic variation and on the individual mechanisms and mental representations involved in linguistic choices.

In Chapter 9 "'The hand is no banana!' On communicating natural kind terms to a robot", Julia Peltason, Hannes Rieser, and Sven Wachsmuth provide an example of how human-robot communication can be analyzed deeply using methods from dialogue theory, conversational analysis and pragmatics. Analyzing a setting where the robot "Flobi" acquires natural kind terms, the authors identify interaction problems arising from the human's mis-attunement to the social facilities of the robot. It is discussed how investigating artificial minds affords methodological advantages and thus can contribute to the development of a theory of communication.

In Chapter 10 "Interactive alignment and prediction in dialogue", Simon Garrod and Martin Pickering consider an alignment mechanism between speakers and listeners which is based on interweaving between processes of language production and comprehension within each interlocutor. Specifically it is argued that actors construct forward models of their actions before they execute those actions, and that perceivers of others' actions covertly imitate those actions, then construct forward models of those actions. This account is then discussed in relation to the contributions of this volume.

Finally, Chapter 11 "What is the link between emotional and communicative alignment in interaction?", by Petra Jaecks, Oliver Damm, Martina Hielscher-Fastabend, Karoline Malchus, Prisca Stenneken, and Britta Wrede, examines the

differences and influences between emotional and communicative processes on all linguistic levels and in particular based on the empirical evidence from studies with neurological patient groups. A three-layered computer model of emotional alignment is proposed which aims at explaining how emotional alignment could be computationally modelled and transferred to a human-robot interaction setting.

The authors and the editors hope that the contributions of this volume will stimulate further discussion of the alignment in communication perspective and that it will inspire further research towards a comprehensive theory of communication in human-human and human-machine environments.

Acknowledgements

References

Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.

Levinson, S. C. (2006). On the human 'interactional engine'. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition, and interaction* (pp. 39–69). London: Berg.

Mann, W. C. (1988). Dialogue games: Conventions of human interactions. *Argumentation, 2*, 511- 532.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue [with commentary]. *Behavioral and Brain Sciences, 27*, 169–225.

Poesio, M., & Traum, D. R. (1997). Conversational actions and discourse situations. *Computational Intelligence, 13*, 309–347.

Rickheit, G., & Wachsmuth, I. (2008). Alignment in communication – Collaborative Research Center 673 at Bielefeld University. *Künstliche Intelligenz, 2008*(2), 62–65.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27*, 379–423 and 623–656.

Wachsmuth, I., Lenzen. M., & Knoblich, G. (Eds.) (2008). *Embodied Communication in Humans and Machines*. Oxford: Oxford University Press.