

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/375889257>

Bodies to suit the music : Depictions in opera rehearsals

Book · November 2023

DOI: 10.3384/9789180754026

CITATION

1

READS

13

1 author:



Agnes Löfgren

Université de Neuchâtel

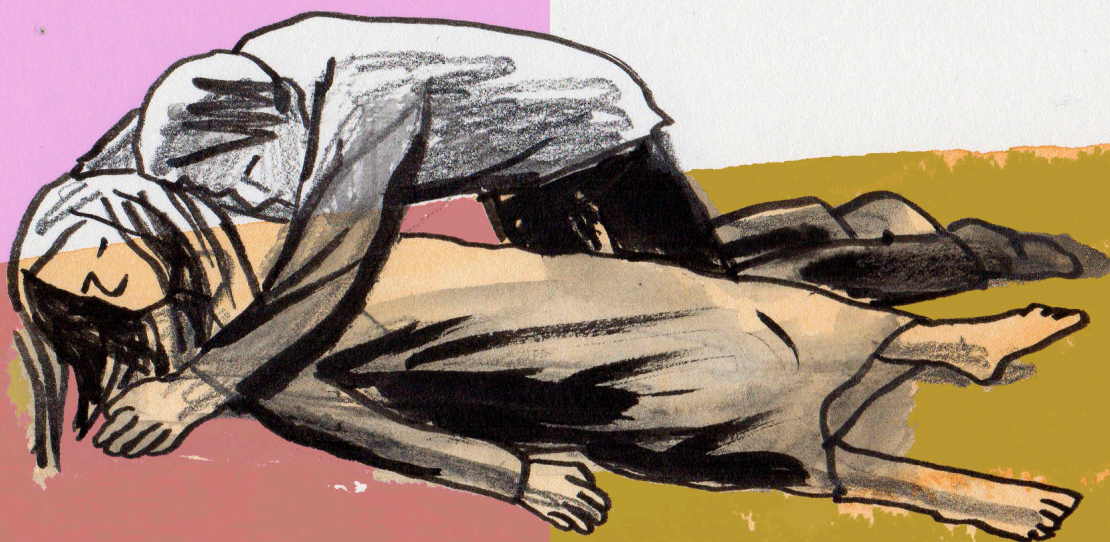
7 PUBLICATIONS 29 CITATIONS

SEE PROFILE

Linköping Studies in Arts and Sciences No. 863
Studies in Language and Culture No. 37

Bodies to suit the music

DEPICTIONS IN OPERA REHEARSALS



Agnes Löfgren

li.u LINKÖPING
UNIVERSITY

Linköping Studies in Arts and Sciences No. 863
Studies in Language and Culture No. 37

Bodies to suit the music

Depictions in opera rehearsals

Agnes Löfgren



Department of Culture and Society
Division of Language, Culture and Interaction
Faculty of Arts and Sciences
Linköping University, SE-581 83 Linköping, Sweden

Linköping 2023

This work was funded by the Swedish Research Council (VR-2016-00827, Vocal practices for coordinating human action).

© Agnes Löfgren, 2023

This work is licensed under a Creative Commons Attribution-NonCommercial International 4.0 License. Articles in the printed version have been reprinted with permission from the respective copyright holder.

Printed in Sweden by LiU-tryck, Linköping 2023

ISBN 978-91-8075-401-9 (print)

ISBN 978-91-8075-402-6 (PDF)

<https://doi.org/10.3384/9789180754026>

ISSN 0282-9800

ISSN 1403-2570

Cover image: © Anneli Furmark

Abstract

This thesis investigates how participants in 20 hours of video-recorded scenic opera rehearsals make use of *depictions* (Clark, 2016), a communicative strategy based on iconicity, to create *performance bodies*, i.e. what the performers should do on stage to music. The method is grounded in ethnomethodologically informed conversation analysis and interactional linguistics (EMCAIL). The thesis aims to reveal how participants in opera rehearsals construct and respond to depictions, what interactional and semiotic functions depictions carry, and the nature of the relationship between depictions and descriptions, and in extension between language and the body in social interaction.

The thesis comprises three individual articles. Article I focuses on how performers and the director deploy depictions in proposals of performance bodies. It is argued that depictions reference both themselves as the current state of the artwork, and prototypes of mundane behaviour (distal scenes). The article compares the self-referentiality, or *introversive semiosis*, of depictions with how interactional practices in general develop over time. Article II focuses specifically on how performers make proposals with depictions. The article concludes that depictions are multimodal gestalts whose interpersonal coordination reflects the distribution of deontic rights during the rehearsals in a visuospatial way, beyond the adjacency pair. Article III focuses on changes in turn design, and the relative deployment of depictions and descriptions, over joint decision-making micro-histories. It is shown how proposals move from descriptive to increasingly depictive states as the participants assure that there is displayed epistemic access to, and alignment and agreement with, the proposed performance bodies. The use of language early in the process secures conditionally relevant responses to the proposed ideas and thereby successful outcomes of proposals. The article reveals the essentially joint nature of the decision-making process on performance bodies.

The thesis uncovers the temporally heterogenous nature of depictions. They are achieved in stepwise manners: both in terms of their moment-by-moment realization in turns, and in terms of their development over interactional histories. They are dialogically achieved both in the local and historical sense: their successful realization is dependent on cooperation from co-present participants who are also intrinsically involved in their development over time. Further, it is argued that depictions are both an interactional practice for creating opera performances and the very same performances at their current states. The thesis contributes to a holistic and integrated view of social interaction where no resources, whether traditionally conceived of as linguistic or not, are considered more important than others for the local constitution of social action.

Keywords: opera rehearsals, social interaction, multimodal interaction analysis, ethnomethodological conversation analysis, depictions, multimodal gestalts, proposals, joint decision-making, deontic rights, joint activities, co-operative action, interactional histories.

“But it is difficult for any of us to gauge the scale on which others register our acts and words; for fear of seeing ourselves as over-important, and by magnifying hugely the dimensions to which other people’s memories must stretch if they are to cover a lifetime, we imagine that all the peripheral aspects of our speech and gestures make little imprint in the consciousness of the people we talk to, let alone stay in their memory. It is this sort of assumption which makes criminals retrospectively emend statements they have made, in the belief that no one will ever be able to compare the new variant with an older version. However, it is quite possible that, even in relation to the immemorial march of humanity, the newspaper’s columnist’s philosophy that everything passes away into oblivion may be less reliable than the opposite prediction, that all things will last.”

– Marcel Proust, *In the Shadow of Young Girls in Flower*, (1919/2003), p. 52

Acknowledgements

The realization of this project is indebted to the guidance and support from many people, whom I wish to thank here.

First and foremost, I want to extend my sincerest thanks to the participants of the study, the anonymous opera ensemble who kindly allowed me to take parts of their rehearsal process, and record and analyse them for this thesis. Without your generosity and openness, none of this would have been possible. I am ever so grateful that you gave me this opportunity to study your fascinating life at work.

I am entirely indebted to my excellent supervisors who share a contagious passion for science. My deepest thank you to Leelo Keevallik for giving me the chance to be a part of your fascinating research, for being an excellent critic, and for your reassuring guidance. To Angelika Linke for expanding my views and inspiring me, for your merciless and indispensable critique, and for valuable advice. To Emily Hofstetter for all the brilliant and critical remarks on the thesis, but above all for the immensely generous and multifaceted support over the years. It has been a pleasure and a privilege to realize this work under the supervision of the three of you.

A big thank you to my senior colleagues in the Non-lexical vocalizations project. To Sally Wiggins for the critical comments on my work, the support, and friendliness. To Adrian Kerrison for the encouragement and the inspiring scientific discussions. And thank you both for letting me learn from your intriguing research.

I feel so fortunate to have conducted this PhD at the research environment for Language and Culture, characterized by its genuine care for PhD students combined with unfailing scientific rigor. I wish to thank all

my senior colleagues for the inspiring discussions and developing critique. A special thanks to Carin Franzén and Nigel Musk who, consecutively, managed the inheritance of this wonderful environment. A big warm thanks to Agnese Grisle for the excellent administrative support and thoughtful eagerness to help, and to Veerle Bergqvist for reliant and friendly support in institutional administrative matters.

To my former and present Language and Culture and Pedagogical Work PhD colleagues: Anna Lindström, Jakob Lien, Ellen Söderblom Saarela, Lars Liljegren, Lars Jämterud, Camilla Wallin Lämsä, Emile Farmer, Sergio Passero, Wille Siwers, Amanda Hoskins, Simon Wessbo, Helen Winzell, Stina-Karin Skillermarck, Gunilla Ståhlberg, Fredrik Olsson, and Pontus Larsen. My sincere thanks for the improving remarks on my work, the support, and for making me laugh. I especially wish to thank Johanna Vernquist and Ragnild Lome for your wise advice, encouragement, and friendship, Lene Asp Fredriksen for the laughter and the tears, and finally, Hannah Pelikan for being my proof-reader, toastmaster, main support, PhD sister, and friend.

To my amazing former colleagues in the SLP unit. Thank you for letting me be a part of your team. A special thanks to Charlotta Plejert for recruiting me to the unit, and for encouraging and acknowledging my contribution there, to Inger Lundeborg Hammarström for your supervision and support, and Lovisa Elm, Sophia Lindeberg, and Elias Ingebrand for the PhD friendship.

I had the benefit of receiving excellent feedback from several readers for my milestone seminars: Christina Samuelsson, Oliver Ehmer, and Stefan Norrthon. I am so grateful for the time you devoted to my work, and for the challenging and insightful comments. I especially wish to thank Mathias Broth for the thorough and helpful critique of my arguments (and especially of the concept *depiction*), not only as a reader

for the half time seminar, but also during presentations and data sessions over these six years. I am also particularly indebted to my main opponents at half time and near end: Asta Cekaite, for the helpful conversation that guided and clarified the focus of my project, and Melisa Stevanovic for the meticulous and sharp remarks that significantly improved the thesis.

Many people within the EMCAIL field have contributed to the quality of the thesis, and to my PhD experience. Thank you to everyone in the SIS community at Linköping University for the vibrant discussions in a friendly atmosphere. To all the workshop facilitators and teachers, especially to Anna Lindström for teaching me the foundations of CA in a pedagogical and inspiring way. And to all the PhD students for excellent conference and workshop company, with all that it entails. I cannot mention you all but to Andra Rumm (the first), Maria Johansson (the most recent), and everyone else in between. To Klara Skogmyr Marian for supportive assistance on future career paths, Simon Magnusson for invaluable practical advice and the most kind-hearted support, and Sara Rönnqvist for helpful and cherished debriefings on academia, research, and life.

To the Skärholmen parents, Klara Persson, Jakob Ljungdahl, Karin Engqvist, Aron Jarrick, Hannah Wiker Wikström, Kim Ekberg, Tilde Svantesson, Evelina Ohlsson Fahlman, Sebastian Tallén, Kajsa Edqvist, Johan Edqvist, Erika Ershammar, Stefan Ershammar, Cajsja Unnbom, and Jojo Ramey Louis. Thank you for the wonderful community and for being stand-in parents when I needed time to work or get my mind off it with diverse self-care activities.

To my lovely friends who have supported me in so many ways these years, and beyond. A particular thanks to Joanna Herskovits, Alice

Luther Näsholm, and Ivonne Contardo. To Karin Wändahl for not only being a best friend but also an outstanding proof-reader and informal editor. To Vera Sundin with whom I have shared my entire PhD journey and who is the most loyal and wise company. To Julia Andersson for generously sharing your home in Norrköping and for thoughtfully caring for me in a difficult time. And to Xheni Dhani, Maria Elena Ramelli, Valentina Macario, and Elena Bernazzani for ancient Italian translations, and, with Nora Ueberhorst, for the love.

A warm thank you to Alfonso Corvera for letting me stay at your house in Norrköping, and for providing an away-from-home preschool. Thank you also to Natalie Gridell. And thank you Edgar Corvera, for believing in me, and for our (mostly) smooth cooperation that means so much.

Thank you to my brilliant aunt Anneli Furmark for illustrating a cover to this thesis that wonderfully captured the essence of a performance body, and of the entire work.

To my main support team, mamma and Toivo. Words cannot express the gratitude I feel for everything you do for us, and that has helped me achieve this work. To pappa, for your most valuable advice and support, and to Laura and sweet Leonardo, for the love and delight. To mommo for being my first and foremost teacher. To William and Lovisa, I hope you understand how much I appreciate your presence in our lives, and your generous and selfless practical and emotional support.

And finally, to my son Boris. Thank you for your patience with me when I work too much and for your delightful company when I close my laptop. I enormously treasure our little family.

With gratitude,
Agnes, Värmdö, 8 October 2023

List of articles

The thesis consists of the following articles. They are authorized for publication in this thesis by their respective editing companies.

- I. Löfgren, A. & Hofstetter, E. (2023). Introversive semiosis in action: Depictions in opera rehearsals. *Social semiotics*, 33(3), 601–620. <https://doi.org/10.1080/10350330.2021.1907180>
- II. Löfgren, A. (2023). Relocating to depict: Managing the interactional agenda at opera rehearsals. *Research on Language and Social Interaction*, 56(3), 209–230. <https://doi.org/10.1080/08351813.2023.2235968>
- III. Löfgren, A. (Under review). On the pursuit of *bodies* to suit the music: From describing to depicting proposed ideas at opera rehearsals.

Other relevant publications

- I. Hofstetter, E., Keevallik, L. & Löfgren, A. (2021). Suspending syntax: Bodily strain and progressivity in talk. *Frontiers in Communication*, 6. <https://doi.org/10.3389/fcomm.2021.663307>
- II. Keevallik, L., Hofstetter, E., Löfgren, A. & Wiggins, S. (Under review). Repetition for real-time coordination of action: Lexical and non-lexical vocalizations in collaborative time management.

Contents

Abstract.....	iii
Acknowledgements.....	vii
List of articles	xi
1. Introduction	1
1.1 Aim	7
2. Multimodal interaction	11
2.1 Conversation analysis and interactional linguistics	12
2.2 Multimodality.....	14
2.2.1 Modalities and resources: a note on terminology	15
2.2.2 Multimodal gestalts	17
2.3 Participation and joint activities.....	21
2.3.1 Visuospatial arrangements and joint activities	23
2.4 Joint decision-making, proposals, and deontics.....	27
2.4.1 Proposals	28
2.4.2 Deontics	31
2.5 Interactional histories.....	35
3. Depictions.....	39
3.1 Depiction as a communicative method.....	39
3.1.1 In favour of the term <i>depiction</i> in EMCAIL.....	42
3.2 Depictions in interaction	47
3.2.1 Iconic phenomena in interaction	47
3.2.2 Depictions and descriptions in interaction	54
3.2.3 The multimodal design of depictions	55
3.2.4 Interactional functions of depictions	57
4. Method.....	63
4.1 Multimodal interaction analysis	63
4.1.1 Fieldwork and video recording.....	64
4.1.2 Transcription.....	67
4.1.3 Analysis.....	68
4.2 Material	71
4.2.1 Participants	71
4.2.2 The rehearsal space.....	72
4.2.3 The opera house.....	73
4.2.4 The opera production	74

4.2.5 Producing an opera.....	74
4.2.5 The scenic opera rehearsal	75
4.3 Ethics.....	76
4.3.1 Consent.....	76
4.3.2 Participant anonymity	82
5. Summary of the articles	85
5.1 Article I – Introversive semiosis in action.....	85
5.2 Article II – Relocating to depict	87
5.3 Article III – On the pursuit of <i>bodies</i> to suit the music	90
6. Discussion	93
6.1 The multimodal design of depictions in scenic opera rehearsals	94
6.1.1 Depictions as multimodal interpersonal gestalts	95
6.1.2 Depictions as locus of negotiations of deontic rights	97
6.2 Interactional functions of depictions in scenic opera rehearsals	101
6.2.1 Depictions as tools in the creation of performing arts	101
6.2.2 Depictions in the service of joint decision-making.....	102
6.2.3 Depictions as manifestations of interactional histories	105
6.3 Depictions as semiotic resources.....	107
6.3.1 Depictions as performances.....	107
6.3.2 Depictions beyond opera rehearsals	111
6.4 Depictions and descriptions	114
6.4.1 What is a description?.....	114
6.4.2 Holistic sense-making in opera rehearsals and beyond	118
7. Svensk sammanfattning	123
References.....	129
Appendices.....	153
Appendix 1: Transcription conventions	153
Appendix 2: Consent information	154
Appendix 2.1 First e-mail to the ensemble.....	154
Appendix 2.2 E-mail to the choir	155
Appendix 2.3 Final e-mail to the ensemble	156

1. Introduction

This thesis investigates how an opera ensemble collaboratively creates an opera production during scenic rehearsals. The operatic expression is a product of the combination of music and performed behaviour¹, as highlighted by a quote from the director, during one of the rehearsals: “When it gets good, when opera gets good, it’s when performers, and the director too for that matter, manage to find a situation, a body, that suits the music.” In this thesis, the term *performance bodies*² will be used to refer to what the participants are creating during the rehearsals – artistic correlates of human behaviours to suit the music. In the process of finding performance bodies, the director, assistant director, and the performers collaboratively explore movements, facial expressions, and postures in coordination with different vocal expressions such as song, spoken libretto, or reported hypothetical words or thoughts of their characters. They are faced with the task of creating an Italian tragedy that produces a relevant aesthetic experience for the future audience. The artistic correlates of behaviours that they create in performance bodies must be coherent with how people behave in general, and they must resonate with the, often emotional, character that is conveyed with the music. Furthermore, the behaviours must be portrayed in a way that is coherent with the aesthetics of the production, which is under continuous

¹ Scenography, light design, and costume are also a part of that aesthetics, but this is outside of the scope of the present thesis which focuses on the creation of the fictional behaviours of the characters.

² This is a theoretically informed member’s term (see Sacks, 1992, on member’s categories). The notion of *performance bodies* is opted for rather than *bodies*, to discern this metaphorical use from the mundane understanding of the word *body*. In previous literature, the terms *blockings* (Norrthon, 2019), *representational choreographic objects* (Hazel, 2018), and *performance routines* (Schmidt & Deppermann, 2023) have been used to refer to similar phenomena.

development during the scenic rehearsals. In this thesis, I take an interest in how this unfolding aesthetics reflexively reveals itself in the interactive creation of bodies to suit the music. Although factors such as the expressive quality of the music, opera as an artform, and physical and technical feasibility are all important for the organization of the social interaction at the rehearsals and for the creation and negotiation of performance bodies, I am particularly interested in the balance between credible behaviours and the aesthetic expression of the production. In this thesis, these factors are conceived of as the difference between *extroversive* and *introversive* referents to performance bodies (cf. Jakobson, 1971; see section 3.1.1 below).

The production at hand is about a father and his daughter and a tragical life experience of theirs. When the daughter falls in love with a man who is not entirely of good intentions, the father is in despair and decides to have the daughter's love interest murdered. Due to a series of unfortunate events, however, the daughter ends up being killed instead, and the opera ends with the father mourning her death in complete desolation. To create an artistic representation of such a series of events, the director and performers draw on intuitive accounts of normative expectations of behaviour (Hazel, 2015, 2018; Lefebvre, 2018; see also Sacks, 1984a). For instance, in a real-world situation where a father realizes that his daughter is about to die, certain behaviours on his part are more expectable than others. Whereas one would find it reasonable for him to run about searching for help in vain, until he finally accepts the immensely tragic situation at hand and chooses to remain as close as possible to his daughter while she is still alive, one would probably hold the father *accountable* for sitting next to her having coffee and cigarettes, casually waiting for her to pass. The participants of the scenic opera rehearsals under scrutiny in this thesis continuously discuss appropriate

behaviours in the diverse situations that the characters of the opera face. The objective is not to portray the characters in isolation, but to make their fictive social interaction with other characters appear credible. The performance bodies should portray the characters as social actors with intentions, emotions, and desires.

The performance bodies must also match the developing aesthetics of the production. In this thesis, the aesthetics of the production is conceived of as its style, i.e. the “how” of the performance bodies that cannot be related entirely to other factors such as narrative credibility or physical and technical feasibility. Instead, it is a product of decisions that draw on forms of artistic expression. In any given situation, there is a panoramic repertoire of possibly adequate behaviour at hand, and the problem that the ensemble needs to solve is which one fits this production. For instance, a young girl who is dying could be portrayed by her intermittently lying down on the floor in a breathless manner, and intermittently regaining strength to sit up to properly sing an aria (see Extract 5 of Article 1 of this thesis). However, this solution could be rejected by the ensemble in favour of a more realistic performance body of a dying girl, even if this would entail sacrifices in the quality of the song. When creating performance bodies, the participants thus negotiate their intersubjective understanding of the developing production as they negotiate how to portray human behaviours in it.

The scenic rehearsals take place in a rehearsal space, where a substitute stage area on which the performers can act is set up. When the scenic rehearsals begin, the performers have already rehearsed the music and are well acquainted with it. Together with the director and assistant director, they now need to figure out what to do on stage in coordination with that music. The director is oriented to as an authority during the rehearsals, both in terms of establishing their agenda, and in having a

final say on performance bodies, by means of her professional identity as an expert (see Goodwin, 1994; Stevanovic, 2021). The assistant director is responsible for the practical details of the scenic rehearsals. For instance, she keeps track of aesthetic decisions, keeps everyone informed about rules and regulations at this opera house, and in other ways assists the director in her creative work. The performers are soloist singers with many years of experience of portraying opera characters on stage to music. They are experts on how the singing body works and how it can be exploited for aesthetic purposes.

Structurally, scenic rehearsals consist of two main activities (cf. Haddington et al., 2014 on *multiactivity*; and Robinson, 2013 on *overall structural organization* of interaction): performance and discussion. The performance activity happens when the performers rehearse scenes to music, in the substitute stage area, and while the other participants, mainly director and assistant director, are seated in an area that corresponds to the place where the audience would sit in the theatre. The performance activity is formally started and interrupted by the director, who announces that a performance will commence (cf. Schmidt, 2018). The discussion activity, on the other hand, occurs in between these performances, can be launched by anyone, and serves to evaluate the scenes just performed, or plan scenes that are about to be performed. The focus of the present thesis is on the discussion activity during which central elements of the developing performance are both created and negotiated, and during which performers make proposals of performance bodies.

The rehearsals are achieved through multimodal interaction (Nevile, 2015). The participants in scenic opera rehearsals make use of different visuospatial and vocal/aural resources (cf. Enfield, 2005), such as language, prosody, voice quality, gaze, facial expressions, postural configurations, movements, and the surrounding material to create,

discuss and negotiate the production. Interactional research has shown that the “peripheral aspects of our speech and gestures” that Proust’s narrator (1919/2003, p. 52) speaks of, in the quote that prefaces this thesis, in fact concern important constituents of the ways in which participants in interaction shape their behaviours in *complex multimodal gestalts* (Mondada, 2014a). These are assemblages of vocal/aural and visuospatial behaviours such as speech, gesture, gaze, movements, and postures, that are mobilized in the service of *social actions*; the things that participants do that impact their co-present others and oblige them to respond in constrained ways (Levinson, 2013). For instance, when pleading to a prison guard to release his daughter from prison, a father might hunch his back, adopt a facial expression of concern, and stretch his arms out towards the man while saying, in a worried tone of voice, “please have mercy and let my daughter go”. These different resources – posture, facial expression, gestures, and lexical choices are all mobilized in a multimodal gestalt towards one and the same social action: to request/plead. The request/plead in turn introduces a normative constraint on the prison guard, who is expected to respond in an *adequate* way (cf. Garfinkel & Sacks, 1970; Levinson, 2013), typically by granting or rejecting the request, in whatever interactional format they deem suitable for *that* interactional instance. In certain research traditions, some of these resources have indeed been considered peripheral and even irrelevant for the study of communication. Research on multimodal interaction has nevertheless proven that they are crucial for interpersonal sense-making (see for instance Nevile, 2015, for an overview). This thesis argues that it is precisely these peripheral aspects which constitute performance bodies as *artistic correlates* of multimodal gestalts, and which the participants aim to create. By artistic correlates, I mean that these multimodal gestalts are not spontaneously produced. Instead, they are

invented, rehearsed, and performed, and thus produced with a certain amount of self-reflexivity, although they draw on deeply rooted intuitive accounts of human behaviour.

Previous interactional literature has targeted how such artistic correlates of multimodal gestalts are created in theatre rehearsals (Hazel, 2018; Lefebvre, 2018; Lefebvre & Mondada, 2023; Norrthon, 2019; Norrthon & Schmidt, 2023; Nissi & Stevanovic, 2021; Schmidt, 2018; Schmidt & Deppermann, 2022, 2023). Although scenic opera rehearsals bear many similarities with theatre rehearsals, a couple of features distinguishes the two. Firstly, as the vocal/aural parameters in opera are predetermined by the music, what is created are the visuospatial aspects of multimodal gestalts. In theatre rehearsals, on the other hand, the participants often elaborate with prosody, pauses, overlaps, and sometimes even text as a part of the artistic labour (see Hazel, 2018; Norrthon, 2019). Further, in contrast to naturalistic theatre, the performance bodies created during opera rehearsals must suit the emotional character of the music. The participants need to understand why the music sounds tragic, frightening, or cheerful at a certain moment in the script and incorporate that rendition in their performance bodies. Finally, opera as an artform is traditionally characterized by its bombastic and often exaggerated and extreme character, not least in terms of the song (see Atkinson, 2006). This contrasts with naturalistic theatre (Hazel, 2018) where the aesthetics of the performance bodies more closely resemble how people interact in real life.

This thesis specifically focuses on how the participants of opera rehearsals propose performance bodies to be included in the production. Proposals are social actions that invoke a relatively weak deontic claim as they initiate processes in which decisions are jointly established, in contrast with for instance instructions and orders (see Couper-Kuhlen,

2014; Stevanovic, 2015; and Thompson et al., 2021, on proposals in interaction). To make proposals of performance bodies, participants make use of two different communicative strategies, or interactional practices as they will be referred to in this work – *depiction* and *description* (Clark, 1996, 2016, 2019). Whereas depictions are based on iconicity and *show* performance bodies, descriptions are based on conventionality and *tell* about the performance body, using language. For instance, as revealed in Extract 1 of Article II of this thesis, a participant may either describe a performance body in a proposal by saying “If I choose to remain on stage I can walk away much later”. However, they may also depict it by walking away from the people with whom they are involved in a conversation, adopting a limping style of walking characteristic of their character, and giving words to hypothetical thoughts of that character, such as “Damn it, now things are starting to happen”. In this thesis, I focus on how specifically depictions are used as practices in scenic opera rehearsals. What are the multimodal resources participants use to construct depictions? Why are they used? And what is the division of labour between depictions and descriptions, i.e. when is it necessary for the participants to speak about the performance bodies, and when can they perform them?

1.1 Aim

The aim of this thesis is thus to explore the role of the interactional practice of depiction when an opera ensemble creates performance bodies to suit the music of the opera production. With the thesis, I set out to investigate the psycholinguistic and semiotically grounded concept of depiction (see Clark, 1996, 2016, 2019; Peirce, 1955), from a qualitative and emic perspective using multimodal interaction analysis (Broth & Keevalik, 2020; Stivers & Sidnell, 2013). Clark (2016) offers a holistic account

of the concept of depiction and uses it to refer to different kinds of iconic communicative practices (see chapter 3 of this thesis). In this thesis, I aim to discuss the potential of such a concept for the social interaction of scenic opera rehearsals, and beyond.

The analysis looks at the millisecond-by-millisecond deployment of interactional resources to create depictions. Thereby, I aim to contribute to emerging knowledge on how iconic phenomena are multimodally achieved in interaction (Cantarutti, 2020, 2021; Ehmer & Mandel, 2021; Holt & Clift, 2007; Niemelä, 2010; Sidnell, 2006; Thompson & Suzuki, 2014). Further, I aspire to add to previous literature in which the role of recipients of depictions is considered (Cantarutti, 2020, 2021, 2022a, 2022b; Ehmer & Mandel, 2021; Keevallik, 2014b; Thompson & Suzuki, 2014; Stukenbrock, 2012), and in doing so, I hope to contribute more generally to studies on the multimodality of participation, cooperation, and joint activities (cf. Deppermann & Streeck, 2018; Deppermann et al., 2021; C. Goodwin, 1979, 1980, 1981, 2000, 2007a, 2018; C. Goodwin & M. H. Goodwin, 2004; M. H. Goodwin, 1980; Kendon, 1990; Mondada, 2009, 2012; Stevanovic & Monzoni, 2016; Stivers, 2008; Stivers & Sidnell, 2016; Szczepek Reed et al., 2013). With the use of multimodal interaction analysis, I investigate “how *all* participants’ activities contribute to bring off the event in the precise multimodal and sequential shape it gets” (Deppermann, 2013, p. 3, italics in original). If, as Clark (2016) argues, depictions are staged scenes for which there are spectators, how do participants stage such scenes, and how do the recipients of depictions contribute to that staging?

In the thesis, I focus particularly on how depictions contribute to joint decision-making on performance bodies. Consequently, I hope to contribute to literature on proposals and joint decision-making (cf. Couper-Kuhlen, 2014; Houtkoop, 1987; Lindholm et al., 2020;

Lindström, 2017; Nissi, 2015; Stevanovic, 2012, 2015, 2021; Stevanovic & Monzoni, 2016; Stivers & Sidnell, 2016; Thompson et al., 2021), particularly on how visuospatial resources become relevant, both for the formation of proposals, and for negotiations of deontic rights (Clifton et al., 2018; Kuroshima, 2023; Magnusson, 2022; Stevanovic, 2015, 2018, 2021; Stevanovic & Peräkylä, 2012, 2014; Van de Mierop et al., 2020).

Further, I aim to examine how joint activities and projects such as collaborative creation, idea formation and decision-making (Due, 2016; Yasui, 2013) develop over time, beyond immediate interactional sequences (cf. Schegloff, 2007). I take an interest in how the design of proposals of performance bodies changes over interactional histories (Broth et al., 2017; Deppermann, 2018a; Deppermann & Pekarek Doehler 2021). I particularly wish to contribute to the emerging understanding of how performing arts come into being in interaction over time (cf. Deppermann & Schmidt, 2021b; Hazel, 2015, 2017, 2018; Lefebvre, 2018; Lefebvre & Mondada, 2023; Norrthon, 2019, 2021; Norrthon & Schmidt, 2023; Schmidt, 2018; Schmidt & Deppermann, 2022, 2023). I do this by focusing specifically on how depictions are a tool in the creative process and how they figure in relation to language (see also Schmidt & Deppermann, 2022). By examining the relationship between descriptions and depictions in interaction when artistic non-verbal communication, namely an opera production, is created, I wish to contribute to a more general discussion on the role of language vs. the role of the body, both when creating performing arts, and in human sense-making in general (cf. Deppermann, 2013; Deppermann & Streeck, 2018; Harjunpää et al., 2023; Jakobson, 1971; Keevallik, 2013a, 2018; Mondada 2016; Nevile, 2015 Schegloff, 1984).

In the thesis, I aim to answer the following inquiries on the nature of depictions in scenic opera rehearsals:

- 1) How do participants construct and respond to depictions?
- 2) What are the interactional and semiotic functions of depictions?
- 3) What is the relationship between depictions and descriptions, and what does this relationship tell us about the roles of language and the body in human sense-making?

The thesis begins with two chapters that outline the theoretical background of multimodal interaction analysis (chapter 2), and the theoretical background of the concept of *depiction* (chapter 3). Thereafter follows an account of the method and material of the study (chapter 4), a summary of the articles it comprises (chapter 5), a concluding discussion (chapter 6), and a summary in Swedish (chapter 7).

2. Multimodal interaction

This thesis has its background in studies of multimodal interaction, in the two distinct albeit connected research traditions of Conversation Analysis (CA) (Sidnell & Stivers, 2013) and Interactional Linguistics (IL) (Couper-Kuhlen & Selting, 2017). These disciplines' conception of multimodal interaction differs radically from the notion of *multimodal interaction* treated in the field of social semiotics (see Kress & van Leeuwen, 2001) and sociolinguistics (see Norris, 2004), by virtue of its emic, naturalistic, and data-driven approach, outlined in section 2.1 below.

Studies in CA, IL, and related fields have provided a dialogical and situated account of language and communication, in contrast to the monological and decontextualized view of language that has dominated linguistics as a discipline ever since Saussurean structuralism (Linell, 2005, 2009, 2020). Studying language and communication in their natural occurrence, with sensitivity to their fine-grained structural details and orderliness, has revealed how contextual factors both shape, and are reflexively shaped by, how we make sense of each other in interaction (Heritage, 1984a; see also Goodwin, 2000). Meaning is not simply inherent in words and grammar, as some linguistic disciplines have it, but the *meaning potentials* (Linell, 2009) in words, grammar, and other resources, are exploited in interaction through both turn design and “the turn's placement in relation to a preceding turn” (Stivers & Sidnell, 2005, p. 3). The shift has gone from conceiving of language as an abstract system to an “always-evolving (living) product of creative problem-solving and invention” (Streeck, 2018).

The present chapter outlines the theoretical and empirical framework of multimodal interaction and covers aspects of social interaction that are relevant for the understanding of depictions in opera rehearsals.

It begins by describing the fields of conversation analysis and interactional linguistics (2.1), before developing on the concept of multimodality (2.2), participation and activity in social interaction (2.3), joint decision-making, proposals and deontics (2.4), and interactional histories (2.5).

2.1 Conversation analysis and interactional linguistics

Conversation analysis and interactional linguistics study language and social interaction in naturally occurring material with qualitative, inductive methodology (Sidnell & Stivers, 2013). The analytical claims are based on the *emic orientations* of the participants, i.e. on how the participants themselves interpret each other's behaviours. This analytical mindset stems from Garfinkel's ethnomethodology (see Garfinkel, 1967), hence the terminology ethnomethodological conversation analysis and interactional linguistics (EMCAIL). Within the EMCAIL framework, social interaction is conceived of as orderly and sequentially organized (Sacks et al., 1974). This means that participants in social interaction produce *turns-at-talk*, utterances that follow each other in specific ways.

The emic orientation manifests itself in one of the core analytical tools deployed in conversation analysis: the *next-turn proof procedure* (Sacks et al., 1974; Schegloff, 1996a). The latter stipulates that a specific turn-at-talk can only be interpreted as performing a certain social action if the other participants treat it as that specific action in the conversational turn that immediately follows (Maynard, 2013). Social actions are essentially "the main business" of a turn (Levinson, 2013, p 107),³ i.e. what the turn accomplishes (for instance greeting, proposal, assessment,

³ In contrast to its "collateral effects" such as displaying epistemic/deontic/emotional stance or indexing identity and relationships (Sidnell & Enfield, 2014, p. 426). See also Levinson 2013, on "primary" and "secondary" actions.

etc.). A proposal of what a performer can do on stage, can thus only be analysed as a proposal if it is treated as such, that is, as a matter that should be treated interactionally and decided upon jointly, by the other participants.

Producing a first turn (for instance a proposal) induces a normative constraint, or *conditional relevance* (Schegloff, 2007), where the recipient is expected to produce a second turn that conforms to that first turn (for instance acceptance/rejection of the proposal, cf. Houtkoop, 1987). Together, the first turn and its conditionally relevant response constitute an *adjacency pair*, a basic sequential structure. The second part of the adjacency pair, if absent, is “officially” so (Schegloff, 1968, p. 1083); participants can make each other accountable for not producing a (relevant) response. According to EMCAIL research, participants in interaction continuously display to one another what they are doing, and what they understand each other to be doing, in the pursuit of intersubjectivity, i.e. mutual understanding (see Garfinkel 1967; Sacks et al., 1974). It is these displays and their orderliness that researchers within these traditions investigate (Couper-Kuhlen & Selting, 2017), and that are analysed in the present thesis.

Participants in interaction accomplish social actions through *practices*, recurrent assemblages of behaviours, in specific sequential positions and contexts (Deppermann & Haugh, 2022; Heritage, 2010; Schegloff, 1996b, 1997, 2007). These practices help participants ascribe social actions to each other’s turns, and thereby to design appropriate responses (Levinson, 2013). In this thesis, I will show how depiction is one such practice to accomplish the social action of proposals, and in extension create the performance through joint decision-making on performance bodies. The practice of depiction, in turn, is constructed through several concurrent *resources* (see section 2.2.1 below), both vocal/aural

and visuospatial. In EMCAIL literature, assemblages of resources in practices are conceived of as *multimodal gestalts* (Mondada, 2014a), a concept that will be outlined in the next section.

2.2 Multimodality

Although early research on social interaction within the CA tradition focused mainly on linguistic resources in “talk”,⁴ an embodied turn can now be said to have taken place within the research field (Nevile, 2015). The increasing use of video recordings as material (in contrast to the telephone calls analysed in the earliest years of CA), beginning in the 1980s (for instance C. Goodwin, 1979; 1980; M. H. Goodwin, 1980; Heath, 1986), allowed for investigations of how the organization of turns-at-talk are informed and shaped by gestures, body movements, gaze, facial expressions and the surrounding environment, in addition to lexicosyntax and prosody, that is, the only behaviours captured by audio recordings. Social actions reveal themselves in what is often referred to as several *modalities* (see section 2.2.1 below) simultaneously (Deppermann & Streeck, 2018; Mondada, 2014a; Stivers & Sidnell, 2005). These modalities may support, extend, or modify each other in sense-making, as well as contribute in complementary ways to the management of turn-taking and co-operative action (Stivers & Sidnell, 2005). This section begins by describing the terminology used in research on multimodal social interaction, and motivating terminological choices made for this thesis (section 2.2.1). Thereafter, I discuss the concept of multimodal gestalt and its importance for the present analysis (section 2.2.2).

⁴ Though see for instance Sacks & Schegloff (2002, based on a paper from 1975).

2.2.1 Modalities and resources: a note on terminology

In this work, the term *resource* will be used to refer to different aspects of human behaviour that contribute to sense-making (cf. Goodwin, 2013, for a similar use of the term *resource*). Such resources may include for instance bodily configurations (Kendon, 1990), gaze (Rossano, 2013), gestures (Streeck, 2009), facial expressions (Kaukomaa et al., 2015), grammar (Keevallik, 2020a), prosody (Szczepek Reed, 2010), voice quality (Persson, 2020), non-lexical vocalizations (Keevallik & Ogden, 2020), and touch (Cekaite & Mondada, 2020). The different resources that participants mobilize towards social actions in interaction are commonly grouped into *modalities*, hence the term *multimodal interaction* (Deppermann & Streeck, 2018, Mondada, 2014a; Stivers & Sidnell, 2005). Although in some work, a *modality* is defined in a similar way to a *resource* (cf. Deppermann & Streeck, 2018), the term *resource* is chosen over *modality* in this thesis as the former stresses the usage-based approach of the EMCAIL methodology. Multimodal resources are mobilized by participants in different practices to accomplish social actions.

Nevertheless, I continue to deploy the term *modality* to refer to differences in how groups of resources are mainly produced and perceived. In doing so, I follow Enfield's (2005), distinction between the *vocal/aural* (spoken language, including prosody) and *visuospatial* modalities (gesture, gaze, and body postures). These modalities are multimodal themselves (cf. Deppermann & Streeck, 2018). For instance, the vocal/aural modality divides into two separate "channels": the lexicosyntactic and the prosodic (Stivers & Sidnell, 2005). These are argued to be separate precisely because they can achieve different interactional functions simultaneously, whereas this is not the case with, for instance, the lexicosyntactic channel only. Further, visuospatial behaviour can be

conceived of as multimodal as well, as it can simultaneously provide propositional information through some gestures, and build participation frameworks through postural orientations (Goodwin, 2000). Importantly, neither visuospatial behaviour nor the prosodic channel are sequentially organized in the same way as lexicosyntactic turns-at-talk. In addition, visuospatial behaviour and prosody is not always treated as accountable (Schegloff, 1984). The prosodic channel, as well as the entire visuospatial modality, could thus be considered as separate interactional orders from turns-at-talk (Deppermann & Streeck, 2018).⁵

The terminology of visuospatial and vocal/aural is opted for rather than the commonly prevailing distinction between linguistic and bodily resources (reflected in the terms *embodied resources* vs. *linguistic resources*), within EMCAIL. As has already been pointed out, language is produced by living bodies (Streeck, 2013). The separation of language and the body is thus unmotivated and misleading (see also Keevallik, 2018). It is to be noted, however, that the distinction between vocal/aural and visuospatial is not entirely discrete. In fact, visuospatial behaviour can be heard and felt, and vocal/aural behaviour seen, and possibly also felt (Mondada, personal communication, June 26, 2023; further evidenced by the famous McGurk effect, McGurk & Macdonald, 1976). Further, in general, some EMCAIL researchers argue for caution in conceiving of resources as separate channels and stress the holistic nature of social interaction (see Deppermann, 2013; Streeck, 2013). Deppermann (2013) argues that there is a risk of reifying modalities as “analytic entities”, rather than as made relevant from a participant perspective (see also Ford 2004; and Fox, Ford & Thompson, 1996, on the dangers of

⁵ However, some researchers (see for instance Stukenbrock, 2018) argue in favour of multimodal adjacency pairs, multimodal actions that are sequentially organized and subject to conditional relevance.

working with pre-conceived categories when conducting emic analysis). Being mindful of this danger in the analyses, I continue to use the terminology of *visuospatial* and *vocal/aural modalities* to reflect the main difference in which sense-making occurs. As will be shown, the practical consequences of the different features of visuospatial and vocal/aural resources (see prior paragraph) are relevant for the participants when constructing and responding to depictions.

2.2.2 Multimodal gestalts

Some visuospatial resources, such as facial expressions and gaze, and vocal/aural resources, such as emotional prosody and voice quality, are possibly what Proust (1919/2003, p. 52) aims at when he speaks of the “peripheral” aspects of speech and gesture. In a scene of *In the Shadow of Young Flowers in Bloom*, the narrator, an adolescent in love, speaks to an acquaintance of his father who promises to mention the narrator’s name to the mother of his love interest. Upon hearing this, the narrator describes how he becomes so grateful to this man, who is essentially a stranger to him, that he feels he could kiss his hands. On another occasion, when they meet again, it becomes evident that the father’s acquaintance had, somehow, been able to read the narrator’s (slightly inappropriate) joy of some “peripheral” aspect of his behaviour in that moment. Something in his gaze, facial expression or tone of voice had revealed his deep gratitude, although he did not intend to reveal it.⁶

In some linguistic disciplines, such as psycholinguistics, these “peripheral” resources have been referred to as *para-linguistic*, tangential to the “message”, that is conveyed via language (see for instance Dietrich

⁶ These resources are a type of index in Peirce’s (1955) sense. By virtue of being unintentionally meaningful, Peirce (1955) refers to them as *natural signs* (see also Clark, 1996).

et al., 2008; Trager, 1958). However, research on multimodal social interaction (Broth & Keevallik, 2020; Mondada, 2016; Nevile, 2015; Stivers & Sidnell, 2005) has evidenced that there is no need to assign language any *a priori* dominance vis-à-vis the other resources that participants use to make sense of each other in interaction.⁷ Consequently, within EMCAIL,⁸ language has come to be conceived of as but one, and not necessarily the most important, sense-making resource in a wider web of semiotic fields or systems (Goodwin, 2000; C. Goodwin et al., 2002) – a *multimodal ecology* (Mondada, 2018a). The concept of *complex multimodal gestalts* is a combination of multimodal resources all in the service of “coherent courses of action” (Stivers & Sidnell, 2005, p. 1; see also Mondada, 2014a). The depictions under scrutiny in this thesis will be argued to be precisely such multimodal gestalts, and performance bodies to be their artistic correlates.

Exactly how multimodal resources are tied together to form coherent courses of action, and how recipients ascribe actions to them, is still debated (Deppermann & Haugh, 2022; Deppermann & Streeck, 2018; Levinson & Stolk, 2019). In other words, it is not evident which aspects of a person’s behaviour are a part of a specific multimodal gestalt, and which are not. Temporal coordination has been suggested to be highly influential in interpreting different multimodal resources as coherent and in the service of one and the same action (Deppermann & Streeck, 2018; Stivers & Sidnell, 2005). However, whereas different multimodal resources targeting an action may be synchronically deployed, most patterns of coordination “involve asynchronous relationship between

⁷ This is reflected in the deployment of the term *interaction* rather than *conversation*, the former favorizing a non-logocentric approach to human sense-making (Goodwin, 2000; Mondada, 2014a; 2016; Nevile, 2015).

⁸ Although IL work is partially characterized by giving language a particular focus (see Couper-Kuhlen & Selting, 2017; Fox et al., 2013).

component actions” that “project, prepare, complement, elaborate, modulate, complete, etc. other component actions” (Deppermann & Streeck, 2018, p. 12). Thus, it is hard to specify (even for the participants themselves) where an action gestalt’s resources begin deployment, or end. There is substantial evidence that visuospatial resources project and prepare for vocal/aural actions (Mondada, 2009; Mondada & Schmitt, 2010; Schegloff, 1984; Streeck, 1993, 2009; Streeck & Hartge, 1992; Stukenbrock, 2018), which has also been evidenced for the relative deployment of visuospatial and vocal/aural resources in depictions (Cantarutti, 2020; Ehmer & Mandel, 2021; see also chapter 3 of this thesis). Stukenbrock (2018, p. 41) speaks of the *multimodal compaction zone* – a temporal area where multimodal resources converge to index semantically rich information, the *informational climax*. As such, gaze coupled with a deictic expression, produced with emphatic prosody, can help orient the recipient’s attention towards an iconic gesture that carries such informational value. Much as will be evidenced for depictions in opera rehearsals, Stukenbrock’s research evidence how gestalt phenomena are not to be understood as static objects but as “multi- and hetero-temporal configurations that have an interactional temporality that is flexibly adapted to the local context” (Stukenbrock, 2018, p. 64).

Whereas Proust’s (1919/2003) distinction between peripheral and central aspects of speech and gesture hints at a distinction between intentional and unintentional, within EMCAIL it is reasoned that this distinction is not necessarily relevant for the participants. Consequently, researchers within the EMCAIL tradition typically seek to remain at least “agnostic” (Hopper, 2005) about intentionality, thoughts, and rationales behind behaviour (see for instance Sacks, 1992; and discussion in Potter & Edwards, 2013). It is argued that the analyst (and participants) cannot know what a person intended or felt, they can only observe what they

choose to, or involuntary, reveal in interaction. As for the participants in interaction, they can draw on a worried tone of voice as a resource to make sense of a person's behaviour, and as a part of a multimodal gestalt, regardless of whether the person wished to express worry, or if, in contrast, they wished to not do so.

To create performance bodies, i.e. artistic correlates of multimodal gestalts, opera professionals draw on their intuitive understanding of all of the multimodal aspects that have traditionally been considered peripheral to language – both the resources participants actively mobilize to accomplish social interactions, such as a turn-at-talk performing a request, *and* those that may be happenstance, and that index for instance, a certain type of emotion, such as worry that the request will not be granted. The focus is on the “peripheral” *visuospatial* aspects of multimodal gestalts, as the turns-at-talk, and other vocal/aural resources such as prosody, are predetermined by the libretto and musical notation. The participants of opera rehearsals can only choose to nuance the fictive social actions that these turns-at-talk imply, to reveal the characters' motives, intentions, emotions, and desires, and to make their social interactions appear credible, recognizable, and touching, for the audience.

As evidenced above, research on multimodal gestalts reveals that social interaction is not only sequentially ordered but consists of “sequentially ordered simultaneities” (Mondada, 2018b) in which participants co-operatively achieve and coordinate their actions (see Hofstetter & Keevallik 2023; Keevallik, 2020b; and Szczeppek Reed, 2023, for the co-operative achievement of simultaneity). Further, participants may collaborate in the achievement of gestalts, making them interpersonal phenomena (Stukenbrock, 2018). A central aim for this thesis, beyond exploring the multimodal nature of depictions, is to explore their co-operative, joint, nature. It will be shown that for a depiction to be

successful, the recipients of it must participate in it in one way or another. The next section reviews previous research on such cooperation, or participation, in joint activities in social interaction.

2.3 Participation and joint activities

Social interaction is a joint achievement on a fundamental level. Participants in interaction “inhabit each other’s actions” (Goodwin, 2018, p. 11). They reflexively and continuously shape each other’s turns through concurrent visuospatial resources,⁹ in *micro-sequential actions* (Deppermann & Streeck, 2018; Deppermann et al., 2021; C. Goodwin, 1979; 1980, 1981, 2000; M. H. Goodwin, 1980;). This has been conceived of in terms of *participation*, a term borrowed from Goffman (1981) but developed by C. Goodwin and M. H. Goodwin (C. Goodwin 2007a; C. Goodwin & M. H. Goodwin, 2004). This section reviews the ways in which participation, co-operative action, and joint activities are achieved, and reveals how these achievements are important for the understanding of depictions in opera rehearsals.

Stivers (2008) has described two different ways in which participants may collaborate with an ongoing turn-at-talk. They can *align* with the structural implications of the turn-at-talk, for instance by producing continuers that allow for suspension of the normal rules of turn-taking that permits the other participants to pursue an initiated story. Importantly, *alignment* does not necessarily imply that the recipient(s) *affiliate* with the emotional stance conveyed by the story; it is simply a means to acknowledge a contribution to the *local interactional agenda* (Stevanovic & Peräkylä, 2014), for what it is worth (see also Stivers et al.,

⁹ For evidence on co-operative vocal/aural phenomena, see for instance Stivers, 2008, and Szczepek Reed, 2012.

2011). In that sense, alignment is ultimately tied to deontic rights in interaction (see section 2.4.2 below). In opera rehearsals, as in many other settings (notably in Stivers, 2008), alignment may be achieved through different visuospatial orientations and configurations of the interactional space (see also Ehmer, 2021; and Krug, 2022, on synchronization in dance lessons and dance rehearsals respectively).

Participants in interaction continuously position themselves in ways that calibrate the interactional space and co-operate with the actions and projects of other co-participants (Fatigante et al., 2010; Mondada, 2009, 2012, 2018a; Usatch, 2000). Configurations of the interactional space can be achieved through postural changes (Streeck, 2018) and whole-body spatial relocations (Broth & Mondada, 2013, 2019; Mondada, 2009). For example, Kendon (1990) described how participants configure in F-formations during conversation: they position themselves so that they orient to a mutually shared space. These configurations are a part of the participation framework of interaction (C. Goodwin, 1981; M. H. Goodwin, 1990, 1997; Heath, 1986; Kendon, 1990). They are “displays of mutual orientation made by the actors’ bodies” (Goodwin, 2000, p. 1492) and “part of the interactive ground from which actions emerge, and within which they are situated” (ibid., p. 1496). This framework is not static and fixed for each interaction, but always “open to challenge, negotiation, and modification” (ibid., p. 1496). In opera rehearsals, visuospatial arrangements are an important way to coordinate alignment and the participation framework around depictions. When a participant initiates a depiction through the visuospatial resource of a relocation in space (see Article II of this thesis), the other participants must choose to align, or not, with the change in participation framework, and reconfiguration of the interactional space, that is suggested by the initiation of a depiction. The ways in which

participants configure the interactional space and participation framework allow them to manage *joint activities*.

2.3.1 Visuospatial arrangements and joint activities

There is no precise definition of *activity* within EMCAIL, and the “scale and scope [of activities] can vary greatly” (Betz et al., 2020, p. 3; see also Robinson, 2013). According to Heritage & Sorjonen (1994), activities span over multiple sequences: they are larger structures than sequences. Further, they are characterized by an overall structural organization, that is, “smaller units (actions, sequences) [that] relate to each other to form a recognizable whole” (Betz et al., 2020, p. 3). A main activity may consist of several sub-activities (Robinson & Stivers, 2001), and participants may be faced with multiple activities in interaction, although it is not always easy to disentangle what constitutes multiple activities rather than one complex main activity (Haddington et al., 2014).¹⁰ Multiple activities can occur either simultaneously or serially, where one activity follows the other (Haddington et al., 2014). They are managed through intra- and interpersonal coordination, meaning that participants must both organize their multimodal behaviour with respect to the different activities at hand, as well as coordinate these activities with the other co-present participants (Deppermann, 2014). In terms of intrapersonal coordination, for instance, participants can orient to one activity with one modality (e.g. vocal/aural) and to another activity with another modality (e.g. visuospatial) (Mondada, 2014b). Activity transitions can be indexed linguistically, such as verbally asking someone to wait while another activity

¹⁰ Haddington et al. (2014) argue that this judgement should be based on participant orientations to two or more activities as in some way attended to separately.

is carried out (Keevallik 2001; Keisanen et al., 2014) or they may be indexed by gaze and bodily orientations (Hofstetter, 2020b, 2021; Ticca, 2014).

This thesis studies one main activity: rehearsing the scenic aspects of an opera production. During the rehearsals, however, participants do not perform all the time, they also discuss these performances, and other relevant practical questions. The main activity is thus in turn divided into two primary sub-activities: discussion and performance (see also chapter 1, and section 4.2.5 below). This makes the scenic opera rehearsal a multiactivity setting where two main types of activity contribute in different ways to the achievement of the greater activity, or the overarching project of the interaction.¹¹ The depictions that are the focus of this thesis occur during the discussion activity of the rehearsals. However, by performing a depiction, the participants orient briefly to the performance activity (see Article II of this thesis). Introducing a depiction, i.e. a performance, can thus be conceived of as a momentary shift in joint activity.¹² This shift is co-operatively achieved through the reconfigurations in participation framework and interactional space, as a practice in the service of social actions (in this thesis, mainly proposals, see section 2.4.1 below).

Previous research has shown how participants manage the sustainment, transition between, and abandonment of joint activities through different visuospatial orientations. Participants may, for instance, hold gestures throughout distinct activities, indexing the activities as an entity (Streck, 2018; see also Rönnqvist & Lindström, 2021, on gestural managements of turn-taking). Further, spatial relocations such as walking

¹¹ The conversations that the participants hold during the discussion activity are mostly “activity-oriented” talk (Usatch, 2000), that is, talk that is related to accomplishing the goal of creating an opera performance.

¹² In a wide definition of the concept of activity, as a unity (the depictions) that consists of smaller parts (the resources participants deploy in the service of the depiction) (see Betz et al., 2020, p. 3).

away may project the end of an activity (Broth & Mondada, 2013, 2019; Reed, 2015). These visuospatial achievements of joint activities are often a collaborative endeavour. Although single participants may “propose” changes in activity through visuospatial means, the ultimate realization of these changes implies that the other participants align with the “proposal” for activity transition (Stevanovic & Monzoni, 2016; Stivers & Sidnell, 2016). If there is a mismatch in verbal and visuospatial behaviour during joint activities, participants orient primarily to the visuospatial realm, suggesting a “primacy of embodied behaviour in the management of joint activities” (Stevanovic & Monzoni, 2016, p. 15; see also Szczepek Reed et al., 2013, p. 32, for a similar finding).

A central argument of this thesis is that depictions are collaboratively achieved performances that are constructed in interaction through diverse multimodal resources. A central topic will thus be how the collaborative transition into performance is achieved. In instructional and creative artistic settings, previous research has demonstrated how participants transition between non-performance and performance activities in collaborative ways through different visuospatial orientations (see below). It has revealed how visuospatial arrangements and multimodal orientations help construe the participation framework as *doing performance* vs. *doing conversation*. To commence a rehearsal activity in a theatre setting, the participants go through visuospatial preparations, for instance “shaping a specific bodily formation” that is aligned with the scene rehearsed (Schmidt, 2018, p. 231). In jam circles, dance couples negotiate the right to be the next dancing couple by means of visuospatial preparations that display couplehood, imminent entry, and attempts to occupy the performance space (Keevallik & Ekström, 2019). Participants in dance classes begin to form dance couples after directives by the teacher’s *practice projectors*, such as “let’s try it out” (Broth & Keevallik,

2014). To transition from instruction to performance during musical master classes, the teacher typically leaves the stage area, and the pianist may start to play (Reed, 2015; Szczepek Reed et al., 2013). After performance, to provide instructions, the teacher returns to the stage area, or the “engagement space” (see Reed & Szczepek Reed, 2013, p. 320). Further, the transitions from performance to discussion can also be achieved in verbal ways, through assessments (Reed & Szczepek Reed, 2013, 2014).

The transitions described in the prior paragraph are often achieved in a stepwise manner and reflect a change in activity. However, transitioning into performance may be achieved during much shorter time fragments than changing the overall activity from discussion to performance. For instance, Keevalik (2010, 2013a) has demonstrated how dance teachers use fragments of dance to fill grammatical slots in ongoing turns to demonstrate to students how they should achieve a dance practice. Similarly, Weeks (1996) discusses how an orchestra conductor uses *illustrative expressions*, essentially fragments of sung music, to instruct the orchestra how to play. Further, Szczepek Reed et al. (2013) show how singers differentiate between “local vs. restart-relevant directives”, the former being instructions on aspects of the performances that are to be briefly implemented within the discussion (or instruction) activity, rather than later in a restart of the performance activity. This thesis focuses on how smaller fragments from the performance activity are jointly and momentary achieved within the discussion activity. In this work, these fragments are conceived of as depictions of tentative performance bodies.

Activities reflexively shape the interaction while they are being shaped by it (Betz et al., 2020; Levinson, 1979). The overarching activity investigated in this thesis, scenic opera rehearsals, conditions the

interaction in several ways. Opera rehearsals are an institutional setting (Arminen, 2005) with predetermined roles and responsibilities¹³ of the participants, such as director, assistant director, and performer. These roles, and their associated rights to determine self's and other's local and future action, i.e. *deontic rights* (Stevanovic, 2018; see section 2.4.2 below) are frequently oriented to in interaction. The relative distribution of deontic rights between participants may vary between different forms of performing arts, and even between different settings of the same performing arts. Whereas the prominent and leading role of the director, or teacher, is more evident in some settings (see for instance Köhler, 2007; Messner, 2020; Reed, 2015; Reed & Szczepk Reed, 2013, 2014; Schmidt, 2018; Szczepk Reed et al., 2013), a highly collaborative and dialogic endeavour is evidenced in other settings (Norrthon, 2020). In more informal settings, such as musician friends informally playing together, there may be no evident leaders at all (Balantani, 2022). This thesis contributes to the understanding of how deontic rights are made manifest and distributed in interaction. Instructions and proposals from the director, and proposals from the performers, mutually inform and build on each other in a dialogical process of making decisions on performance bodies that extends over time, that is, interactional histories. The two following sections cover previous research on proposals and joint decision-making in interaction (2.4), and interactional histories (2.5).

2.4 Joint decision-making, proposals, and deontics

In the collaborative process of creating an opera production, the participants share ideas, develop them together and make preliminary

¹³ Or *membership categories* (Sacks, 1992).

decisions on which ideas to implement in the production. This manifests itself interactionally in the joint decision-making sequences, involving proposals from the performers, that are in focus of this thesis. This section targets how decisions are jointly achieved in social interaction by first reviewing previous research on proposals (2.4.1), before focusing on the orientations to deontic order that are manifest in joint decision-making processes (2.4.2).

2.4.1 Proposals

A proposal is a social action that puts forth an idea for a joint activity or project that will be of benefit to both proposer and recipient of the proposal (Couper-Kuhlen, 2014; Thompson et al., 2021), and whose actualization depends on the recipient's approval (Stevanovic, 2012). Proposals are distinguished from other directives, that is, "social actions that all count as an attempt to induce the recipient to perform (or not to perform) some action", for instance orders, invitations, suggestions, requests, offers, etc. (Stevanovic & Svennevig, 2015, p. 2; see also Couper-Kuhlen, 2014). Suggestions, for instance, although being like proposals in that they evoke some future action, do so only for the benefit of the recipient of the suggestion, who is also the one who ultimately makes the decision (Couper-Kuhlen, 2014). Suggestions and proposals have in common, however, that they both imply "mild attempts to determine future actions" (Stevanovic, 2012, p. 779), in contrast to, for instance, instructions and orders. In opera rehearsals, the activity and project that the proposals target is the collaborative achievement of an opera production, and the benefit of creating a successful production resides with the performers, director, and assistant director. Furthermore, for some proposed version of a performance body to be implemented in future

iterations of the production, it needs to be accepted by the other co-present participants.

Proposals may be distal/remote, targeting activities to be performed in the future, or proximal/immediate, targeting activities to be performed in the here-and-now (Houtkoop, 1987; Lindström, 2017; Stivers & Sidnell, 2016). Proposals in opera rehearsals have both a distal component, targeting arrangements for the future production, and a proximal, as the current state of the production is made available through depictions. What is proposed during the rehearsals, performance bodies, is often implemented in direct relation to the proposals through depictions (cf. Lindholm et al., 2020; Magnusson, 2021; Mondada, 2015, 2018c; Nissi, 2015, for similar cases with written artifacts).

Couper-Kuhlen (2014) argues that proposals are recognizable as such through their grammatical construction. Different grammatical formats for proposals have been explicitly targeted in literature on American English (Couper-Kuhlen, 2014; Thompson et al., 2021), Canadian English (Stivers & Sidnell, 2016), Polish (Zinken & Ogiermann, 2011) and Finnish (Stevanovic, 2013b). These different formats may index varying stances towards the proposals, depending on the likelihood of a proposal's acceptance or the kind of activity proposed, or display different deontic stances (see section 2.4.2 below). Although these grammatical constructions may be canonical in designing proposals, the latter may also be achieved in more implicit ways (Stevanovic, 2012). Proximal proposals are sometimes entirely embodied (Stevanovic & Monzoni, 2016; Stivers & Sidnell, 2016). Further, visuospatial resources may be used to make proposals in the absence of verbal descriptions, as evidenced by research on theatre rehearsals (Schmidt & Deppermann, 2022, 2023). The proposals in opera rehearsals are sometimes constructed using canonical proposal formats, but there are nevertheless many implicit ways

to make proposals of performance bodies, some not involving lexicosyntax at all (see Article II of this thesis).

As evoked above, proposals initiate processes of joint decision-making. Typically, proposals are responded to by means of the mutually exclusive categories of acceptance and rejection (Asmuß & Oshima, 2012; Houtkoop, 1987; Lindström, 2017; Stevanovic, 2012; though see for instance Yasui, 2013, on more nuanced ways of responding to a proposal). In general, there is a preference for accepting proposals – they are rarely overtly rejected (Houtkoop, 1987; Ekberg, 2011; Stevanovic, 2012, 2015; Thompson et al., 2021). Before being accepted or rejected, however, proposals characteristically move through a process of joint decision-making. Stevanovic (2012) identified three components in the process of arriving at a joint decision during meetings: i) access, ii) agreement, and iii) commitment. Her research shows that merely complying with a proposal does not make the decision joint: participants need to display access to the content of the proposal, agreement with the suitability of it, and commitment to carry out, in the future, what the proposal puts forth.¹⁴ Article III of this thesis focuses specifically on how participants at opera rehearsals go through these different stages of the joint decision-making process and how they are relevant for the turn design of proposals, and implementations of depictions in them.

When making and responding to proposals, participants reveal the relative roles, rights, and responsibilities they assign to one another in the decision-making process. The next section zooms in on the notion of *social deontics* and explains how it is relevant for understanding how the participants at opera rehearsals collaboratively create the production.

¹⁴ Although commitment to future action is not always explicit, or even mandatory (Huisman, 2002; Magnusson, 2021; Stevanovic & Monzoni, 2016).

2.4.2 Deontics

The notion of *social deontics* refers to the capacity of participants in interaction to decide over one's own and others' proximal and distal courses of action (Stevanovic, 2013a, 2015, 2018, 2021; Stevanovic & Peräkylä, 2012, 2014). This capacity is manifested in participants' deontic rights. It is determined by their deontic statuses – the authority that others assign to them, and their deontic stances – the authority that they themselves claim (Stevanovic, 2018). The participants in opera rehearsals, as in all face-to-face interaction where decisions are established, unavoidably display orientations to self's and others' deontic rights (cf. Stevanovic & Peräkylä, 2012). According to Stevanovic (2018, p. 375), deontic authority is “potentially at stake every time [a person's] behaviour makes relevant compliant, responsive behaviour from another person”. Orientations to deontic rights are reflected in the turn design of proposals – in how and when performers choose to depict ideas for performance bodies, and in how the director manages the performer proposals.

Although some work refers to deontic statuses as in part a real-world circumstance, through which certain people have preassigned authority based on, for instance, institutional roles or membership categories, deontic authority is negotiated in interaction (Stevanovic & Peräkylä, 2014; see also Asmuß & Oshima, 2012; Thompson et al., 2021). Deontic authority, as well as membership categories (Sacks, 1992) such as *director* and *performer*, must be “accountably talked into being” (Heritage, 1984a, p. 290). Authority is thus based on “*other people* treating someone's power to determine action as legitimate” (Stevanovic, 2018, p. 372, italics in original). Formal authority may be challenged and rejected by other participants (Ekberg & LeCouteur, 2015; Kent, 2012). Further, deontic authority is related to epistemic authority in that the

person who is most knowledgeable on some concern may be the person who is most suitable to decide on future actions on the same concern, and participants use knowledge as an argument when defending their deontic rights (Stevanovic & Peräkylä, 2012). There is, however, no one-to-one relationship between epistemic and deontic authority (see Stevanovic & Peräkylä, 2014; see also Stevanovic, 2018, on *epistemic-deontic incongruence*). In opera rehearsals, there is a hierarchical relationship between the director, assistant director, and the performers that the participants orient to in different ways and that conditions the nature of the interaction. The director is oriented to as having deontic authority in both proximal and distal matters during rehearsals. She is mostly oriented to as having the final word on performance bodies and on the interactional agenda of the scenic opera rehearsals (see next paragraph on *distal vs. proximal* deontic rights). The performers are generally oriented to as having authority in matters regarding the technicalities of song, and the assistant director in the practicalities of this opera house and over schedules. This deontic relationship between the participants is manifested in the types of social actions that they perform, where the director often instructs the performers, and the performers mainly make proposals. When performers propose performance bodies, they thereby do so in a situation with asymmetric recipient-tilted distribution of deontic rights (Stevanovic & Peräkylä, 2012).

Deontics divide between *distal* and *proximal*, that is, the right to determine future actions vs. the right to determine local sequences of actions (Stevanovic, 2015). The proximal deontic claim of a turn-at-talk may be accepted even in the face of rejection of the distal deontic claim of the very same turn, and vice versa (Stevanovic, 2015, 2021; Versteeg, 2018). There may thus be an acceptance of the content of a proposal, despite a rejection of the relevance of the topic in the here-and-now, or

there may be a rejection of the content of a proposal, despite an acknowledgement of the relevancy of the proposal (Stevanovic, 2015). As will be evidenced in Article II of this thesis, this holds for proposals in scenic opera rehearsals as well. In opera rehearsals, participants make both proximal and distal deontic claims when proposing ideas for performance bodies. The proximal claims involve making a proposal through a depiction in the here-and-now, whereas the distal claims involve the relevance of a depicted performance body for the production as it will materialize when it is ready for premiere.¹⁵

Prior research has mostly focused on how different verbal formats express deontic stance, and the “key locus of negotiation over deontic rights” has been argued to be the adjacency pair (Stevanovic, 2018, p. 370; see also Stevanovic & Svennevig, 2015). As evidenced above (see section 2.2.1), it is still debated whether visuospatial behaviour can be organized in adjacency pairs, and thus whether it can reveal orientations to deontic stances and statuses. However, some research has revealed visuospatial managements of deontic rights (see for instance Asmuß & Oshima, 2012; Barske, 2009; Clifton et al., 2018; Kuroshima, 2023; Stevanovic & Monzoni, 2016; Tranekjær et al., 2022; Tuncer, 2015; Van De Mieroop et al., 2020). This thesis adds to that research by investigating how participants negotiate rights to propose and depict through visuospatial resources. As is evidenced in Article II of this thesis, participants initiate depictions as joint activities through visuospatial initiations that could be considered *proximal proposals* of the very same activity (see also Stevanovic & Monzoni, 2016; Stivers & Sidnell, 2016; see also Kendrick & Drew, 2016, on *recruitments*). These proposals of the

¹⁵ This difference can be summarized through these fictive questions to the other participants: “Can I add a proposal/depiction to the interactional agenda?” (proximal claim) vs. “Can I perform behaviour X on stage to music?” (distal claim).

immediate joint activity of a depiction are in turn in the service of, i.e. a practice for, making proposals on performance bodies to be implemented in the production. It is argued that by initiating depictions, performers make deontic claims that the director accepts or rejects. The management of the deontic rights of making depictions as proximal joint activities is achieved through initiating and responsive relocations (see Article II of this thesis).

In summary, orientations to the deontic order of social interaction become prevalent in any setting where decision-making is at stake, and scenic opera rehearsals are no exception. Proposals open processes of *joint*, rather than unilateral, decision-making (cf. Stevanovic, 2012). By making a proposal, the performers make a deontic claim that gets granted, challenged, or even rejected by the other participants. The proposals, however, may also intersect with other social actions (Thompson et al., 2021; see also Rossi, 2018, on *composite social actions*) and carry multiple functions in the dialogic process (cf. Linell, 2009) of creating the production. Whereas they bring out new ideas, they simultaneously build on previous ideas and display understanding, and ultimately alignment, with them. This becomes particularly evident in depictions, as they increase epistemic access (cf. Stivers et al., 2011) to performance bodies in the making. Further, any depiction, or description, of an idea, unavoidably displays understanding of the production aesthetics thus far. In that way, the decision-making process becomes an accumulated shared history. The next section discusses previous research on interactional histories, and how these are relevant for the participants' endeavours at opera rehearsals.

2.5 Interactional histories

As evidenced above (section 2.4), proposals initiate joint decision-making processes in which performance bodies become elaborated and agreed upon in different ways. This thesis is interested in the unpacking and developing of performance bodies in proposals over time. Whereas some of this development may occur within proposal sequences (Due, 2016; Magnusson, 2021; Stevanovic, 2012; Yasui, 2013), other research has focused on how the design of turns changes over time, beyond the sequence, within *interactional histories* (Deppermann, 2018a).

On the very first day of scenic opera rehearsals, the performers, director, and assistant director do not yet have intersubjectivity on the aesthetics of the production. Over the course of the rehearsal period, however, they increasingly form shared beliefs on that aesthetics, the motives, and psychological natures of their characters and how they, as a specific collective, best create the production together. This shared interactional history (or *common ground* to borrow a term from psycholinguistics, Clark, 1996) begins to reflect in the turn design of different social actions during the opera rehearsals, and in the ways in which the participants speak about the production. In this work, such histories will be referred to as *joint decision-making (micro-)histories*. This section reviews previous literature on how social interaction is shaped by interactional histories.¹⁶

¹⁶ The study of interactional histories stems from, and is closely related to, studies on longitudinal CA (Deppermann & Pekarek Doehler, 2021). In this thesis, the term *interactional history* is used to refer to the phenomenon of shared accumulated experiences that condition the here-and-now interaction, whereas *longitudinal CA* is used to refer to the specific method with which shared experiences and interactional changes can be studied in interactional data.

Interactional histories are shared lived experiences beyond the local interactional sequence, that nevertheless influence that sequence, for instance in terms of turn design (cf. Deppermann, 2018a). Whereas some interactional histories span over several years (cf. Mondada, 2015, 2018c), others may refer to much shorter time spans such as weeks (Norrthon, 2019), days (Broth et al., 2017; Deppermann, 2018a), or even minutes (Schmidt & Deppermann, 2023). To refer to relatively short time spans of interactional histories (seven minutes in their case), Schmidt & Deppermann (2023) coined the term *interactional micro-history*. The histories explored in this thesis cover different time spans – the six-week-long rehearsal period of the opera production in question – but also histories of the development of specific scenes in the rehearsals that may accumulate over days, hours, or minutes. In Article III of this thesis, which has a special focus on interactional micro-histories, the cases analysed cover development over 10–20 minutes of rehearsal time.

Studies on interactional histories have mainly focused on instructions in pedagogical settings (Broth et al., 2017; Deppermann, 2018a; Chazal, 2020; Keevallik, 2020; Stukenbrock, 2021). This strand of research has demonstrated how instructions become shorter and syntactically less complex, more indexical, and ultimately, more infrequent overall, over time. The focus in this study is not on verbal instructions, however, but on proposals, and more precisely, on how depictions as a practice for accomplishing proposals develops over time.

A branch of studies on interactional histories has focused on interactional histories in creative settings and settings where performing arts is created. It has been explored how proposals develop dialogically over sequences (Yasui, 2013), as well as how participants gradually develop mutual understandings of aesthetic concepts (Deppermann & Schmidt, 2021b). Further, research on theatre rehearsals has investigated how

participants develop theatrical scenes over time (Hazel, 2018; Lefebvre, 2018; Norrthon, 2019; Savijärvi & Ihalainen, 2021). Much as over interactional histories in general, interaction in theatre rehearsals is characterized by a move from verbal instructions to embodied implementations over time (Norrthon & Schmidt, 2023; see also Deppermann & Schmidt, 2021b; Lefebvre & Mondada, 2023; Norrthon, 2019, 2021; Schmidt & Deppermann, 2023).

In all three articles of this thesis, I focus on how performing arts, namely performance bodies of an opera production, come into being in social interaction. I focus on agentive performer proposals on performance bodies, rather than on the instructions and implementations evidenced in prior research on theatre rehearsals (cf. Deppermann & Schmidt, 2021b; Schmidt & Deppermann, 2022, 2023). Much as in previous studies on theatre rehearsals, I look at the interactional practices that the participants deploy to collaboratively create the production, and on how these resources may change over time. To do so, I draw on Clark's (2016) notion of *depiction* (see also Schmidt & Deppermann, 2022). This concept will be outlined in the next chapter.

3. Depictions

A central concept for this thesis is the communicative method of *depiction* (Clark, 1996, 2016, 2019). In this chapter, I aim to describe the roots of this concept, relate it to phenomena that have been studied within the EMCAIL tradition, and argue for the relevance of the term depiction for the purposes of the current thesis, as well as for future studies on iconic communicative strategies in social interaction within the EMCAIL framework.

3.1 Depiction as a communicative method

Clark (2016) coined the term *depiction* and argued that it is a communicative method, alongside *description* and *index* (see also Clark, 1996), that is relatively undealt with in the subject of linguistics, and that should be integrated in theories of language processing. The tripartite distinction between *depiction*, *description*, and *index* has its origins in the semiotic model of Peirce (1955), that defines *icon*, *symbol*, and *index*. Whereas Peirce would refer to them as different relationships between a sign and its referent, Clark terms them communicative methods and thus stresses how the signs are deployed in communication. *Depiction*,¹⁷ or *icon*, is a method of communication/sign relationship that is based on similarity between the sign and the referent, such as a painted portrait of someone, or communicating by means of demonstrating what something looks/sounds/feels like. *Description*, or *symbol*, is a method of communication/sign relationship based on conventionality, i.e. the tacit agreement that X should mean Y, as with lexical semantics and

¹⁷ In earlier work (see Clark, 1996; Clark & Gerrig, 1990), *depictions* are referred to as *demonstrations*.

conventional/emblematic gestures. *Index* is a method of communication/sign relationship based on contiguity and proximity between sign and referent, such as a pointing hand locating something in the immediate space, or dark clouds conveying the possibility of rain. Depicting in communication is to *show* rather than to *tell* (describe), or *locate* (index) (Clark, 1996, 2016, 2019). However, signs and methods of communication are often composite, and depictions may have non-depictive (i.e. descriptive or indexical) elements (Clark, 1996; Peirce, 1955).

Depictions can occur in all human sensory modalities and in combinations of them (Clark & Gerrig, 1990). They can be visuospatial, such as a gesture depicting hitting an imaginary tennis ball with an imaginary racket, vocal/aural, such as a conductor producing a fragment of song to illustrate a specific voice quality for the choir, or haptic, such as a child pinching their caregiver in the arm to demonstrate how they were pinched by a sibling (cf. Clark & Gerrig, 1990; see also Reed, 2019, on *demonstrative touch*). Enfield (2005) argues that vocal/aural signs are prototypically symbolic whereas visuospatial signs are prototypically iconic. In this thesis, however, none of the modalities are conceived of as *a priori* better suited for any of the communicative methods, since many examples from the opera rehearsals, and other interactional settings (cf. Cantarutti, 2021; Ehmer & Mandel, 2021), have shown that depictions are often multimodal, and not *a priori* visuospatial.

To illustrate what a depiction is, Clark (2016, p. 324) offers a written example from a news magazine, in which a Hollywood director discloses to a journalist that, due to falcons resting on a rooftop intended as a filming location, they could not pursue the filming as planned. According to the news article, the movie director was reported to have said: “In L.A., they would have’ – He leveled a finger at some imaginary nestlings and made a gun-cocking sound” (Clark, 2016, p. 324). In this example,

the movie director produces a depiction, by leveraging a “finger at some imaginary nestlings and making a gun-cocking sound”, at the end of a clause that projects a verb phrase (for instance “shot those falcons”, cf. Clark, 2016). The example illustrates the role of depictions in communication and demonstrates how they can be integrated in linguistic clauses (see Keevallik, 2014a, 2015, 2018, on *syntactic-bodily units* in video-recorded naturally occurring data).

Clark (2016) uses what he labels *staging theory*, theatre as a metaphor, to understand depictions. He argues that “staging a scene [in depictions] is the same type of act that is used by children in make-believe play and by the cast and crew in stage plays” (Clark, 2016, p. 324). A depiction introduces a scene (see section 3.1.1 below) from a “there-and-then” for the recipient to perceive in the here-and-now. Staging a scene in depictions implies that the recipients become spectators to it (Clark, 2016, 2019). It contrasts with describing that scene, where the status of the other participants is rather conversational partners. In this work, in line with Clark (2016), the “there-and-then” of the depicted scene is understood as its referent or its *distal scene*. The distal scene is distinguished from the *proximal scene* – the communicative sign in which a baritone embracing a soprano becomes understandable as a father mourning a daughter. The distal and proximal scenes, in turn, are distinguished from the *base scene* – the resources in the here-and-now, such as prosodic alterations and spatial configurations, that participants use to construct a depiction. This work is particularly interested in the base scene of depictions: what are the multimodal resources participants use to construct a depiction and index a set of behaviours as depictive, and how are these resources responded to by other participants?

3.1.1 In favour of the term *depiction* in EMCAIL

The concept of *depiction* stems from a psycholinguistic research frame (Clark, 2016), and is in turn based on a concept from Peircean semiotics (Peirce, 1955). These research traditions have different epistemological frameworks and methodologies to EMCAIL research. Whereas the latter studies language and social interaction as it occurs in naturalistic data and takes a qualitative and emic approach to data analysis (Sidnell, 2010), psycholinguistics and semiotics instead work with theoretical and etic categories, often, though not always, using experimentative and quantitative methodologies (Chandler, 2017; Traxler, 2012). In this thesis, *depiction* is shown to be an emic category that is oriented to by the participants in opera rehearsals. Recipients of depictions treat them as scenes, in which the *origo* (Bühler, 1934/2011, see section 3.2.1 below) of the depicter (see below) is displaced from the depicter in the here-and-now to their character in a fictional there-and-then, and to which they position themselves as spectators or co-performers. Further, as will be evident in section 3.2.1 below, iconic phenomena have long been of interest for EMCAIL researchers, even though they have been described using different terminology. In this work, depiction is used as an umbrella term that encompasses all these different terms (see Schmidt & Deppermann, 2022, for a similar use of the term *depiction*). In coherency with the term *depiction*, the thesis coins the term *depictive speech* as an equivalent to reported speech, that is, speech in which the origo of the speaker is displaced into somebody else (real/fictional) at some other (real or fictional) place (see section 3.2.1 below). The term *depicter* is used to designate the participant who performs a depiction at a certain moment, whereas the other co-present participants who focus on the

depiction (either by watching it as spectators or by becoming recruited to play characters in the depiction), are *recipients* of it.

Depiction as a term has several affordances. This thesis looks at both visuospatial and vocal/aural aspects of depictions, and the term is neutral with regards to the modalities implied (not just vocal/aural as with reported *speech* or visuospatial as with *iconic* gesture). Further, what is represented in depictions, the distal scenes, is fictive and does not have a prior event on which it is based. Therefore, depiction is a suitable term, in contrast to the term *reenactment* that implies such a prior existing event (see also Stukenbrock, 2017). In contrast to the term *enactment*, that might otherwise suit the setting of opera rehearsals that indeed encompasses acting, depiction has the benefit of keeping consistency with Clark's (2016) framework and the theoretical roots of the concept of *iconicity* (cf. Peirce, 1955). This allows the thesis to contribute to Clark's (2016, 2019) ambition to provide a complete and holistic account of the role of iconic signs in human communication and language processing by facilitating comparisons between iconic phenomena in different social environments and activities. In addition, the term facilitates engagement with semiotic concepts that contribute to the understanding of depictions in opera rehearsals (see below).

Another important affordance of the term *depiction* for this work is its compatibility with the artistic setting of opera rehearsals, due to the artistic connotation of the term. In layman's terminology, works of art are typically described to *depict*, or as *depictions of*, chains of events, objects, emotions, states of mind, and the like.¹⁸ In that sense, the

¹⁸ Clark (2016, p. 324) explicitly highlights the artistic connotation of the term depiction, by putting forth that when people think of the term *depiction*, they "usually think of artifacts created at one place and time and exhibited at another – paintings, sketches, blueprints, maps, statues, movies, animated cartoons, TV sitcoms, or radio plays". He refers to these latter as *exhibited*

production, that is, the Italian tragedy at hand, is a depiction of a chain of events. The performance bodies that the participants create to constitute different scenes of the production are also depictions (of prototypes of mundane behaviour, see below, expressed in the aesthetic style of this production). However, in this work, to avoid confusion, the term *depiction* is reserved for the interactional *practice*, whereas *performance body* is used for the interactional *outcome*. In the global sense, however, performance bodies are depictions, and depictions are performance bodies. Depictions can therefore be argued to be both interactional practice and outcome.

In addition to Peirce's (1955) semiotic model, Clark (2016) draws on literary theories on poetic production originating in the works of Plato and Aristotle, via Bakhtin (1981) and Lodge (1990). The relationship between depictions and descriptions is described as the relationship between *mimesis* and *diegesis* in ancient Greek dramas (Clark, 2016; Clark & Gerrig, 1990; see Aristotle, 1994, 1997; and Plato, 1992). Whereas *mimesis* is showing/enacting parts of the drama, *diegesis* is describing it with words (cf. *enacted* vs. *narrated* parts of stories, Thompson & Suzuki, 2014). According to Aristotle, *mimesis* is a method to depict human action (*praxis*) that gives the audience of the drama a sense of human nature (Aristotle, 1994, 1997). It is such *praxis*, that the members of the opera ensemble aim to represent in the production. They do so by drawing on prototypes of (mundane) behaviours (see Article I of this thesis; see also Lefebvre, 2018) – their intuitive accounts of human behaviour. These are then portrayed in momentary fictional distal scenes, through depictions (cf. Clark, 2016).

depictions but claims that in “everyday discourse” people use their bodies to make *staged depictions*.

In this thesis, I also draw on a development of Peirce's (1955) tripartite semiotic model by Roman Jakobson (1971). Jakobson's semiotic model incorporates a fourth type of sign beyond icons, symbols, and indexes – *connotative signs*. These are like icons in that there is a similarity between the sign and the referent, but whereas this similarity is factual in the case of icons, it is imputed in the case of connotative signs. Connotative signs are characterized by a lack of external referent, that is, *extroversive semiosis*. Instead, they refer to themselves in what Jakobson terms *introversive semiosis*. A connotative sign thus references itself, which means that the sign and the referent merge (cf. reasonings above on the amalgamate nature of depictions and performance bodies). According to Clark (2016), however, this is not the case for depictions (icons). Rather, these follow the *pas-une-pipe*-principle,¹⁹ meaning that they are not what they depict.

According to Jakobson (1971), connotative signs are typical of music and abstract visual art, but they can also operate in other kinds of media and alongside extroversive semiosis (for instance in poetry). Article I of this thesis draws specifically on Agawu's (1991) development of Jakobson's concept of introversive semiosis in his work on music semiotics. He showed how musical phrases can create internal references and structures that mutually inform each other. The establishment of a musical key, for instance, condition the understanding of subsequent notes. Introversive semiosis is relevant for depictions in opera rehearsals as it will be argued that these reference both distal scenes *and* themselves as the performance bodies that reflect the aesthetics of the production. In

¹⁹ This principle has borrowed its name from René Magritte's famous painting entitled "Ceci n'est pas une pipe", in English "this is not a pipe". The painting features a painted pipe accompanied by the text that lends itself to the title of the painting. The statement serves to draw the observer's attention to the fact that the pipe on the painting is not a real pipe, merely a representation (depiction) of one.

performance bodies, i.e. artistic correlates of human behaviours, the latter can be tweaked *vis-à-vis* reality, for aesthetic effects. Performance bodies are *simulacra*, rather than *copies*, of human behaviours (Deleuze, 1969; see also Smith, 2006). The way this twist occurs is, in part, determined by introversive semiosis. Introversive semiosis is reminiscent of accumulated common ground over interactional histories (Deppermann, 2018), as previous interactional experiences of the production guide and condition the participants' behaviour when developing it.

Figure 1 summarizes the relationship between the terms described in this section. When creating *performance bodies*, i.e. the fragments of performances that taken together constitute the production, participants draw on intuitive understandings of human behaviour. They deploy the interactional practice of *depiction*, that draws on both intuitive accounts

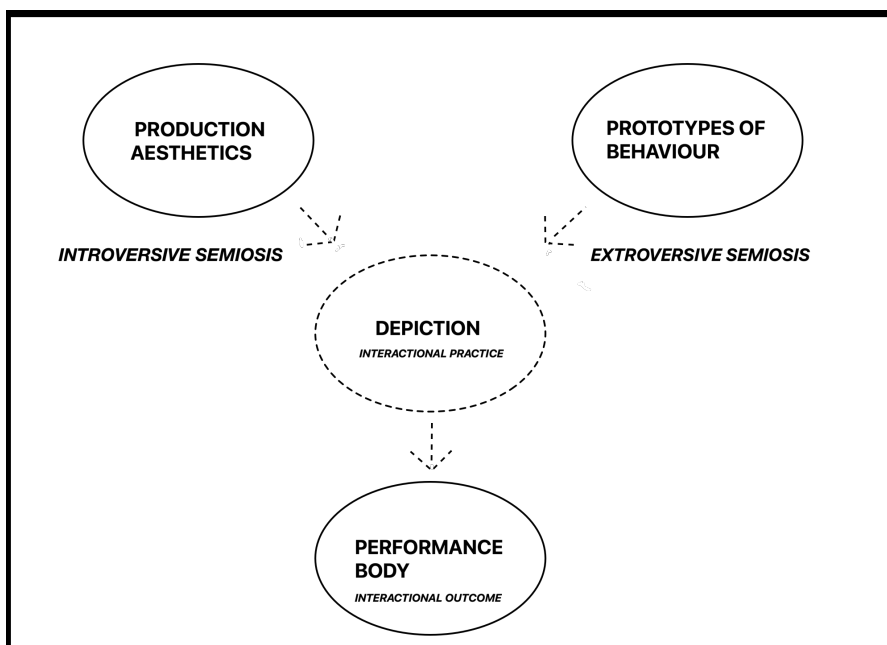


Figure 1. Overview of the concepts: *production aesthetics*, *prototypes of (mundane) behaviour*, *extroversive/introversive semiosis*, *depiction*, and *performance body* as they are used in this thesis.

of behaviour (*extroversive semiosis*), and understandings of the developing aesthetics of the production (*introversive semiosis*) to create, demonstrate, and negotiate ideas for performance bodies.

3.2 Depictions in interaction

Research on social interaction within the EMCAIL tradition has long been interested in interactional phenomena in which iconicity is involved. In this thesis, *depiction* is used as an umbrella term that encompasses all these phenomena. This section first reveals a range of iconic phenomena in social interaction (3.2.1), before reviewing research on how depictions are designed in interaction in terms of the base scene, that is, the multimodal resources deployed to construct it (3.2.2), and finally research that has targeted the question of *why* participants use depictions in interaction, and what communicative functions they may carry (3.2.3).

3.2.1 Iconic phenomena in interaction

In contrast to the psycholinguistic literature on depictions, where examples are mostly invented or anecdotal (cf. Clark, 2016, 2019; Clark & Gerrig, 1990), EMCAIL takes an empirical stance towards depictions and scrutinizes them on a millisecond-by-millisecond time frame in naturally occurring video-recorded interaction (on the benefits of recording for analysis, see section 2.1, on conversation analysis and interactional linguistics, above). This has led to detailed studies on the fine-grained particulars that participants deploy to construct depictions in time and space (see section 3.2.3).

Depictions have been shown to occur in mundane conversation (Cantarutti, 2021; Sidnell, 2006), telephone calls (Couper-Kuhlen,

2007), and text-mediated interaction (Tolins & Samermit, 2016). They have also been evidenced in many different institutional settings such as health care (Heath, 2002; Nishizaka, 2017; see also Due & Lange, 2020 for depictions in video-mediated healthcare interactions), martial arts training (Stukenbrock, 2014), dance training (Keevallik, 2018), orchestra rehearsals (Tolins, 2013), and theatre rehearsals (Schmidt & Deppermann, 2022). Examples of the diverse ways in which iconicity may figure in interaction are provided in Table 1.

Whereas some of these phenomena involve both visuospatial and vocal/aural resources, some involve only visuospatial or vocal/aural resources. Depictions associated with the vocal/aural modality are iconic *non-lexical vocalizations* and *reported speech*. Non-lexical vocalizations are sounds made in the vocal tract that are on a continuum with conventional words in languages, although some seem to be almost purely iconic and/or indexical and bear very little conventional meaning (Keevallik & Ogden, 2020; see also Dingemanse, 2020, on *liminal signs*). Not all non-lexical sounds are iconic, however. Non-lexical sounds that are iconic are so in virtue of mimicking sounds, shapes, or the like from the environment (cf. Dingemanse, 2011, on *ideophones*). An example of a non-lexical depiction in opera rehearsals is an occasion when the director imitates the character of a part of the music through the non-lexical vocalization “wham”.

Another form of vocal/aural depiction is *reported speech*, quoting the words of others (Holt & Clift, 2007; see also Jönsson, 2005). When reporting others’ talk, participants often manipulate parameters such as prosody and voice quality to depict *what* somebody else has said, and *how* they said it (Günthner, 1999). In the current work, the term *depictive speech* is used as an equivalent to reported speech (see section 3.1.1 above). Although the term itself focuses on the vocal/aural modality

(speech), research on reported speech has acknowledged that participants often also mobilize visuospatial resources towards the reported event (Holt, 2007). Therefore, some researchers rather speak of *(re)enactments* (see, for instance, Sidnell, 2006; Thompson & Suzuki, 2014), or *animation* (Cantarutti, 2020, 2021, 2022a, 2022b). Some of these refer to only visuospatial involvement in the depiction, whereas others include both modalities.

In pedagogical and creative contexts, the term *demonstration* is frequently deployed to designate when participants show, for instance, visuospatial behaviour, to one another, to instruct or create. As with, (re)enactments, demonstrations can be visuospatial and/or vocal/aural, depending on the nature of what is demonstrated. In orchestral rehearsals, demonstrations are typically vocal/aural (Tolins, 2013), whereas in dance classes, they are predominantly visuospatial (although they are sometimes accompanied by non-lexical vocalizations (Keevallik, 2014a). They are often used in settings where the aim is some kind of bodily performance, be it sports, arts, positionings of bodies in medical examinations, etc.

The term *depiction* has been used in EMCAIL research, and often refers to a phenomenon similar to demonstrations. In Schmidt & Deppermann (2022), depiction is used as an umbrella term encompassing the two different subcategories of *demonstrative depiction* and *illustrative depiction*, whereas Szczepek Reed (2021) presents *depiction* and *demonstration* as two different overarching categories (corresponding to Schmidt & Deppermann's, 2022, *illustrative depiction* and *demonstrative depiction*, respectively). Szczepek Reed (ibid.) demonstrates participant orientations to the difference between conceptual instructions through metaphors that target bodily actions (depictions), and concrete models to bodily actions (demonstrations).

Iconic gestures are visuospatial depictions, typically performed with the hands. They mimic their referents, or qualities of them (such as shape or speed) (cf. Schegloff, 1984; Streeck, 2009). They can be distinguished from other types of gestures, such as indexical (pointing) and conventional/emblematic gestures (e.g. “thumbs up” in western context) (Streeck, 2009).

Finally, other, less frequent terms have been proposed for iconic interactional phenomena (see Table 1). They are often used interchangeably with demonstrations and include *bodily shadowing*,²⁰ *bodily quoting*, *gestural (re)embodiment*, *illustrative expression*, *imitation*, *metaphorical hand gesture*, *mimicable body part highlighting*, *staging*, *sung quotation*, and *vocal imitation*. Shifts in *authorship* and *footing* (Goffman, 1981) as well as song as an interactional practice (Frick, 2013; Stevanovic & Frick, 2014) also relate to iconicity in interaction.

Terminology	Articles
(Re)enactment	Ehmer & Mandel (2021), Evans & Lindwall (2020), Fox & Robles (2010), Harjunpää et al. (2023), Holt (2007), Melander (2012), Pfeiffer & Weiss (2022), Sandlund (2014), Sidnell (2006), Streeck (2002), Stukenbrock (2017), Thompson & Suzuki (2014), Tolins & Samermit (2016), Tutt & Hindmarsh (2011), Wilkinson et al. (2010)
Demonstration	Ehmer (2021), Ehmer & Brône (2021), Evans & Lindwall (2020), Evans & Reynolds (2016), Goffman (1974), Haviland (2007), Hsu et al. (2020), Ivaldi (2016), Jokipohja & Lilja (2022),

²⁰ Bodily shadowing refers to a specific practice whereby a demonstration by a conductor teacher is directly adopted by his students in a quasi-simultaneous manner (Sunakawa, 2018; cf. Ehmer, 2021).

	Keevallik (2010, 2013a, 2013b, 2014b, 2015, 2018, 2021), Messner (2020, 2022), Nishizaka (2017), Råman (2019), Råman & Haddington (2018), Reed & Szczepek Reed (2013), Schmidt & Deppermann (2022), Stoeckl & Messner (2021), Stukenbrock (2012, 2014a), Szczepek Reed (2021), Szczepek Reed et al. (2013), Tolins (2013)
Depiction	Tolins & Samermit (2016), Brandenberger & Hottiger (2018), Keevallik (2018, 2021), Emerson et al. (2019), Szczepek Reed (2021), Urbanik & Svennevig (2021), Schmidt & Deppermann (2022), Harjunpää et al. (2023)
Iconic gesture	Greer (2013), Gudmundsen & Svennevig (2020), Hayashi (2003), Enfield (2009), Schegloff (1984), Streeck (2008), Urbanik & Svennevig (2021), Wu (2022)
Reported speech	Barnes & Moss (2007), Berger & Pekarek Doehler (2015), Bolden (2004), Burdelski (2015), Calabria (2023), Good (2015), Günthner (1999), Holt (2007), Holt & Clift (2007), Niemelä (2010), Reber (2021), Reed & Szczepek Reed (2014), Soulimani (2018), Wu (2016)
Animation	Cantarutti (2020, 2021, 2022a, 2022b)
Bodily quoting	Keevallik (2010)
Bodily shadowing	Sunakawa (2018)
Gestural (re)embodiment	Heath (2002)

Illustrative expression	Weeks (1996)
Imitation	Reed & Szczepek Reed (2013)
Metaphorical hand gestures	Sambre & Feyaerts (2017)
Mimicable body part highlighting	Due & Lange (2020)
Non-lexical vocalization	Albert & vom Lehn (2022), Balantani (2022), Brandenberger & Hottiger (2018), Hofstetter (2020a), Keevallik (2021), Keevallik & Ogden (2020), Tolins (2013)
Shifts in authorship and footing	Goffman (1981)
Singing in interaction	Frick (2013), Stevanovic & Frick (2014)
Staging	Iversen & Flinkfeldt (2020) ²¹
Sung quotation	Reed & Szczepek Reed (2013)
Vocal imitation	Messner (2022)

Table 1. Overview of EMCAIL articles on iconic phenomena in interaction.

Some of these iconic phenomena imply *origo shifts*, translocating the reference point of indexical expressions from oneself in the here-and-now to another real or fictive person in another real or fictive time and place (Bühler, 1934/2011; Stukenbrock, 2012, 2014). This displacement, elsewhere referred to as shifts in *footing* (Goffman, 1981), is mostly unproblematic for the recipient, who rarely mislocates the origo and thereby mis-ascribes social action (Cantarutti, 2020). This is related to

²¹ Original terminology is the Swedish *iscensättning* which translates closest to *staging*.

the pas-une-pipe-principle of depictions (see section 3.1.1 above). Depictions are not treated as the thing that they are depicting, but rather as a performance, joint-fictionalization, or play (Clark, 2016).²² For instance, during an opera rehearsal, when a performer utters the following phrase during a conversation with the director: “Why the hell would you do that?”, it is not meant to be taken as the performer questioning the director in a rude way, as herself, but rather as subtext of her character at that time (see Schmidt & Deppermann, 2022 on *subtext* in theatre rehearsals). Different multimodal resources serve to index the origo shift that is necessary for the participants to be able to ascribe the correct social action.

The depictions that are under scrutiny in this thesis are complex and heterotemporal multimodal gestalts involving origo shifts of the performers into their characters in a fictional time and space (*full-scale-demonstrations* in Clark’s, 2016, terminology). They are complex in the sense that they involve several interactional resources, from both the visuospatial and vocal/aural modality, such as gaze, voice, facial expressions, and body movements. They are heterotemporal in the sense that they do not always involve both modalities, or a multiple set of resources, and their nature changes over time, both millisecond-by-millisecond as they are constructed, but also over time as the performance bodies develop. The depictions in this work thus more closely resemble (re)enactments, reported speech, animations, and demonstrations than brief monomodal non-lexical vocalizations or iconic gestures.

²² In Clark & Gerrig (1990) this is described as the *non-seriousness* of depictions.

3.2.2 Depictions and descriptions in interaction

In EMCAIL research, depictions have often been discussed in relation to descriptions. Schegloff (1984) speaks of the *lexical affiliate* of iconic gestures (see also Urbanik & Svennevig, 2021), Weeks (1996) of *verbal vs. illustrative* expressions in orchestra rehearsals, Thompson & Suzuki (2014) of *narrated vs. enacted* parts of stories, and Stukenbrock (2017) about *explanations vs. enactments*. The lexical affiliate often follows the iconic gesture and disambiguates it (see also Stevanovic & Frick, 2014, on language disambiguating song), but depictions may also disambiguate descriptions, such as an iconic gesture specifying the meaning of a word (Urbanik & Svennevig, 2021). Depictions and descriptions should thus not be considered direct translations of one another, as they may contribute in distinct ways to sense-making, in supplementary, rather than redundant ways.

Descriptions may be used simultaneously to depictions, for instance to highlight certain aspects of the depiction (Stukenbrock, 2017). Schmidt & Deppermann (2022) show how the director of a theatre ensemble makes use of both depictions and descriptions when instructing the performers. In their material, depictions are always accompanied by descriptions in instructions, and descriptions can be used as a part of the depictions (cf. section 3.1 above, on *composite signs*). Descriptions in theatre rehearsals are used to index depictions within an utterance, to index a contrasting relationship between two depictions (see also Keevallik, 2010), and/or to highlight the focus of an instruction that is performed through a depiction. The present work adds insights into the relationship between depictions and descriptions. Rather than focusing on director instructions, however, the focus is on performer proposals with their unique contingencies.

3.2.3 The multimodal design of depictions

Clark (2016) discusses how people can use their bodies (voice, movements, face, posture, etc.) to build a unity of a depiction: “what producers are trying to do is stage scenes and not simply collections of aspects. If so, the aspects they combine should form coherent scenes, and they seem to” (ibid., 2016, p. 334). Within EMCAIL, such *collections of aspects* are conceived of as assemblages of multimodal resources in *multimodal gestalts* (Mondada, 2014a, see section 2.2.2 of this thesis). This section reviews previous research on how participants in interaction assemble multimodal resources so that they become recognizable as depictions.

It has been suggested that depictions are multimodal gestalts that become recognizable as invoking distal scenes through various vocal/aural and visuospatial cues (Ehmer & Mandel, 2021; see also Cantarutti, 2020, 2021; Couper-Kuhlen, 1996; Günthner, 1999; Holt & Clift, 2007; Niemelä, 2010; Sidnell, 2006). Depictions may be framed by lexicosyntactic quotatives (see, for instance, Eriksson, 1995). Keevallik (2013a; 2015; 2018) has shown how depictions can be embedded in linguistic clauses and occupy positions where syntactic elements are projected.²³ Often, however, other resources besides lexicosyntax are used to index a specific set of behaviours as depictive. Gaze has been pointed out as particularly important for the understanding of depictions as depictions, gaze withdrawal from the recipient indicating depiction, and gazing at recipient indicating end of depiction and return to the here-and-now

²³ Keevallik’s studies provide empirical examples of Clark’s (2016) account of how depictions may take the place of descriptions (words) in linguistic clauses. These are termed *embedded* depictions. Depictions may also be *indexed* by means of deictic expressions within linguistic utterances, produced *adjunct* to linguistic utterances, or produced *independently*, in the absence of linguistic utterances (see Clark, 2016). Keevallik (2018) argues, in a similar vein to Clark (2016), that visuospatial resources need to be included and accounted for within any conceptualization of language.

(Pfeiffer & Weiss, 2022; Sidnell, 2006; Thompson et al. 2014). In the case of reported speech, prosodic alterations (changes in pitch height, loudness, voice quality and articulation), sometimes present even before the onset of the fragment of reported talk, serve to index the origo shift that the depiction implies (Cantarutti, 2020, 2021; Günthner, 1999). Further, visuospatial configurations such as changes in posture or facial expression “prepare the stage” (Streeck, 1995) for depictions that become ultimately realized in combination with vocal/aural resources (see also Cantarutti, 2020; Ehmer & Mandel, 2021).²⁴

As evidenced in section 3.2.3 above, depictions may occur in the visuospatial modality alone while the vocal/aural modality is mobilized for some other purpose. A participant may, for instance, depict something with visuospatial resources while describing the depicted scene with vocal/aural resources (see, for instance, Stukenbrock, 2012, 2017). The visuospatial and vocal/aural modalities may in that way orient to different perspectives (observer’s vs character’s). These incongruent orientations serve different interactional functions, as also evidenced by Goodwin (2007b) (see section 3.2.4 below). The distal scene of depictions is thus not a static entity, it may transform over time as participants organize their different interactional resources towards the here-and-now and the there-and-then to different degrees. In other terms, the “animation space”, or the *distal scene*, has “fuzzy boundaries” (Cantarutti, 2020, p. 124).

Depictions may also be multimodally achieved across participants. Research has shown how depictions in mundane interaction are frequently responded to by means of further depictions, in what has been

²⁴ This reflects a more general feature of visuospatial behaviour to project and prepare for the vocal/aural modality (see Deppermann & Streeck, 2018; Streeck, 1995; Stukenbrock, 2018; see also section 2.2.2 above).

termed *co-animation* (Cantarutti, 2020, 2021, 2022a, 2022b), *co-enactment* (Ehmer & Mandel, 2021) or *collaborative reported speech* (Calabria, 2023), and what will here be referred to as *co-depiction*. In these, the recipient assumes the same origo as the original depicter by coordinating various multimodal resources in a way that matches those mobilized by the depicter in first position. Other studies have shown how recipients of depictions are recruited as co-performers, assuming a different origo in the same distal scene, that of another character (Keevallik, 2014b, Stukenbrock, 2012; Thompson & Suzuki, 2014). One aim of this thesis is to provide an account of the collaborative nature of depictions. This is particularly in focus in Article II of this thesis. The ways in which recipients respond to depictions in opera rehearsals partially differ from previous accounts of co-depiction. As reported in Article II, and much has been evidenced by prior research on co-depictions (see above), the recipients of depictions in opera rehearsals may co-depict by assuming the same origo of the original depicter but in a *parallel* distal scene, or alternatively, by assuming *another* origo as the original depicter in the *same* distal scene. However, in opera rehearsals, a third way of responding to a depiction, that has not been previously reported in the literature on depictions, is by explicitly positioning oneself as a spectator to the staged scene. In this thesis, I argue that this contribution is an integral part of the depiction, although the recipient who positions themselves as a spectator is not a part of the distal scene of the depiction, but merely present in the base and proximal scenes.

3.2.4 Interactional functions of depictions

In social interaction depictions are deployed as practices in the accomplishment of different social actions – they carry different interactional

functions. Whereas some of these functions are unique to the setting in which the depiction is employed, such as demonstrating an incorrect performance of a dance step in dance training (cf. Keevallik, 2010), others, such as making sensory qualities directly available for other participants in interaction (cf. Ehmer, 2021; Harjunpää et al., 2023; Iversen & Flinkfeldt, 2020), are observed across different settings. This section reviews accounts of different interactional functions that depictions may carry, in both mundane and institutional settings, as well as generally, before concluding on the interactional functions of depictions in opera rehearsals.

In everyday conversation, depictions are often deployed in storytelling (see Berger & Pekarek Doehler, 2015; Holt & Clift, 2007; Thompson & Suzuki, 2014). They are frequently implemented in climaxes of stories to illustrate certain aspects of the ongoing story (Cantarutti, 2020; Couper-Kuhlen, 2007; Ehmer & Mandel, 2021; Holt, 2007; Sidnell, 2006; Thompson & Suzuki, 2014). Co-depictions in response to tellings that involve depictions allow participants to affiliate with the stance conveyed in the telling (Cantarutti, 2020).

In instructional settings, depictions serve to demonstrate that which is taught, either by serving as models that the learners can directly copy with their bodies (cf. Ehmer, 2021; Keevallik, 2010; Sunakawa, 2018), or by functioning as metaphors that aim to deduce certain behavioural qualities (cf. Szczepek Reed, 2021). Depictions are also deployed in creative settings, either to instruct, to propose ideas, or to implement instructions and display understanding with them (Due, 2016; Norrthon & Schmidt, 2023; Schmidt & Deppermann, 2022, 2023; Yasui, 2013).

In general, depictions are efficiently used to illustrate matters that are not directly available to others in the here-and-now, such as subjective experience, emotion, stance, and intention (Couper-Kuhlen, 2007;

Harjunpää et al., 2023; Iversen & Flinkfeldt, 2020; Tolins & Samermit, 2016). They allow for joint pretence, imagination, and play (Cantarutti, 2020, 2021; Ehmer & Mandel, 2021; Holt, 2007; Stukenbrock, 2014a), and facilitate heightened focus, joint attention, and co-participation (Cantarutti, 2020; Ehmer, 2021; Sidnell, 2006; Thompson & Suzuki, 2014). Further, depictions can accomplish simultaneous orientations to a here-and-now and there-and-then (see also Clark, 2016). When depicting, the depicter can coordinate their multimodal resources so that some, for instance movements, depict, whereas others, for instance facial expression, simultaneously convey a stance to what is depicted (Berger & Pekarek Doehler, 2015; Cantarutti, 2020; Goodwin, 2007b; Günthner, 1999; Holt & Clift, 2007; Niemelä, 2010). In opera rehearsals, depictions involve insertions of performance within the activity of discussion and allow for simultaneous orientations to here-and-now (discussion) and there-and-then (the production under creation as well as the distal scenes that it represents).

It has been proposed that depictions reduce speaker agency and thereby accountability and responsibility (see Reber, 2012; Stevanovic & Frick, 2014). The origo displacement, at least partially, allocates the authorship to a person other than the depicter (see discussion in Cantarutti, 2020). What the depicter is saying/doing is not entirely attributed to them, and therefore they do not bear full responsibility for it. Further, depictions are often vague and ambiguous in comparison with descriptions (Messner, 2020; Reber, 2012; Stevanovic & Frick 2014). The correct interpretation of depictions may depend on the ability to draw inferences (see Deppermann, 2018b). In the face of possible challenge, the depicter may reject responsibility for an inference that another person has drawn, and thereby avoid taking responsibility for it. Cantarutti (2020), however, partially rejects the idea that depictions reduce agency

and responsibility, as the displays of stance that they invoke are often treated as accountable. She (ibid.) prefers the term *detachment* over *reduced responsibility*.

In this thesis I am particularly interested in how depictions are used by the performers to accomplish proposals. The performer proposals that are in focus differ from previous accounts of instructions and implementations of instructions in theatre rehearsals (cf. Schmidt & Deppermann, 2022, 2023), in their agentive nature. They are agentive claims to take up space on the interactional agenda of the rehearsals. By initiating depictions as part of proposals, the performers implicitly claim to have the right to momentarily introduce a fragment of performance on the discussion agenda, and by formatting their proposals as explicit proposal turns (in contrast to the implicit proposals described in Schmidt & Deppermann, 2022, 2023), they induce normative constraints of conditionally relevant uptakes of the proposal.

Depictions in opera rehearsals thus accomplish turns with high speaker agency and accountability. As initiations of depictions request joint attention from the co-present participants, depictions in opera rehearsals cannot be said to involve reduced agency (see discussion above). They are attempts at creating performance bodies that, when made available, are vulnerable to assessments from the other co-present participants. As expressed by Stevanovic (2015, p. 95), the participants' "senses of self-esteem as contributing participants are at stake at every moment of interaction". This, however, does not mean that introducing a depiction on the interactional agenda is a controversial matter. In contrast, depictions are crucial for the decision-making process at opera rehearsals. They make that which is proposed available to the other participants in interaction and constitute a way to collaboratively create performance bodies. As Article II of this thesis shows, however, depictions may be

more or less relevant on the interactional agenda at different moments of the creative process, and it is mostly the director who ultimately manages this agenda.

The interactional functions of depictions in opera rehearsals are expanded on and discussed in chapters 5 and 6 of this thesis. Before that, however, an account of the methodology and material of the thesis will follow in chapter 4.

4. Method

This chapter explains the method and material of the thesis. Section 4.1 outlines the main method, *multimodal interaction analysis*, as it has been described by Broth & Keevallik (2020) and other methodological accounts, and as it is deployed in the present work. Thereafter, the details of the material of the thesis are presented (section 4.2), followed by an account of the consent procedure and ethics discussion (section 4.3).

4.1 Multimodal interaction analysis

Whereas this thesis is informed by semiotic and psycholinguistic concepts (see section 3.1.1), the main method is video-based multimodal interaction analysis (Broth & Keevallik, 2020). This method is in turn grounded in ethnomethodologically informed conversation analysis (EMCA, see chapter 2 of this thesis). Multimodal interaction analysis allows for a fine-grained scrutiny of the base scenes constituents of depictions and thereby has the potential to offer a conceptualization of depictions as multimodal and interpersonal interactional practices. The aim is to zoom in on depictions as they are described by Clark (2016), and their relationship to descriptions, on a millisecond-by-millisecond time frame, to reveal how they are constructed in time and space and made relevant and oriented to by other participants. The thesis aims to examine depictions in the site where they occur, namely social interaction, and gain a dialogic (Linell, 2009) perspective on them. Therefore, naturalistic data, and next-turn-proof procedure as a core tool for the analysis, are particularly suitable (see 4.1.3 below). The social interaction that is in focus for this study is 20 hours of video recorded scenic opera rehearsals from one Swedish opera production (see section 4.2).

According to Clark (1996, p. 160), in conversation, “most signals are discrete events that leave no physical trace. Words and gestures are audible and visible only while they are produced. This is unlike many of Peirce’s signs, such as the painting or the weathercock, which are static and open to repeated viewing”. With multimodal interaction analytic tools such as transcriptions and video recordings, however, objects that are “static and open to repeated viewing” are *de facto* created out of ephemeral phenomena such as gestures and combinations of phonemes into words. Interaction analysis thus has the potential to analyse ephemeral signs that normally escape the scrutiny that freezing and replaying (Sacks, 1984b) allow for. In the following sub-sections, I will explain the core features of multimodal interaction analysis and how the method has been deployed in the present work.

4.1.1 Fieldwork and video recording

The micro-details of social interaction have been shown to be crucial for the understanding of how participants in interaction make sense of one another. Multimodal interaction analysis must therefore be carried out on material that can be replayed, decelerated, and frozen in time, such as video and audio recordings (Sacks, 1978; Broth & Keevallik, 2020). When studying multimodal interaction specifically, i.e. focusing not only on turns-at-talk,²⁵ but on how different interactional resources such as gaze, gestures, facial expressions, posture, and spatial configurations contribute to sense-making, video recordings must be opted for. It is crucial that these video recordings capture as much as possible of the setting as the researcher cannot pre-emptively decide which aspects of bodily

²⁵ In the early days of CA and IL research, audio-recordings were the go-to material but with technical advances, researchers came to rely more and more on video-recorded material, see chapter 2 of this thesis.

conduct and the surrounding material become relevant for sense-making. Even the slightest and seemingly unimportant detail may in fact be crucial for how the participants structure their actions. Several cameras targeting different angles of a setting may be necessary to capture, for instance, facial expressions and gestures of all the participants involved. Some interaction may have to be captured by a moving-camera man, to follow the participants' movements (Mondada, 2006). Supplementary audio recordings may sometimes be necessary, depending on the sound conditions in the setting the researcher wishes to capture (Broth & Keevallik, 2020).

To assure documentation of all the relevant details of the interaction with video (and audio) recording technology, recordings should be preceded by preparatory fieldwork that serves to familiarize the researcher with the setting (Mondada, 2013). Such preparatory fieldwork ensures a “good enough record of what happened” (Sacks, 1984b) in the interaction. To prepare for video recordings in this study, I made initial observations of the rehearsals of an opera production, to facilitate acquaintance with the environment of scenic opera rehearsals in this particular opera house. Thereafter, I contacted the assistant director of another opera production to request permission to follow the rehearsal period of an entire opera production. This request was accompanied by recounts of the purpose of the study and its methodology and was subsequently granted. Before video recording, I undertook a few days of additional preparatory fieldwork to familiarize myself with the specific production and ensemble, both scenography requirements and work methods.

After obtaining informed consent from the participants (see section 4.3 below) video recordings commenced. A hand-camera on a tripod was placed next to the “control room” (see section 4.2.2 on the rehearsal

space below), to capture the main rehearsal area (corresponding to the “stage”) from a left-rear angle. The rest of the space was captured with a mobile hand-camera. When recording with the mobile hand-camera, I mostly stood behind the director’s table, to reduce invasiveness. However, I sometimes moved around in the room when appropriate to accurately capture the interaction. Due to the relative proximity of the participants to the mobile camera, no additional audio recording was necessary.

In this thesis, the analysis of video material (see section 4.1.3 below) is complemented by ethnographical observations and interviews, to create a deeper understanding for the context of opera in general, and scenic opera rehearsals in particular. Ethnographical information is conceived of as important for EMCAIL analyses as it may provide the researcher with important material that contributes to the subsequent video analysis. It may, for instance, help decipher indexical references, and aid in the understanding of context-specific bodily conduct (see Depermann, 2013). In addition to the observations made during the fieldwork, and the informing conversations with the members of the ensemble, one opera singer and one opera singer student were interviewed in more depth on two separate occasions, to obtain information about the particular opera house, the rehearsal procedure of an opera production, and general characteristics of the professional life of an opera singer. All ethnographic information was documented in the field notes.

The recording process was abandoned once the ensemble moved from the rehearsal hall to the main stage in the opera house, and as the character of the rehearsals transformed (see material section 4.2 below). However, fieldwork progressed with observations and notes. During the final week of the rehearsal procedure, I attended one of the dress rehearsals and the pre-performance talk given by the dramaturge, director,

conductor and one of the lead singers. The aim was to deepen the understanding of the artistic vision behind the production. In addition, I was provided a copy of the video-recorded premiere of the production, to facilitate the understanding of the ensemble's aesthetic vision.

4.1.2 Transcription

When the recording procedure has been terminated, it is customary to transcribe the recorded material (Broth & Keevallik, 2020; Park & Hepburn, 2022). Within EMCAIL, transcriptions are commonly achieved with a detailed system developed by Gail Jefferson (2004), refined to incorporate multimodal detail by Lorenza Mondada (2018) (see Appendix 1 for transcription conventions).²⁶ These transcription systems capture the fine-grained details of turn-construction, turn-taking and multimodal conduct.

Transcribing the video data prepares the material for analysis and presentation in written artifacts.²⁷ However, the researcher must perform analytical work to be able to transcribe (see Ochs, 1979). According to Hutchby and Wooffitt (2008, p. 83) the aim of the transcript is not just “accuracy of detail” but to capture those features of a recording that are interactionally relevant and therefore important for the analysis. As the transcription cannot cover all details of the interaction (see Bogen, 1999, for a critique of the overly detailed CA-transcripts), it becomes necessary to judge what is relevant for the participants (see section 4.1.3 on analysis below) at the transcription stage (this is particularly true for the

²⁶ In addition, transcription of phonetic details can be successfully captured with the use of the GAT2 system for transcription (Selting et al., 2011).

²⁷ Indeed, transcriptions are a tool to present the material in text, such as scientific publications, as video recordings cannot be included in printed text. However, online articles may sometimes provide the reader with links to video clips.

multimodal details of interaction, see Deppermann, 2013). Transcribing everything in detail from the beginning may, on the other hand, help the researcher identify phenomena of interest, and details that are found to be irrelevant may be removed at a later stage of the transcription process (Ashmore & Reed, 2000; Ten Have, 2004; see also Hutchby & Wooffitt, 2008).

In this thesis, moments of the data that were judged interesting (see section 4.1.3 below) were transcribed in great multimodal detail. In fact, the detailed multimodal transcription of these moments allowed for the discovery of several interactional details that would come to be crucial for the analysis in the articles, such as the sometimes subtle relocations to depict, described in Article II of this thesis.

4.1.3 Analysis

When the material has been prepared by video recording and transcription, the analytical procedure continues with a scrutiny of the interactional details (for instance, turn-taking, turn design, social actions). It is important to note that analysis should not be conducted based on transcripts alone. Instead, the researcher is recommended to move back and forth between video recordings and transcripts when conducting analysis (Hutchby & Wooffitt, 2008).

The goal of the analysis is “to describe the intertwined construction of practices, actions, activities and the overall structure of interactions” (Stivers & Sidnell, 2013, p. 2). This aim is neatly captured in the canonical question of “Why that now?” (Sacks, 1992); what evidence is there in the material for why somebody behaves/communicates in a certain way at a particular moment in time? The response to this question is always

grounded in emic orientations, that is, how the other participants treat that behaviour (see section 2.1 on the *next turn proof-procedure*).

The analytical procedure should be characterized by *unmotivated inquiry* (Couper-Kuhlen & Selting, 2017; Maynard, 2013), where the aim is to “approach talk and interaction in a fashion that is absent of (commonsensically- or theoretically derived) presupposition, to the degree that it is possible, and recognizing that inquiry is completely freestanding” (Maynard, 2013, p. 19; see also Ford et al., 1996). However, some interactional research departs from a form, or “vernacular action” (Clayman & Gill, 2004, p. 596-597), especially within the interactional linguistics realm (see, for instance, Heritage’s work on ‘oh’, Heritage, 1984b, 1998, 2002).

Once a phenomenon of interest has been identified (in this case depictions in the service of instructions and proposals), the analysis advances with detailed scrutiny of cases of the phenomenon. This step of the analytical procedure typically aims at developing a collection of cases where interactional conduct is organized in patterns, although sometimes a single case can instead be discussed in greater detail (Fox et al., 2013).

The collection will sometimes include deviant cases, cases that are “noticeable departures from the norm” (Couper-Kuhlen & Selting, 2017, p. 25). These are incidences that go against the pattern that has been identified in the collection, and that are often treated as deviant by the participants themselves. These cases often inform the analysis, and to analyse them has been a core feature of EMCAIL since its early days (Sidnell, 2010). Documenting aspects of accountability is thus a way to cement the robustness of an analysis; if a lack of (a certain) response is treated as accountable on the part of the recipient, this serves to prove the

normative organization of that adjacency pair (see section 2.1), for instance proposal-acceptance/rejection.

In this thesis, the analysis was not entirely guided by unmotivated observation. As this thesis is a part of a research project on how non-lexical vocalizations figure in interaction (cf. Keevallik, 2021), the initial aim was to investigate different vocal practices and their role in the creative process during the rehearsals. The 20-hour material was scrutinized for instances of vocal practices such as song, reported speech, non-lexical vocalizations, and prosodic alterations. However, the mindset towards these vocal phenomena once identified was inductive, which resulted in the steering of the focus of the thesis towards holistic multimodal depictions, rather than vocal depictions alone, or even more broadly, different vocal phenomena.

During the initial analyses, a pattern emerged where the participants deployed vocal depictive practices when proposing, instructing on, and creating what they were going to do on stage in coordination with the music. This guided further readings on the topics of reported speech and song in interaction. In close scrutiny of the data, however, and in exploring earlier research on reported speech and the related phenomenon of (re)enactments (cf. Sidnell, 2006), it became evident that the vocal practices were intimately tied with concurrent visuospatial behaviour. The focus of the thesis then became these holistic multimodal gestalts that invoked distal scenes in the here-and-now interaction, that is, *depictions* (for the choice of concepts and terminology, see section 3.1.1. above).

4.2 Material

The material of the thesis consists of ethnographical observations, interviews, and video recordings of the rehearsal procedure of an opera production in one of Sweden's opera houses. The main period of material collection took place during six weeks, corresponding to the rehearsal period of the opera production. The video recordings constitute the main material of the study. They cover a period of five days of scenic opera rehearsals, during the six-week-long scenic rehearsal period. The recorded segment corresponds to a period in the middle of the scenic rehearsal procedure. In total, 20 hours of scenic opera rehearsals were video recorded. The languages spoken during the rehearsals were Swedish and English, although Italian words and phrases, mainly from the libretto, also occurred. Further details of the material, its participants, locations, and other additional circumstances, are outlined below.

4.2.1 Participants

The participants of the study were individuals who were present during the opera rehearsals: members of the opera ensemble, the artistic team, and the production team. The opera ensemble consists of soloist performers with different degrees of participation; some have leading roles and sing/act extensively during the performance, whereas others have smaller parts and consequently sing/act less.

In the production that this thesis focuses on, there were 14 soloists, a male choir (27 individuals) and eight stand-ins. Members of the artistic team were the director, conductor, scenographer, light designer, costume designer, and dramaturge. Members of the production team were the producer, assistant director, stage manager, light master, sound engineer, two rehearsal assistants, two prompters, two text machine

operators, translator, language coach, choirmaster, assistant choirmaster, and choir coordinator, as well as production assistants (one for props and one for the stage). The music was played by an orchestra associated with the opera house.

Not all team members were present during the rehearsal period, and in the video-recorded material only the following appeared: director, assistant director, soloists, choir members, rehearsal assistants, conductor, light designer, scenographer, stage manager, language coach, choirmaster, assistant choirmaster and production assistants. Other people involved in the production, but who do not feature in the recordings, and thus cannot be qualified as participants of the study, were planning coordinator, communicator, program director, and press officer, as well as head of costume and people responsible for women's and men's costumes respectively. The participants who are in focus of the analyses in the three articles in this thesis are director, assistant director, and the soloist performers.

4.2.2 The rehearsal space

The rehearsal space is located on the ground floor of a large hangar-like building. It is spacious, with a high ceiling, and the décor of the production can be installed in it. At one end of the hall, tables are placed, where the director, assistant director, stage manager, other members of the artistic/production team, or other observers may be seated during performances. The performances are positioned spatially as if these tables were in the place of the audience. Next to the tables there is a small control room, mainly used as a room of rest for the members of the ensemble. At the other end of the space (to where the performers normally have their backs while performing) there is a seated gallery that can be pulled out

and used for certain purposes such as introductory scenic meeting and first-time opera-orchestra rehearsals. On one side of the room there are hangers for clothes and shoes, as well as a space corresponding to the backdrop. On the other side there was a similar space corresponding to the backdrop of that side. The hall was heavily lit at most times, although the light could be, and was occasionally adjusted, upon requests of the members of the ensemble.

4.2.3 The opera house

The rehearsals that were observed in the opera house took place on the main stage. The performers and some other team members were either on the stage or in the backdrop during the rehearsals. The director, assistant director, scenographer, and light designer were mostly seated at provisional tables set up on the seventh row of the auditorium, but they also frequently moved on and off stage, especially the director and the assistant director. The stage manager was seated in a control room in front of the stage, but behind the stalls. Light and sound engineers were seated behind the stalls. Rehearsal assistants were seated by the piano in the orchestra pit when playing, and on rows one and two of the stalls when observing. Other team members and observers were present from time to time, usually sitting in rows one to eight one of the auditorium. The rehearsals in the opera house usually took place in dim light or complete darkness except for stage lighting. The director and assistant director used a microphone to communicate from the stalls to the team members on stage.

4.2.4 The opera production

The production studied in this thesis is a famous opera; a tragedy of three acts. The libretto of the opera is in Italian. The plot centres around a father, his daughter, and the lover of the daughter. At the end of the opera, the daughter tragically dies as she tries to protect her lover, a lover that the father has deemed unworthy of her due to his morally questionable character. The production was two hours and 20 minutes long including one interval.

The production was referred to as a “new production” by team members, meaning it had never been put on at the opera house before by this particular director and with this particular team. The decision to put on this opera was made approximately two years before the premiere. The first production and technical meetings were held around a year and a half before its premiere. These latter were not studied in this thesis, however, as the fieldwork ranged from first scenic meeting to premiere.

4.2.5 Producing an opera

Producing an opera involves technical/scenography aspects, musical rehearsals, and scenic rehearsals. Generally, the first technical meeting, when the director and team meet the technical staff to discuss the technical details of the production, will take place a year before the scenic opera rehearsals start. The choir normally starts rehearsing together about half a year before the scenic rehearsals start. Soloists may start rehearsing up to a year before the commencement of the scenic rehearsals. The first rehearsal with the orchestra occurs a few weeks before the premiere, and the orchestra is thereafter present at some, but not all, of the stage repetitions. This thesis focuses on scenic rehearsals and the

collaborative work of creating performance bodies. The orchestra rehearsals and technical stage work are not analysed.

4.2.5 The scenic opera rehearsal

The scenic rehearsals begin with an introductory meeting, where the artistic team meets the production team and ensemble, to present their vision for the production. The length of the scenic rehearsal period varies from production to production, but rarely extends beyond two months. This is shorter than most theatre rehearsal periods, as the temporality of the opera piece is already predetermined by the music. The first part of the scenic rehearsals always takes place in the rehearsal hall where the director works closely with the soloists, choir, stand-ins, dancers, and other potential participants on stage. This is where the performance bodies are created, although they may potentially be subject to change during the rehearsals in the opera house. For the opera production in focus of the present study, the scenic repetitions took place over three and a half weeks in the rehearsal space, and subsequently two weeks at the opera house, before the premiere. The rehearsals in the opera house are often more technical in their nature and aim at finding correct positions on stage. It is the scenic rehearsals in the rehearsal space that are studied in this thesis.

Two primary sub-activities occur within the larger activity of a scenic opera rehearsal in the rehearsal space: 1) performance, and 2) discussion. Performances are formally initiated and ended, mostly by the director (cf. *relinquishing move*, Reed, 2015), and are performed to music (piano, or orchestra if available). Performance is a manifestation of the aesthetic product in its primary expression. Discussion occurs in between the performance segments. The conversation during the

discussion phase typically targets the performance segments, but it may of course intermittently be of a more informal character (cf. Usatch, 2002, on *topic-* vs. *activity-*oriented talk). The shift between performance and discussion implies a shift in general temporality and projectability of the interaction. However, fragments of performance (although reduced, for instance without musical accompaniment) may occur within the discussion activity. These fragments are in focus of the present study, due to their depictive nature.

4.3 Ethics

Video-recorded material is crucial for the analysis of this thesis (see chapter 4 on methodology), which is the only acceptable motivation for video recording study subjects according to the ethical guidelines *Good research practice* (Swedish Research Council, 2017). This section first describes and motivates the procedure with which consent for video recording and research was obtained from the participants (section 4.3.1), before discussing measures taken to protect the anonymity of the participants (section 4.3.2).

4.3.1 Consent

The General Data Protection Regulation (GDPR), Article 4(11), defines consent as “any freely given, specific, informed and unambiguous indication of the data subject’s wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her” (European Data Protection Board, 2020, p. 6). As the material of this thesis is video recordings of the participants, personal data are processed, and GDPR compliant consent from all participants is necessary. The obtaining of informed consent in

research on social interaction based on video-recordings differs from consent procedures in experimental, and even other types of qualitative research, in that “matters of consent are dealt with inside the research process itself” (Speer & Stokoe, 2012, p. 55), as consent are often continuously discussed in the recordings.

The consent procedure of this research was planned in discussions with the producer of the opera production, who in turn had been in contact with a lawyer associated with the opera house. Further, two ethical codes were consulted in the process of obtaining consent: *Good Research Practice* (Swedish Research Council, 2017) and *Guidelines for Ethical Conduct in Participant Observation* (University of Toronto, 2005, henceforth *PO*). In addition, *The EU Guidelines on consent under regulation 2016/679* (European Data Protection Board, 2020, henceforth *EU guidelines on consent*) have also been consulted.²⁸

Before I began to observe the rehearsals, all members of the team were informed of my presence via an e-mail sent out by the producer a week before rehearsals started. During the rehearsal process, after having discussed the consent procedure with the producer (see prior paragraph), I sent out an e-mail to relevant members of the team where it was stated that I intended to film some of the rehearsals beginning the following day (see Appendix I). Three members of the team did not receive this first e-mail, but they did receive it a few days later. Before filming, I individually asked all team members present on location for oral consent of the recording, after reassuring that they had read the e-mail, and after

²⁸ These ethical codes complement each other. Whereas the EU guidelines on consent cover relevant aspects of the General Data Protection Regulation (GDPR) that researchers in Europe are obliged to follow, it is general in its character and not always tailored for research practices. The *PO* on the other hand specifically targets the obtaining of consent in research methods similar to that of the present thesis, whereas it has the drawback of being tailored to Canadian, rather than European, legal frameworks.

a short recapitulation of the purposes of the recordings. Nearly all the team members stated that they had read the e-mail prior to the obtaining of oral consent. One of the team members never received the e-mail due to troubles with the e-mail address, but that person reassured oral consent after the purposes of the recordings had been provided in person. Questions and concerns from the team members were always answered to in relation to the obtaining of oral consent. One of the soloist singers did not want me to publish videos of them singing or acting. This person consented to publication of anonymous non-video material of them singing and acting. The date and time of each individual oral consent was documented in the field notes.

When the choir was present during the days of video recording, I gathered their attention before the rehearsal started, as e-mails had not been sent to them previously. They were provided with information about the purpose of the recordings, and were invited to approach me, the director, or assistant director during the day if they did not wish to participate in the recordings.²⁹ Later that week, an e-mail containing written information about the research project and the recordings was sent out (see Appendix 2.2). The director highlighted the fact that anybody who was uncomfortable with the recordings should approach me or her, but she also stated that the previous days of recording had caused no trouble or significant intrusion for the team. One choir member approached me and wished not to participate in the recordings and was therefore excluded from them.

²⁹ The PO (University of Toronto, 2005, p. 1) states that “participants should remain free to avoid all interaction with the researcher” and the team was informed that they could contact the producer/director or any other member of the crew in case they wanted to opt out of the study (in fact, this was highlighted by the director at one point when the choir was involved in the consent procedure).

In this research, informed consent has thus been obtained by means of “a clear affirmative action” (European Data Protection Board, 2020, p. 6) on the part of each individual participant, with the exception of the choir where explicit individual consent was not obtained.³⁰ The PO (University of Toronto, 2005, p. 5) states that although written consent is preferred in the Tri-Council statement (a general research ethical code of Canada), oral consent may be preferred in many settings where written consent is not “feasible or desirable”. Recital 32 of GDPR states that consent can be collected “through a written or (a recorded) oral statement” (European Data Protection Board, 2020, p. 18). In the data collection procedure of this thesis, oral consent was preferred over written consent to reduce the researcher’s intrusiveness in the professional lives of the participants (see discussion in Speer & Stokoe, 2012). It was deemed that having the many participants, one by one, read and fill in consent forms would interfere with the workflow of the rehearsals. The continuous consent procedure (described below) was judged to assure sufficient options to opt out of the study. The option of documenting the consent in the field notes, rather than by asking the participants to consent while being video recorded, was chosen in accordance with the regulations of the GDPR, stipulating that individuals should consent before any data processing occurs (European Data Protection Board, 2020, p. 19), and that the obtaining of consent should not result in excessive documentation.

After the entire rehearsal process, an additional e-mail was sent out with a recapitulation and specification about how the video-recorded material was going to be used (see Appendix 2.3). This e-mail also included an example of a transcription, with anonymized frame grabs, to

³⁰ The video recordings where the choir is present are not analysed in this thesis.

show the participants what the presentation of the results might look like. All consent e-mails were written in both English and Swedish, as there were non-Swedish speaking members of the team.

It was continuously repeated throughout the rehearsal period (both in the e-mails sent out and in person) that participation was completely voluntary, and that the participants had the right to withdraw their consent at any time by speaking to me, the director, or assistant director in person, by phone, or by e-mail. According to the EU guidelines on consent (European Data Protection Board, 2020, p. 7), consent is not valid if the participants “have no real choice, feel compelled to consent or will endure negative consequences if they do not consent” and “consent should be a reversible decision” (European Data Protection Board, 2020, p. 7). During the fieldwork of this thesis, one individual chose to refuse their consent, and one participant consented with reservations (see above), which validates that the refusal and/or retrieval of consent was an actual option for the participants.

The PO (University of Toronto, 2005, p. 1) highlights that the informed consent procedure in participant observation should be “dynamic and continuous” and in constant dialogue with the participants, as it often involves long-term presence among them. The fact that I was constantly present in the rehearsal space when recording was in progress assured this continuous negotiation of consent. Similarly, it provided the participants with a reminder that they were recorded, so that they could make choices on what information to disclose. On a few occasions members of the ensemble asked me to stop filming, or to delete a certain video-recorded segment. In cases where I hesitated on whether the participants were comfortable with the video recordings, the participants were briefly asked to consent to the continuation of the recordings at that point. As the material is professional in its character, and the

conversations target the creation of performance bodies and not leisurely conversation, the material does not evoke any sensitive details about the participants. Accordingly, no ethical vetting was required before the onset of the data collection.

The EU guidelines on consent (European Data Protection Board, 2020, p. 10) stress that if the data will be processed through different operations and for different purposes, the data subjects should be free to choose to which of the processes they consent to. The participants were informed by e-mails and in person before the recordings that these were going to be used for research and educational purposes, and that there were different options that the participants could consent to. After the recording process, the e-mail that recapitulated and specified the purposes of the video recordings explicitly stated the different purposes and urged the participants to respond if they wished to withdraw consent from any of them. The EU guidelines on consent (European Data Protection Board, 2020), however, stipulate that this information should normally precede consent, although recital 33 allows for “flexibility to the degree of specification and granularity of consent in the context of scientific research” (European Data Protection Board, 2020, p. 30).³¹ In the data collection of this thesis, I was asked by the producer to formulate a short e-mail early in the process, because otherwise she feared that the ensemble would not read it, and thereby complete information was assured in a stepwise, rather than ask-first manner.³² The option of withdrawing consent was stressed in the final e-mail providing the specific

³¹ Recital 33 acknowledges that “it is often not possible to fully identify the purpose of personal data processing for scientific research purposes” and that research subjects should therefore be able to consent to a more general “area of scientific research” (European Data Protection Board, 2020, p. 30).

³² Another important aspect of GDPR (European Data Protection Board, 2020, p. 16) is that “when seeking consent, controllers should ensure that they use clear and plain language in all cases”.

information on the different ways in which the data would be used. Research on informed consent when video recording social interaction has discussed the benefits of providing detailed information on consent at the end of research encounters, rather than in the beginning of them, as ethical concerns may be easier to address in a position where ethical discussion do not hinder the progressivity of the interactions (see Speer & Stokoe, 2012). In this way, the researcher remains sensitive to the “institutional business” of the participants she or he is studying and acknowledges “the possible effects of [hers or his] interactions on participants” (Speer & Stokoe, 2012, p. 70), an important aspect of ethical research practice.

4.3.2 Participant anonymity

Although only one participant explicitly stated that they wanted to have their identity concealed when presenting the data of the study in oral and written presentations of the analyses of this thesis, the identities of all the participants have been consistently concealed, or ‘de-identified’ (see discussion in Profazi & Miecznikowski, 2023) when presenting the data and analysis. This is in accordance with standard procedures within the EMCAIL research field (Broth & Keevallik, 2020). Assuring the anonymity of the research participants has been a golden standard of this work, as far as it has been possible without compromising the analysis. Accordingly, it has never been revealed in what opera house the production took place, which opera piece the production targets, or when the rehearsal procedure and premiere of this opera production happened. Some details around the plot have been revealed, however, and there have not been any efforts to conceal the libretto. These have been necessary to reveal for analytical purposes as the plot is often a topic of conversation,

and the timing of the multimodal details is often done in close coordination with sung libretto. However, all names of the characters of the opera have been replaced with fictional names. Further, in all publications and presentations where this data is involved, images have been de-identified through stylization and all real names have been replaced with fictional ones, to de-identify the participants as much as possible.

5. Summary of the articles

The thesis comprises three articles, which are summarized below. Whereas I am the sole author of Articles II and III, Article I has been written in collaboration with Dr Emily Hofstetter.³³

5.1 Article I – Introversive semiosis in action: Depictions in opera rehearsals

Article I introduces the concept of *depiction* and sets the stage for further research, in Articles II and III, on this iconic interactional resource in the context of creating the “excessive and transgressive” (Atkinson, 2006, p. 187) art form of opera. The article shows how opera is created in the mundane routines of the rehearsals. It provides new insights on depictions in interaction. Before the article, these had mainly been researched in pedagogical settings (Clark & Gerrig, 1990; Keevallik, 2017; Stukenbrock, 2014b). Combining multimodal interaction analysis with semiotic theory, the article shows how depictions are deployed to accomplish proposals, and thereby the ongoing creation and aesthetic negotiation – the *artistic labour*, during the rehearsals. The article adds to research that has studied the sequential and multimodal nature of depictions (cf., for instance, Cantarutti, 2020, 2021; Ehmer & Mandel, 2021; Sidnell, 2006).

In the article, it is argued that the social action is “split” when participants depict, as they both create an imaginary scene *and* propose that that scene should be a part of the production.³⁴ In terms of semiosis, depictions in opera rehearsals are argued to be both self-referential and

³³ I am the first author of Article I.

³⁴ In Articles II and III, this is conceived of as depictions being practices accomplishing the social action of a proposal.

referential of prototypes of mundane behaviour (Lefebvre, 2018); in other words, they operate by both introversive and extroversive semiosis (see Jakobson, 1971; see also section 3.1.1 above). Extroversive semiosis is the prototypical sign relationship that holds between a sign and its referent, in which a sign is not its referent (cf. Clark, 2016 on the pas-une-pipe-principle), but merely a representation of it. Introversive semiosis, in contrast, is a type of sign relationship that does not rely on a factual similarity between the sign and the referent, but on an imputed similarity between the two – a *connotative sign*.

The article draws on Agawu's (1991) interpretation of the notion of *introversive semiosis* in which internal references and structures of art works (music in his case) mutually inform each other and condition subsequent understandings of signs. The referents of the depictions are thus both distal "scenes", such as a father mourning a dying daughter (extroversive semiosis) and the depictions themselves, as the current state of the artwork under creation that relates to the developing aesthetics of the production as it has been manifested in the interaction thus far (introversive semiosis). Prior research has implicitly targeted extroversive referencing in theatre rehearsals: how participant embody scripts by drawing on intuitive resources of how naturalistic interaction runs its course (Hazel, 2015, 2018; Lefebvre, 2018; Norrthon, 2019). In Article I, it is shown how both opera as an artform, and this particular production, conditions how participants build introversive references during rehearsals.

The article argues that participants orient to the distinction between extroversive and introversive semiosis as important for the accomplishment of successful performance bodies. By negotiating extroversive (how one could portray a father mourning a dying daughter) and introversive references (how a father mourning his dying daughter should be

portrayed in a way that is appropriate and coherent with the aesthetics of the production), through depictions, the participants collaboratively create the production. Participants orient to introversive semiosis by referring to prior iterations of ideas for performance bodies and how performance bodies from other productions are explored to develop the introversive semiosis of this production. In orienting to introversive semiosis, the participants manage the coherency of the art piece, that is, its aesthetics, or style.

Finally, the article argues that the concept of introversive semiosis is analogous to how meanings emerge over time in social interaction in general.³⁵ Introversive semiosis is thus a helpful tool in the understanding of how interactional projects develop over the long term.

5.2 Article II – Relocating to depict: Managing the interactional agenda at opera rehearsals

Article II studies how the performers, with relatively low deontic authority *vis-à-vis* the director, introduce depictions on the interactional agenda of opera rehearsals, in the service of making proposals on performance bodies. The performer proposals that are studied occur during the discussion activity of the scenic opera rehearsal, when the performers, director, and assistant director discuss the production (see section 4.2.5 above). The article presents a specific practice for initiating depictions: relocations in space. These relocations break the F-formation of conversation (Kendon, 1990) that the participants are typically in during the discussion activity and increase the distance between the depicting

³⁵ The conception of the relationship between introversive semiosis and interactional histories has been revised since the publication of this article. According to the revised conception, introversive semiosis is a relationship between sign and referent that *builds on* shared common knowledge, interactional histories.

performer and the other co-present participants. This distance achieves a performance-audience-formation that allows for depictions as momentary fragments of the performance activity to be introduced in the discussion activity. With the use of multimodal interaction analysis, the article studies how the performers launch depictions through relocations and how these relocations are responded to by the other participants. The analysis builds on a collection of 80 performer proposals on performance bodies that involve depictions.

It is argued that when relocating to depict, the performers make proximal deontic claims (Stevanovic, 2015) to influence the local agenda of the rehearsal by having the other participants watch and evaluate their depictions as proposals for performance bodies. In turn, it is shown how the director manages these claims through visuospatial resources, that is, responsive relocations, or the absence thereof. The director may choose to endorse the proximal deontic claim by a responsive relocation by which she positions herself as a spectator to the initiated depiction. She may also, however, reject or challenge the proximal deontic claim by remaining in her current position and refraining from assuming a spectator position, or by breaking the spectator position during an ongoing depiction. The director's acceptance or rejection of the performers' proximal deontic claims are relatively independent of the director's attitude towards the performer ideas for performance bodies, i.e. the distal implications of the proposals. A claim to depict a proposal may thus be rejected even though the director displays a positive stance towards the content of the proposal as such.

Relocations may also be used by the director to invite proposal depictions pre-emptively. The director may relocate to a spectator position in response to a description of a performance body, to invite a depiction of it, or even before a proposal has been made. In these cases, the director

provides space on the interactional agenda for the performers without them having to claim it for themselves. Relocations by the director can even be used to index performers' behaviours *as* proposals for performance bodies in the absence of descriptions of proposals.

Relocations thus constitute a way to negotiate proximal deontic rights in a visuospatial manner. Through initiating and responsive relocations, the depicting performer and the director collaboratively create a participation framework that supports the depictive scene, a brief fragment of performance, in the service of a performer's proposal. The article thus contributes to literature on the responses to depictions (cf. Cantarutti, 2020, 2022a, 2022b; Ehmer & Mandel, 2021; Keevallik, 2014b; Stukenbrock, 2012; Thompson & Suzuki, 2014) by demonstrating a third alternative to respond to depictions, beyond i) co-depictions in a parallel distal scene, and ii) co-depictions in the same distal scene, namely iii) assuming spectatorship to depictions. It is argued that the spectator position is an integral part of the depiction, where the spectator is a part of the depiction's base and proximal scenes, but not of the distal scene. Relocations are a practice to microscopically transition into an activity that otherwise necessitates formal transitions, where the director sits at the director's table, the conductor and pianist get involved, etc. Initiating and responsive relocations reflect the dialogic nature of the creative work of this production where the director acts in ways to facilitate an atmosphere where performer proposals are invited and considered.

The article concludes by suggesting that future research should target cross-comparisons of depictions between different datasets to disentangle the relevancy of the scene metaphor for depictions in general, and to investigate differences in terms of leadership and distribution of deontic rights when creating performing arts.

5.3 Article III – On the pursuit of *bodies* to suit the music: From describing to depicting proposed ideas at opera rehearsals

Article III expands on how performers propose performance bodies by focusing on how the design of proposals changes over time, as ideas become developed in decision-making processes. By using multimodal interaction analysis, the article analyses 36 distinct ideas for performance bodies. The ideas go through multiple iterations (typically two to five) over the course of micro-histories (Schmidt & Deppermann, 2023).

The analysis builds on previous research on the different steps of a truly *joint* decision-making process (Stevanovic, 2012), as well as studies on the importance of alignment for collaboration and creation (Stivers, 2008; and Article II of this thesis). Further, the article draws on previous research on changes in turn design over interactional histories (Broth et al., 2017; Deppermann, 2018a) and the development of creative projects over time (Hazel, 2018; Norrthon, 2019; Schmidt & Deppermann, 2023; Yasui, 2013).

The analysis shows that proposals of ideas for performance bodies move from descriptive to increasingly depictive modes over time as the joint decision-making process progresses. Whereas the performers only describe the idea when it is first introduced, its subsequent developments contain simultaneous depictions and descriptions. In the final stages of the decision-making process, the idea gets implemented and coordinated with song. During these late stages of the process, descriptions are scarce, or even entirely absent.

The move from description to depiction is dependent on features of the decision-making process. Before depicting their ideas, the performers do interactional work to secure access to, and alignment and

agreement with, the ideas (cf. Stevanovic, 2012; Stivers, 2008). As Article II has shown, alignment from the recipients is crucial for depictions, and article III shows how the performers secure that alignment in ambiguous cases. Whereas the participants do not depict ideas before there is preliminary access to them, the depictions themselves increase access to the ideas and are often oriented to as necessary for the director to agree with and commit to the ideas presented in them. The recipients of proposals are thus involved in several steps of the decision-making process of an idea.

The fact that depictions become more complex over time in opera rehearsals contrasts with previous accounts of how depictions decrease in size and detail over time, and as they become more conventionalized (Gerwing & Bavelas, 2004; see also Norrthon & Schmidt, 2023; Schmidt & Deppermann, 2023). This contrast is related to the nature of depictions in opera rehearsals. Depictions in opera rehearsals are not merely interactional practices, they are simultaneously the interactional outcome: the performance bodies at their current states, and these naturally reach more developed stages over time. In that sense, depictions in opera rehearsals violate the *pas-une-pipe*-principle of depictions (see section 3.1.1 above) as they *are* what they represent (introversive semiosis).

Article III thus shows the collaborative nature of how performance bodies are created. By constantly affirming co-participation during idea proposals, the performers distribute the agency for the ideas over a collective, even though they are themselves responsible for the progressivity of the proposal. The collaborative nature becomes even more apparent in cases that deviate from the basic pattern, where other participants directly influence performer ideas, notably through depictions. Further, when performer proposals align with previous ideas and arrangements,

they are depicted at earlier stages as access, alignment, and agreement are inferred.

Finally, the article argues that all proposals of ideas during opera rehearsals display understanding of the aesthetics of the production, as it has been manifested in the interaction thus far, not only over specific decision-making micro-histories, but over the macro-interactional history of creating this production. Displays of understanding become particularly evident in the cases where they more directly originate in previous ideas. This understanding of the aesthetics of the production, the introversive semiosis of the depictions, may be subject to challenges by the other participants. Each contribution to the interactional agenda in the form of depictions and descriptions shows sensitivity to a context, that is, an opera production in the making, that the contribution itself renews (cf. Heritage, 1984a).

6. Discussion

This thesis has investigated how participants in scenic opera rehearsals make use of depictions to create performance bodies, that is, what the performers should do on stage to portray the characters of the opera production. The aim was to study the local and multimodal construction of the iconic communicative strategy of *depiction* (Clark, 2016) in naturally occurring video recorded social interaction. The thesis reveals how depictions are a useful tool for participants in opera rehearsals when they create and negotiate the production.

The theoretical background of this thesis resides within multimodal interaction analysis. This methodology has its roots in ethnomethodologically informed conversation analysis and interactional linguistics (EMCAIL). The thesis builds on previous studies on iconic phenomena in interaction and borrows terminology from the psycholinguistic framework on the three distinct communicative methods of depiction, description, and index (Clark, 1996, 2016, 2019). Further, the analysis is informed by concepts from semiotic theory (Agawu, 1991; Jakobson, 1971; Peirce, 1955). The thesis contributes to previous literature on iconic phenomena in social interaction, on studies of multimodal gestalts, joint activities and participation frameworks, joint decision-making, proposals and deontics, and prior studies on performing arts in interaction, as well as theorizations on the relationship between language and other communicative resources.

The thesis set out to study three questions, pertaining to the nature of depictions in scenic opera rehearsals:

- 1) How do participants construct and respond to depictions?

2) What are the interactional and semiotic functions of depictions?

3) What is the relationship between depictions and descriptions, and what does this relationship tell us about the roles of language and the body in human sense-making?

These research questions were explored in the three different articles that comprises this thesis. In the following, I will address each of these questions.

6.1 The multimodal design of depictions in scenic opera rehearsals

Scenic opera rehearsals are, as all face-to-face communication, characterized by multimodality and sequentiality (see sections 2.1 and 2.2 above). Through various multimodal resources, mobilized in quasi-simultaneous manners, participants accomplish social actions that are sequentially adjacent to their responses. As evidenced in this thesis, depictions are constructed using several multimodal resources variably over time, much as multimodal gestalts (Ehmer & Mandel, 2021; Mondada, 2014a, 2016; Stukenbrock, 2018). This thesis contributes to the study of the multimodal design of iconic phenomena that can be categorized under the umbrella term *depiction* (cf., for instance, Cantarutti, 2020, 2021, 2022a, 2022b; Ehmer, 2021; Ehmer & Mandel, 2021; Keevallik, 2010, 2013a, 2013b, 2013a, 2014b, 2015, 2018; Sidnell, 2006; Stukenbrock, 2012, 2014a, 2014b, 2017; Thompson & Suzuki, 2014). These contributions are discussed below in sections 6.1.1 and 6.1.2.

6.1.1 Depictions as multimodal interpersonal gestalts

Article I shows how participants transition into depictions in a stepwise manner, where more and more interactional resources are mobilized towards the distal scene over the milliseconds during which the depiction is produced. In Extract 2 of Article I, for instance, where a baritone depicts his character lifting his dying daughter, the baritone first mobilizes his visuospatial resources towards the depicted scene by placing his hands on her body and beginning to lift her. Only when these visuospatial configurations are in place does he start to sing as his character, thereby also mobilizing the vocal/aural modality towards the distal scene. When both visuospatial and vocal/aural resources are mobilized towards the depictive scene, they reach their climax (see also Stukenbrock, 2018, on *informational climax*). In the climax of the depiction, no resources are oriented to the here-and-now of the interaction (in this case the discussion phase of a scenic opera rehearsal). In other words, the performers are performing. As evidenced in article II of this thesis, these performances are introduced as brief activity transitions within the overall discussion activity of the scenic rehearsal (see section 4.2.5 on the distinction between the *discussion* and *performance* activities of the rehearsals).

Interestingly, depictions that reach such a climax tend to occur later in the decision-making process they are in the service of, as evidenced in article III of this thesis. Depictions thus not only become increasingly multimodal over milliseconds, but also over 10–20 minutes-long joint decision-making micro-histories. In early stages of these micro-histories, depictions are often reduced in form and accompanied by simultaneous descriptions, whereas they reach their most complex form (i.e. the climax), incorporating both visuospatial and vocal/aural

resources, towards the end of them. This replicates earlier findings on interactional histories in instructional settings where the aim is an embodied performance. The embodied performance of a learner becomes more and more elaborate over time, while the descriptions of that performance steadily decrease (Broth et al., 2017; Deppermann, 2018a; Schmidt & Deppermann, 2023). However, it contrasts with earlier findings on depictions and gestures that are shown to become reduced in size, form, and even frequency, over time (cf. Deppermann, 2018; Gerwing & Bavelas, 2004; see also Norrthon & Schmidt, 2023; Schmidt & Deppermann, 2023 on how depictions become routinized over time). This discrepancy can be explained by the specific nature of depictions in opera rehearsals, as both interactional practice for creating performance bodies, and the current states of the performance bodies themselves (see section 6.3.1 below).

Previous research has targeted the collaborative aspects of depictions and shown how other participants can be recruited as participants, or “props” (cf. Clark, 2016) in the same distal scene (Keevallik, 2014b; Stukenbrock, 2012; Thompson & Suzuki, 2014). They can also position themselves as co-depictors in a parallel distal scene (Ehmer & Mandel, 2021; Cantarutti, 2020, 2021, 2022a, 2022b). This thesis continues to show evidence of both these positions vis-à-vis depictions. In addition, however, Article II evidences a third, hitherto undescribed alternative, namely how recipients of depictions explicitly position themselves as spectators to the distal scene. The thesis thus shows empirical evidence of the assumed spectatorship to depictions (cf. reasoning in Clark, 2016, 2019) by demonstrating the specific and situated resources the participants use to do “being spectators” (cf. Sacks, 1984a). It is argued, in line with Clark (2016), that the spectator position is an integral part of the depiction and that the participants who position themselves as

spectators to depictions are a part of the base and proximal scene of the depiction, but not of its distal scene. The evidence for this argument is that recipient alignment through positioning as a spectator is a prerequisite for continuing the depiction. For instance, Extract 3 of Article II illustrates how performers abandon initiated depictions when the director does not relocate as a spectator, or even moves closer to them to emphasize the F-formation. Depictions that successfully accomplish joint evaluations and considerations of performance bodies thus require both intrapersonal and interpersonal coordination (cf. Deppermann, 2014). Depictions in opera rehearsals are thus conceived of as both multimodal and *interpersonal* gestalts (cf. Clark, 2016 on *joint depictions*; see also Keevallik, 2014a; Stukenbrock, 2012). Although this thesis has mainly focused on the coordination between a proposing/depicting performer and the director, future studies could continue to explore the interpersonal coordination that facilitates depictions, notably the many conducts with which the co-performers position themselves in ways that align with an initiated depiction (cf. Article II of this thesis). This would shed further light on the collaborative nature of the scenic opera rehearsals studied in this thesis (see also section 6.1.2 below), and possibly also of other settings in which performing arts are created.

6.1.2 Depictions as locus of negotiations of deontic rights

As Article II of this thesis shows, depictions require changes in the participation framework from conversation to performance. In other terminology, this is put as the differences between Kendon's (1990) famous *F-formation*, and a *performance-audience-formation* that, at least

hitherto, seems to be specific to scenic opera rehearsals.³⁶ The participants prepare for depictions in visuospatial ways by recalibrating the interactional space, much as what has been observed for activity transitions (cf. Broth & Mondada, 2013, 2019; Mondada, 2005, 2009). By initiating spatial relocations, performers project depictions and thus implicitly recruit (cf. Kendrick & Drew, 2016) the other co-present participants to join them in a change of interactional formation, participation framework, and activity, from discussion to performance. The performers thereby make proximal deontic claims when initiating depictions (cf. Stevanovic, 2015, on *proximal deontic claims*). By relocating to depict, they claim to have the right to influence the *local interactional agenda* (cf. Stevanovic & Peräkylä, 2014), namely by introducing a depiction, and thereby a momentary collaborative performance, in the discussion. The other participants (mainly director) may align, or not, with these changes (cf. Stivers, 2008, on *alignment*). In other terms, they accept or reject the performers' proximal deontic claims (cf. Stevanovic, 2015). The multimodal and interpersonal nature of depictions thus reflects the deontic order of the opera rehearsals, i.e. the implicit order of who has the right to develop ideas for performance bodies, and when they have the right to do so.

As evidenced in Article II of this thesis, it is the director who manages the interactional agenda during the rehearsals. The performers orient to her as a chairperson of the rehearsals, *and* as an expert who has the final word on the performance bodies (see Asmuß & Svennevig, 2009; Clifton et al., 2018; Pomerantz & Denvir, 2007, on the role of chairpersons in meetings; Van De Mierop et al., 2020; Svennevig, 2011;

³⁶ In future research it would be of high interest to investigate whether the performance-audience-formation can also be seen in relation to depictions in settings where participants do not create performing arts, such as everyday conversations in families or between peers.

Tranekjær et al., 2022, on interactional manifestations of leadership; and Stevanovic, 2021 on expertise and authority). The director thus ostensibly has deontic authority over both proximal and distal matters (cf. Stevanovic, 2015; Stevanovic & Peräkylä, 2012). She grants or rejects the performers' proximal deontic claims, irrespective of her attitudes towards the distal implications of the proposal. She may for instance reject a claim to continue to show a depiction despite showing strong affiliation with the idea that the performer depicts (see Extract 4 of Article II). Article II also shows how the director invites performer proposals on the interactional agenda (or “hands over the stage”, cf. Reed & Szczepek Reed, 2013, p. 335), by performing an initiating relocation, and even prefers to make space for performers to propose, rather than instructing them (see Extract 5 of Article II). In that way, the director facilitates an environment where performer proposals are invited and considered.

The setting that this thesis studies, i.e. this particular production of an opera with these particular participants, can thus be conceived of as both hierarchical and collaborative. Although the director may reject performer claims to depict, and although mild attempts against the director's “power” can be seen (see for instance Extract 4 of Article II), the rehearsals are not primarily characterized by power struggles. On the contrary, the participants of the scenic opera rehearsals respectfully negotiate whose contributions will best serve the overall purpose of the interaction – their shared goal of creating a successful production.

As evidenced above, the negotiations of deontic rights are carried out with initiating and responsive relocations. This means that they are negotiated in micro-sequential preparations for the macro-sequence adjacency pair of *proposal* + *uptake* (see, for instance, Mondada, 2021, 2022; Deppermann & Schmidt, 2021a; Deppermann & Streeck, 2018, on *micro-sequentiality*). Much as Stukenbrock's (2018) multimodal

summons-answers that precede other social actions, initiating and responsive relocations in opera rehearsals thus serve to pave the way for an adjacency pair. Previous research on deontic rights in interaction has identified the adjacency pair as the main locus of negotiations of deontic rights (Stevanovic, 2018). The adjacency pair is normally associated with turns-at-talk, and the existence of multimodal adjacency pairs is still debated (Deppermann & Streeck, 2018, Stukenbrock, 2018; see also section 2.2.1 of this thesis). However, some research has begun to tease apart how visuospatial behaviour might figure in deontic negotiations (cf. Asmuß & Oshima, 2012; Barske, 2009; Clifton et al., 2018; Kuroshima, 2023; Stevanovic & Monzoni, 2016; Tranekjær et al., 2022; Tuncer, 2015; Van De Mierop et al., 2020; see also Reed & Szczepek Reed, 2013, for an example of embodied negotiations of the interactional agenda in musical masterclasses). This thesis contributes to this emerging understanding by showing evidence of negotiations of deontic rights below the level of the adjacency pairs normally associated with language. Although the initiating and responsive relocations that initiate depictions as part of proposals follow one another in a sequential fashion, rather than being “continuous adjustment” (Mondada, 2022, p. 42), they are not turns-at-talk. In Schegloff’s (2007) terms, deontics may be negotiated in *sequential* organization as well, rather than in *sequence* organization exclusively.

The multimodal and interpersonal design of depiction thus allows for subtle negotiations of deontic rights, in the visuospatial and micro-sequential calibrations of the interactional space that prepare for the macro-sequential action of a proposal. The recipients of depictions, and, in extension, of proposals (see discussion in section 6.2.2 below), thus “inhabit” those depictions and proposals (Goodwin, 2018) by collaboratively achieving them. The interaction in opera rehearsals is joint, on a

fundamental level, a topic that will further be explored in section 6.2 below.

6.2 Interactional functions of depictions in scenic opera rehearsals

This section discusses the different interactional functions of depictions in opera rehearsals – the social actions that the practice of depictions are demonstrated to support in this setting. It reviews the role of depictions in creating the opera production (section 6.2.1), more specifically how depictions accomplish proposals and support joint decision-making (section 6.2.2) and how they reflect the interactional history of the process of creating an opera production (section 6.2.3).

6.2.1 Depictions as tools in the creation of performing arts

Previous research on depictions, has shown them to fulfil interactional functions such as illustrating stories (Cantarutti, 2020, 2021, 2022b; Ehmer & Mandel, 2021; Sidnell, 2006; Thompson & Suzuki, 2014) or making embodied skills to be learned perceptually available (Ehmer, 2021; Keevallik, 2010, 2013a, 2013b, 2013a, 2014b, 2015, 2018; Stukenbrock, 2012, 2014a, 2014b, 2017; Szczepek Reed, 2021; Szczepek Reed et al., 2013; see also section 3.2.4 above). The function of depictions in scenic opera rehearsals are particularly explored in Articles I and III of this thesis, and resemble that of the function of depictions in theatre rehearsals (cf. Norrthon & Schmidt, 2023; Schmidt & Deppermann, 2022). Depictions are shown to be a tool in the collaborative process of creating performance bodies. Depictions make performance bodies available for the other participants to directly perceive and simultaneously propose that they should be a part of the production. In other

words, they allow the participants to evaluate the aesthetic quality of the performance bodies. Article III shows how depictions are crucial for the decision-making on performance bodies, evidenced by orientations of the director, and other participants. The director explicitly asks the baritone, repeatedly in Extracts 1a and 1b in Article III, to show her what his idea for a performance body looks like, before she makes a final evaluation of the extroversive and introversive semiosis of the performance body that is proposed. Much as in previous research on joint decision-making where the object that is decided upon is jointly manipulated in the here-and-now (see for instance Lindholm et al., 2020; Magnusson, 2021, in the case of written artifacts), depictions are the performance bodies at their current states, the interactional practice by which these become tangible (see discussion in section 6.3.1 below).

The performance bodies that the depictions show are not to be conceived of as pre-existing to the interaction, i.e. to the moment that they are expressed with a depiction. On the contrary, they come into being in interaction, and the participants use depictions as exploratory devices, much as has been evidenced for markings in dance (Kirsh, 2011). Some depictions in the data are not even produced to be watched by their spectators, which provides evidence for their exploratory character. In the future, these depictions could be interesting for a topic of inquiry as it may shed light on differences in design between “private” and “public” behaviour.

6.2.2 Depictions in the service of joint decision-making

In this thesis, the focus is specifically on how depictions accomplish *proposals* on performance bodies, and more specifically on how the opera

performers make these proposals.³⁷ As Article III of this thesis explains (see also section 2.4.1 on proposals), proposals differ from other directive social actions (see Couper-Kuhlen, 2014; Stevanovic & Svennevig, 2015) in that they initiate processes of joint decision-making and thus involve weaker deontic claims than for instance orders and instructions.

Previous research on theatre rehearsals has proposed that actor implementations (through depictions) of the director's instructions constitute proposals of performance bodies (see Norrthon & Schmidt, 2023; Schmidt & Deppermann, 2022, 2023; see also Stevanovic & Monzoni, 2016; and Stivers & Sidnell, 2016 on embodied proposals). Depictions that directly follow a director's instruction can indeed be considered a kind of proposal, but these differ from the proposals that the performers make in the scenic opera rehearsals studied in this thesis. The proposals studied in this thesis are designed as verbal turns-at-talk that get a verbal response. As Article III shows, the proposal turns are sometimes constructed using descriptions only, and sometimes using syntactic-bodily units (cf. Keevallik, 2014a, 2018) comprising both descriptions and depictions.

Constructing a proposal as a turn-at-talk has important implications for the joint decision-making process that is characteristic of the social interaction of the scenic rehearsals of this opera production. It imposes conditional relevance on the director to (more or less) immediately provide an uptake of the proposals (cf. discussion on adjacency pairs and multimodal behaviour in sections 2.2.2 and 6.1.2 above). In constituting the first part of an adjacency pair, proposals formatted as turns-at-talk embody stronger proximal deontic claims than do implementations of

³⁷ Although this was beyond the scope of the present thesis, depictions may occasionally be used by the director to instruct, and by the performers to implement those instructions (cf. Norrthon & Schmidt, 2023; Schmidt & Deppermann, 2022, 2023).

proposals. By making a proposal in a turn-at-talk, the performers make sure that the director explicitly agrees with their ideas for performance bodies, or contrarily, has been given an “official” chance to veto to it. Thereby, although the performer proposals are agentive in that they involve stronger proximal deontic claims than merely implementing an instruction, the authorship, and accountability, for the idea becomes distributed over a collective. If the theatre rehearsals observed in Schmidt and Deppermann (2022, 2023) and Norrthon & Schmidt (2023) are characterized by a top-down procedure in which the director instructs and the performers implement, the scenic opera rehearsals studied in this thesis are in turn characterized by a bottom-up procedure in which the performers propose ideas that, via the approval of the director, become implemented through a joint procedure. For a continued discussion on the affordances of including depictions in verbal turns, see section 6.4.1 below.

As Article III of this thesis shows, it is not only the decision-making process that is joint, but also the mere development of ideas (see Extracts 2a-b of Article III). Performer ideas may be modified by the director through depictions or modified by the performers themselves through depictions in response to evaluative descriptions from the director. The fact that ideas become jointly modified in this manner once again reveals the dialogical nature of the joint decision-making process of these scenic opera rehearsals (cf. Linell, 2009; see also section 6.1.2 above). Future studies could explicitly target differences in work methods between different artistic ensembles, with the aim of identifying specific affordances of different director strategies (for instance top-down vs. bottom-up, see prior paragraph).

6.2.3 Depictions as manifestations of interactional histories

In making the current state of parts of the production available in interaction, depictions also inevitably show an understanding of the opera production that the participants are currently creating, as it has been manifested in the interaction thus far through both descriptions and depictions. This also means that depictions demonstrate understanding of the current state of the production (and its aesthetics), while simultaneously proposing to add to it. Depictions thus support two actions (cf. Rossi, 2018) – demonstrations of understanding, as a second pair part or as a response to a larger structural organization (cf. Robinson, 2013), and proposals, as a first pair part of a new adjacency pair. In that sense, they are both responsive and initiating (see below on their nature as context-sensitive and context-renewing).

The participants are creating a tragedy, that is, a “place of the tension-charged relationship between tragic experience and playful implementations” (Menke, 2004, p. 207). The style with which human behaviours should be playfully implemented, the introversive semiosis of the performance bodies (see section 6.3.1 below), comes into being in social interaction. When creating performance bodies, the participants draw on their understanding of the production thus far, the “production-specific interactional history” (Norrthon & Schmidt, 2023, p. 338). During the rehearsals, the participants both describe and depict performance bodies “into being” (Heritage, 1984a, p. 290). In aesthetic discussions, that is, discussion on *how* to portray behaviours in this production, the performance bodies are negotiated, judged and valued against performance bodies of other productions, ideas of what traditional opera looks like, practical constraints such as scenography and the singing body, and against the participants intuitive understanding of human behaviours

and the psychology behind these, as well as the character of the music of the opera. Creating a production is an aesthetically accountable practice (cf. Lefebvre, 2018).

The discussions and explorations on how to perform become a continuously developing shared interactional history of how this ensemble wishes to accomplish this production, and this accumulated knowledge in turn conditions how performance bodies are achieved interactionally, namely, the design of how they are expressed. Article III of this thesis specifically covers the interactional history of how the production is developed. It shows how each depiction builds on previous descriptions and depictions (one's own or others'). Depictions in opera rehearsals are thus both context-sensitive and context-renewing (cf. Sacks et al., 1974; Heritage, 1984a). They are sensitive to both immediate sequential information (see Extract 2a of Article III) *and* to the interactional history of the production (cf. Deppermann, 2018a). Over the joint decision-making micro-histories of 10–20 minutes described in Article III (see also Schmidt & Deppermann, 2023, on *interactional micro-histories*), this is evidenced by the fact that participants depict their ideas earlier if they are in line with previous ideas and arrangements (see Extract 2a). In sum, the participants draw on interactional histories when creating depictions, and it is these that give the depictions their introversive referents. Ultimately, it is argued that introversive semiosis as a concept is analogous to how participants' shared lived experiences in general are influenced by, and influence, the local interaction, and how interactional practices develop over time, both on micro and macro scales. Introversive semiosis is conceived of as a particular sign relationship that builds on interactional histories.

6.3 Depictions as semiotic resources

According to Clark (2016, p. 327), depictions are “physical analogs” of what they represent – they are signs. One of the aims of this thesis was to investigate the semiotic function of depictions in opera rehearsals. In this section, I discuss what depictions represent and reveal their double nature as interactional practice *and* interactional outcome (6.3.1), and ultimately problematize the divide between depictions and their referents by comparing the representative nature of depictions in scenic opera rehearsals with that of depictions in other settings (6.3.2).

6.3.1 Depictions as performances

Article I of this thesis examines how depictions function as semiotic resources, or communicative signs. It is argued that depictions in scenic opera rehearsals reference both prototypes of mundane behaviour, in extroversive semiosis, and themselves as the current state of the artwork that relates to the aesthetic style of the developing production, in introversive semiosis (see Jakobson, 1971 and section 3.1.1 of this thesis for the distinction between extroversive and introversive semiosis). In this thesis, depictions are thus conceived of as fragments of the production, i.e. artistic representations of distal scenes – performances used for interactional purposes. They are both interactional practice, and outcome (see Norrthon & Schmidt, 2023, for a similar discussion). In that regard, depictions and performance bodies are two distinct terms with one and the same referent. According to Clark (2016) depictions follow the *pas-une-pipe-principle* – they are not what they depict (see section 3.1.1 above). In scenic opera rehearsals, it is true that depictions are not their distal scenes. A baritone hugging a soprano, for instance, is not a father mourning a daughter but a baritone merely performing a playful

implementation of that distal scene. By being manifestations of the production at its current state, however, depictions in opera rehearsals partially violate the *pas-une-pipe-principle* – they are the performances that they aim to create.

The choice to use the distinct terminology of *depiction* (interactional practice) and *performance body* (interactional outcome) is made to precisely reflect the double nature of depictions as communicative signs and performance.³⁸ Further, the choice of the term *performance bodies* for the interactional outcome of the scenic opera rehearsals also reflects the difference between member’s and analyst’s terminologies. Whereas researchers speak about *depictions*, following Clark’s (2016) seminal work,³⁹ this terminology is not used by the participants of the rehearsals to speak about the interactional aims.⁴⁰ In the scenic opera rehearsals in focus of this study, the director defines their interactional aim as a quest for “bodies to suit the music”. The term *performance body* is thus derived from a member’s term (although its metaphorical character is highlighted by my addition of *performance*). The use of the more technical term *depiction*, on the other hand, facilitates the comparison between depictions in opera rehearsals and depictions in other forms of social interaction (see section 6.3.2 below).

Using EMCAIL methodology, performances and performing arts can be defined by looking at participant orientations to behaviour. In interactions in the theatre (cf. Broth, 2011), these orientations are for instance observed through the separation of participants in audience

³⁸ Indeed, it would be confusing to say that the participants deploy depictions to create depictions, even though that is essentially what happens during the rehearsals.

³⁹ Although in EMCAIL work different terminology is used to refer to phenomena that is similar to, or identical with, depictions (see section 3.2.1 above).

⁴⁰ Nor do they use it to speak about the interactional resource, but these are rarely topicalized by participants in social interaction in general (cf. article I of this thesis).

members and actors on stage. This separation of participants makes it easy for the participants to understand that what is happening on stage is a performance by the actors. It is an expectancy from the culturally shared norm on the institution of theatre. In rehearsals however, and perhaps even more so for mundane interaction, there is no stage that separates the actors and the audience. How can we then be sure that the performers are performing when they depict? Participant orientations allow us to answer that question. Depictions become recognizable as performing arts in interaction through different interactional resources. This thesis, as well as previous studies (cf., for instance, Cantarutti, 2020; Couper-Kuhlen, 1999; Sidnell, 2006), show how changes in different vocal/aural and visuospatial parameters create ruptures with prior interactional behaviour so that depictions become recognizable as *something else*, namely, as invoking a scene that is displaced from the here-and-now. These ruptures (or “cesural areas”, Ehmer & Mandel, 2021) may for instance be achieved through postural changes, such as the limping style of walking suddenly adopted by the baritone in Extract 1 of Article II. As the limping style of walking is different from the baritone’s way of walking, it displays his origo displacement (cf. Bühler, 1934/2011; see also section 3.2.1 above). The ruptures may also be achieved through changes in pitch height or voice quality (cf. Couper-Kuhlen, 1999; Günthner, 1999). The ruptures serve to highlight the non-seriousness of depictions (cf. Clark & Gerrig, 1990), and their *pas-une-pipe-reality* (cf. Clark, 2016). They index the multimodal gestalts that they are a part of as not being genuine multimodal gestalts pertaining to the here-and-now, but as artistic correlates of such multimodal gestalts – performances rather than genuine interaction.

This understanding of depictions as performances is also evidenced in recipients’ orientations to depictions. In scenic opera

rehearsals, the recipients of depictions orient to them as scenes from a there-and-then that are to be evaluated for their suitability for the production that is being created. In other words, they orient to the *origo* (cf. Bühler, 1934/2011) of the depicter as displaced to a fictional character in a fictional time and place, rather than pertaining to themselves in the here-and-now. In the rehearsals, recipients even actively position themselves as an audience to the depicted scene (cf. analysis of Article II). The initiating and responsive relocations thereby accomplish a mimetic separation of participants into audience and actors “on stage”, in the absence of an actual theatre building with an actual stage (cf. Broth, 2011; see also section 4.2.2 on the rehearsal space in which the rehearsals take place).

As evidenced by previous studies (cf. Cantarutti, 2020), participants rarely mislocate *origo*, and thereby mis-ascribe social action (cf. Deppermann & Haugh, 2022; Levinson, 2013, for discussions *on action ascription*), in their interpretations of depictions. They do not treat depictive speech as “serious”, pertaining to the here-and-now (cf. discussions in Clark & Gerrig, 1990, on the non-seriousness of depictions). Indeed, such orientations would imply different behaviours than the reactions to relocate as a spectator (see reasonings in Article II of this thesis). In future studies it would be of value to investigate cases of “unsuccessful” depictions in which the *origo* is misinterpreted and the social action mis-ascribed. For such studies however, finding a thorough collection might be problematic and may require examinations across several corpora, due to the relative infrequency of unsuccessful depictions. In the present corpus for instance, although depictions are omnipresent, a depiction is only misunderstood once. Interestingly, this speaks of their omnirelevance for social interaction and communication, and supports Clark’s (2016) argument that they should be subject to profound study,

and that they should be considered in theories of language processing. The final section (6.4) of this thesis expands on the relationship between depictions and language. Before that, however, this section ends with a discussion on the generalizability of the concept of *depiction* to settings beyond the scenic opera rehearsal (sub-section 6.3.2 below).

6.3.2 Depictions beyond opera rehearsals

In this thesis, I sympathize with Clark's (2016) aim to provide a holistic account of depictions as a clearly delimited and generalizable communicative phenomenon, and I wish to spark a conversation on the different ways that depictions in different interactional settings are relatable. I have proven how depictions are a relevant interactional practice in the social interaction of scenic opera rehearsals, and I have shown the many ways in which they resemble iconic phenomena discussed in previous interactional literature, both in terms of their form, and their function (see particularly section 3.2 on depictions in interaction). Some questions remain however, such as whether all depictions in social interaction operate on both extroversive and introversive semiosis. This issue, and related problems, will be discussed in this section.

An issue that remains to be tackled in terms of the generalizability of the concept of depiction is their heterogeneity across different settings in terms of both form and function. Whereas some depictions involve the entire body, and origo displacements, others, like iconic gestures, are merely *prop depictions* (Clark, 2016) in which a depicter uses parts of their body to manipulate a distal scene that they are not themselves "in". Further, whereas some depictions, such as the depictions in scenic opera rehearsals, and reenactments that retell prior events, clearly represent distal scenes, it is more of a challenge to imagine what the distal scenes

of for instance demonstrations in pedagogical settings may be. In some of these settings, it is relatively easy to visualize distal scenes, such as in Stukenbrock's (2012, 2014a, 2014b, 2017) renditions of self-defence training where the teacher indeed invokes "phantasms" in which a hypothetical victim is attacked by a hypothetical aggressor.⁴¹ In Keevallik's (2010, 2013a, 2013b, 2014a, 2014b, 2015, 2017, 2021) studies on dance rehearsals, on the other hand, it is less straightforward what the corresponding distal scenes, to the demonstrations of dance moves that the teachers frequently use for instructional purposes, would be. If these demonstrations are contrasted with descriptions, i.e. showing a dance move vs. describing how to do that same move ('bring your left foot forward...' etc.), it is evident that they are different interactional practices. Interestingly, this suggests that depictions are most easily recognized as a unified phenomenon when contrasted to descriptions (see section 6.4.1 for a discussion on the notion of *description*).

Clark's (2016) definition of depictions, however, states that they represent distal scenes. But again, what kind of scenes are we imagining in the dance practice scenario? A distal scene in which a dancer is performing a certain dance move? Or are they rather simply that dancer performing that dance move in the here-and-now, momentarily suspending syntax (cf. Hofstetter et al., 2021)? To put it in the terminology of this thesis, whereas they operate on introversive semiosis in so much as they are referring to themselves and to the culturally recognized form of a certain dance practice, do they have extroversive referents? Are they signs, or "physical analogs" of the things they represent (Clark, 2016, p. 327), or are they simply the "things" themselves (cf. section 3.3.1 above on the

⁴¹ This feature of the phenomena in Stukenbrock's data is reflected in the fact that she opts for the term *enactment* rather than *demonstration*, which more clearly connotates the existence of a scene to be enacted (cf. Stukenbrock, 2017).

pas-une-pipe-principle)? Future research on depictions should wrestle specifically such issues. Much as Szczepk Reed (2021) forewarned, it might be relevant to distinguish between different forms of iconic interactional phenomena (or alternatively to create sub-categories to the umbrella concept *depiction*, as in Schmidt & Deppermann, 2022). In sum, more research that is sensitive to the fine-grained temporal details of interaction should be conducted on iconic phenomena in social interaction to discern what unifies and separates them.

Finally, this thesis has discussed introversive semiosis as a relevant feature of depictions in scenic opera rehearsals. But to what degree is introversive semiosis a feature of depictions in other contexts, and of social interaction in general? In Article I of this thesis, it was argued that introversive semiosis is “analogous” to how interaction is structured more generally. This relates to studies on interactional histories and longitudinal CA (cf. Deppermann, 2018a; Deppermann & Pekarek Doehler, 2021; Schmidt & Deppermann, 2023). Over interactional histories, participants interactional patterns change, and they may make these histories relevant in several ways in the interaction. In this thesis, it is argued that introversive semiosis conditions interactional patterns, and that participants make it relevant in interaction when they refer to previous interactions on the same topic (see Extract 4 of Article I). Whereas *interactional histories* is a term that reflects a feature of interaction that may condition for instance turn design, *introversive semiosis* is a sign relationship that builds on interactional histories (shared common ground). In this thesis, it is specifically related to artistic signs – *performance bodies*, and communicative signs that are used to create these – *depictions*. However, it is assumed that all communicative signs, or interactional practices, work by means of introversive semiosis, although the context to which they refer may be less specific than a particular opera

production. In that sense, because of the interactional histories, and common ground (cf. Clark, 1996), that participants share, introversive semiosis is omnipresent in social interaction.

6.4 Depictions and descriptions

Although this thesis has focused on depictions, descriptions are ubiquitous in the analyses and one of the aims of this thesis is to examine the relative distribution and roles of depictions and descriptions when creating performance bodies. This section discusses the notion of description against the backdrop of the analytical findings on how these are used in scenic opera rehearsals. It ends by suggesting implications from this thesis for the conception of the relationship between language and the body in human sense-making.

6.4.1 What is a description?

In Article III of the thesis, it was shown that descriptions are used early in joint decision-making processes to secure preliminary understanding of ideas before depictions, a more ambiguous practice, can be used (see Reber, 2012, and Stevanovic & Frick, 2014, on the ambiguity of certain iconic resources). A performer who wishes to propose a performance body through a depiction does interactional work to increase the likelihood that the proposal/depiction will receive 1) an uptake that is 2) positive. This interactional work is partially achieved through language. This section tries to disentangle whether that means that it is partially achieved through *descriptions*, as they have been described by Clark (1996, 2016).

To produce a successful proposal that incorporates a depiction, performers draw on three resources that are all related to language:

lexical semantics, grammar, and features of sequence organization (cf. Sacks et al., 1974; Schegloff, 2007). By using pre-emptive descriptions (see Article III of this thesis), the participants draw on lexical semantics, to make meaning and to secure intersubjectivity on ideas for performance bodies. But the role of language in securing a positive outcome of a proposal goes beyond lexical semantics. In Article II of this thesis, it was shown that visuospatial resources are used to project depictions. However, in some cases, and notably if the participants are already in a performance-audience-formation and no projecting relocation is necessary, syntax may be used to project depictions. The participants thus draw on the projective capacities of grammar (cf. Auer, 2005, 2009) to secure their recipients' attention on depictions and to mark them as relevant on the interactional agenda.

Further, grammar/syntax is used to organize the depictions in turns-at-talk (rather than *independently* of language, cf. Clark, 2016). In Article III, Extract 1c, for instance, the baritone begins an utterance by saying “here I could like-” before disrupting syntax and producing a depiction, in what has been referred to as a *syntactic-bodily unit* (Keevalik, 2013a; 2014a; 2015; 2018; see also Clark, 2016, on *embedded depiction*, and Enfield, 2009, on *composite utterances*). The benefits of including a depiction in a turn-at-talk are (at least) twofold. As evidenced above, grammar has projective capacities. Further, by virtue of being produced in the modality reserved for turns-at-talk in spoken language – it allows for self-selection as the next speaker, in other words, for turn-taking. The turn-taking system (cf. Sacks et al., 1974) is a robust order of social interaction. It serves to secure that a speaker has the attention of their co-present interactional parties. For a depiction to be successful, that is, for it to accomplish a proposed performance body that becomes evaluated by other participants, it must receive attention from the other

participants – the depiction must have spectators (see also Clark, 2016, 2019). Including it in a turn secures such attention. It is unsure, and to be discovered in future studies, whether independent depictions (Clark, 2016), that is, depictions that occur outside of a turn-at-talk, can achieve this.

The role of sequence organization in achieving a successful proposal depiction goes beyond that of self-selection in turns, however. Whereas grammar projects more to come within a turn, the turn itself, upon completion, projects a conditionally relevant response (Auer, 2005, 2009). Visuospatial behaviour alone, on the other hand, does not necessarily achieve such interactional contingencies (cf. Schegloff, 1984; see also section 2.2.1 above). Using turns-at-talk thus increase the likelihood that the performers receive a response from the recipients of their proposals. This allows the performers to check whether there is access to, and alignment and agreement with, the idea for a performance body.⁴²

In sum, it is not only the semantic quality of language that is important for the successful deployment and understanding of depictions, but also the organizational properties of language. Whereas lexical semantics disambiguate depictions (as has been pointed out before by for instance Weeks, 1996; Stevanovic & Frick, 2014), syntax allow recipients to pay attention to depictions as constituents of proposals^{43,44}, and the

⁴² Visuospatial means may, however, serve to further emphasize this conditional relevance in response mobilizing practices (cf. Stivers & Rossano, 2010). In Extract 2a of Article III, for instance, a response from the director is pursued by the baritone's gaze at her by the end of his depiction.

⁴³ This is achieved through resources such as softeners (Stevanovic, 2021), deontic mitigations (cf. for instance Stevanovic, 2013b), or other grammatical formats that have been associated with proposals (cf. Couper-Kuhlen, 2014; Thompson et al., 2021).

⁴⁴ In article II of the thesis, it is shown how visuospatial arrangements make behaviour recognizable as a proposal – at least in retrospect. However, whether visuospatial arrangements alone can *project* something as a proposal remains uninvestigated.

normative constraints of sequence organization nudge them to respond to them. However, this raises an important concern for the question on the relative roles of depictions and descriptions in scenic opera rehearsals – what in the above-described features are a part of the concept of *description*?

In Clark's rendition (2016), descriptions and depictions appear as two contrasting methods of signalling with which semantic content can be expressed. However, many utterances that are analysed in this thesis, and in social interaction in general, have little, or no, semantic content and rely on composite signals in which indexical, symbolical, and iconic means of signalling are combined to make meaning (see also Clark, 1996). In an utterance such as "on my knees is the best way or lie down next to you and sing", it is easy to envisage something that is described, and the description becomes identifiable as an interactional practice (contrasted with a depictive interactional practice in which the baritone sits on his knees or lies down next to his co-performer). However, whereas no one would argue that utterance such as "here I could like-" (see Extract 1c of Article III of this thesis) is not language, it does not really describe anything, and many of its constituents rely on indexes ("here", "I") and pragmatic devices for indexing possibility "could", and a focal part of an utterance "like"). In that sense, it does not become recognizable as the interactional practice of a description, but rather of grammar, and sequence organization (see above). *Description*, then, as an interactional practice, becomes most easily understood in comparison to *depiction*, much as depictions are most easily understood as a unified interactional practice in relation to descriptions (see section 6.3.2 above). Although language is often provided as the *a priori* symbolic sign, or communicative method (cf. Peirce, 1955; Clark, 2016), *description* does not cover all linguistic interactional practices. On the other

hand, it is a useful term to denote an interactional practice in which a projected element of a clause is realized with lexical semantics. This interactional practice can in turn successfully be contrasted with the practice of realizing a projected element of a clause with a depiction, in a syntactic-bodily unit (cf. Keevallik, 2018).

6.4.2 Holistic sense-making in opera rehearsals and beyond

In the previous section, it was concluded that depictions and descriptions are two ways of terminating a linguistically projected clause. This, however, is not the entire picture of how depictions and descriptions figure in interaction. As evidenced by this and previous work (see Article III; Clark, 2016), depictions and descriptions are frequently deployed simultaneously and in concert with other interactional resources, both simultaneously and sequentially. Further, depictions and descriptions should not be considered translations of one another. EMCAIL research has shown how depictions and descriptions may indeed disambiguate each other's meanings (see for instance Stevanovic & Frick, 2014; Urbanik & Svennevig, 2021). However, prior research on multimodal conduct in general (see for instance Goodwin, 2000), or on depictions/demonstrations in particular (cf. Harjunpää et al., 2023; Keevallik, 2010, 2013a, 2013b; Schmidt & Deppermann, 2022), show how they contribute to sense-making in supplementary, rather than redundant, ways.⁴⁵ Depictions have indeed been characterized as vague and in need of *interpretative frameworks* (cf. Clark, 2016; see also Harjunpää et al., 2023; Reber, 2012; Stevanovic & Frick, 2014), but they may sometimes be more precise than lexical descriptions (cf. Harjunpää et al., 2023; Keevallik,

⁴⁵ This is also acknowledged by Clark (1996) in a discussion on the relationship between demonstrations and lexical affiliates.

2013a). Further, according to an ethnomethodological framework, language is entirely dependent on interpretative frameworks itself, namely, the context to which it is indexically tied (cf. Garfinkel, 1967; Heritage, 1984a). Depictions may indeed be an important part of such contexts. In sum, the complex picture of how descriptions and depictions are deployed in social interaction emerges when studying them from a temporal and micro-analytic perspective, with an unmotivated interest in the resources that participants use to construct them. Especially IL research has emphasized the danger of working with predefined categories, such as linguistic units, when trying to disentangle which aspects of social actions that matter most to the participants of it, and how they do so (cf. Ford, 2004; Ford, Fox & Thompson, 1996). In this thesis, although I have been working with predefined dichotomous categories, the results continuously reveal the holistic and integrated nature of social interaction.

To avoid the unmotivated distinction between language and the body (cf. Keevallik, 2018; and Article II of this thesis), I have opted for the terminology of visuospatial and vocal/aural resources.⁴⁶ I argue that these reflect perceptive qualities rather than an artificial distinction between different kinds of motor patterns – movements of the limbs versus movements of abdominal and/or intercostal muscles, the larynx, and the tongue. In my view, language is primarily a fine-grained and precise motor ability. However, neither language/body nor visuospatial/vocal/aural capture the fact that we produce and perceive interactional resources holistically (cf. Deppermann, 2013). As evoked in section 2.2.1 we see talk, and we hear and feel movements (Mondada, personal communication, June 26, 2023). When we interact, we make sense of the entire

⁴⁶ Alternatively, in my opinion, Keevallik's (2018) terminology of bodily-visual and bodily-vocal are beneficial in that they emphasize the embodied nature of both visual and vocal communicative behaviour.

situation, not discrete communicative signs. In that way, it becomes impossible to disentangle what is peripheral, and what is central, and this is what Proust (1919/2003) hints at in the quote that prefaces this thesis. We remember not only what is said, the semantic content, but how it was said, and the emotional resonance the power of relationships creates in us. It is often argued that language can be abstracted from concrete situations but as Proust (ibid.) contends, and as I would like to as well, so can other aspects of our interactional behaviour. It is not only language that becomes routinized over interactional micro- and macro-histories, but so are other interactional resources and practices, as successfully shown in many strands of EMCAIL research (see for instance Broth et al., 2017; Deppermann, 2018a; Gerwing & Bavelas, 2004; Schmidt & Deppermann, 2023; Stukenbrock, 2021; and many other studies on the multimodality of social interaction, see particularly section 2.2 of this thesis).

In scenic opera rehearsals, the participants strive to create performance bodies, artistic correlates of *multimodal gestalts* – the assemblages of interactional resources such as gaze, talk, gesture, movement, and facial expression that the fictional characters of the opera production could have performed in their fictional universe to accomplish fictional social (inter)action. To create performance bodies, the participants draw on their intuitive knowledge of social interaction, as they mobilize their resources in multimodal depictions (see also discussion in Hazel, 2018). It is evident how the members of the opera ensemble value the “peripheral” aspects of our visuospatial and vocal/aural behaviour, and their importance for social interaction (in line with Proust, 1919/2003, p. 52). As the peripheral vocal/aural aspects are already predetermined by the music, the opera ensemble can concentrate on the visual bodies and how they can convey a sense of tragedy to the audience. They create these

bodies, explore them, negotiate them, and finally, make them suit the music. The collaborative nature of the rehearsals reflects the aim to design performance bodies for an audience not fully present during the rehearsals. The objective is to create a production that helps the audience understand themselves and the people around them (see Aristotle 1994, 1997); a production in which the different constituents that comprise the performance bodies, and the emotional resonances they create, last in the memories of the audience members.

7. Svensk sammanfattning

Den här avhandlingen undersöker hur deltagare i sceniska operarepetitioner använder sig av *gestaltningar* (eng. *depictions*) för att skapa *performativa kroppar* (eng. *performance bodies*), det vill säga vad operasångarna ska göra på scen för att porträttera föreställningens karaktärer. Termen *gestaltning* kommer från Clarks (2016) psykolingvistiska ramverk och används här för att beskriva en interaktionell praktik som bygger på ikonicitet – likhet mellan tecken och referent. I avhandlingen studeras gestaltningarnas lokala och multimodala konstruktion i naturligt förekommande videoinspelad social interaktion. Det framkommer att gestaltningar är ett värdefullt verktyg för deltagarna under operarepetitioner för att skapa föreställningen och förhandla om dess estetik.

Avhandlingens teoretiska bakgrund återfinns inom multimodal interaktionsanalys (Broth & Keevallik, 2020). Denna metodik har sitt ursprung i etnometodologisk samtalsanalys och interaktionell lingvistik (Selting & Couper-Kuhlen, 2017; Sidnell & Stivers, 2013). Studien bygger på tidigare forskning i social interaktion samtidigt som den använder sig av Clarks kommunikationsteoretiska ramverk och knyter an till begrepp från semiotisk teori (Agawu, 1991; Jakobson, 1971; Peirce, 1955).

Avhandlingen bidrar till tidigare studier om ikoniska fenomen i social interaktion (Cantarutti, 2020, 2021, 2022a, 2022b; Ehmer & Mandel, 2021; Holt & Clift, 2007; Keevallik, 2010, 2014a, 2014b; Schmidt & Deppermann, 2022; Sidnell, 2006; Stukenbrock, 2012, 2014a, 2014b, 2017; Thompson & Suzuki, 2014), multimodalitet och multimodala gestalter (Deppermann, 2013; Mondada, 2014a, 2016; Nevile, 2015; Stukenbrock, 2018, 2021), gemensamt beslutsfattande, förslag och deontik (Asmuß & Oshima, 2012; Clifton et al., 2018; Couper-Kuhlen, 2014; Houtkoop, 1987; Kuroshima, 2023; Lindholm et al., 2020; Lindström,

2017; Magnusson, 2022; Nissi, 2015; Stevanovic, 2012, 2015, 2021) samt studier om interaktionella historier och longitudinell samtalsanalys (Broth et al., 2017; Deppermann, 2018a; Deppermann & Pekarek Doehler, 2021), särskilt de som studerar kreativa och konstnärliga processer över tid (Due, 2016; Hazel, 2018; Norrthon, 2019; Norrthon & Schmidt, 2023; Schmidt & Deppermann, 2023; Yasui, 2013). Följande fyra forskningsfrågor angående gestaltningar under sceniska operarepetitioner undersöks inom ramen för avhandlingen:

1) Hur skapar och besvarar deltagarna gestaltningar med hjälp av multimodala resurser?

2) Vilka interaktionella och semiotiska funktioner har gestaltningar?

3) Vad är relationen mellan gestaltningar och beskrivningar och vad säger den om relationen mellan språk och kropp i social interaktion?

Dessa frågor besvaras i de tre artiklar som tillsammans med föreliggande kapp utgör avhandlingen.

I artikel I ligger fokuset på hur sångare och regissör använder gestaltningar för att skapa performativa kroppar. Artikeln introducerar gestaltningar som fenomen vid sceniska operarepetitioner och visar hur deltagarna konstruerar dessa med hjälp av olika multimodala resurser för att åstadkomma förslag. Gestaltningarna gör den nuvarande versionen av föreställningen tillgänglig för de andra deltagarna i interaktionen och är ett redskap för förhandlingar om föreställningens estetik. I artikeln förs argumentet att referenterna till gestaltningarna (som kommunikativt tecken) är prototyper av mänskligt beteende, det som i avhandlingen benämns som *extroversive semiosis* (efter Jakobson, 1971).

Gestaltningarna refererar även till sig själva och till föreställningens framväxande estetik, i egenskap av att vara föreställningen i dess nuvarande form, i vad som benämns som *introversive semiosis* (Jakobson, 1971). Slutligen förs argumentet att introversive semiosis bär likheter med hur gemensamma erfarenheter påverkar social interaktion i allmänhet.

Artikel II fokuserar på hur sångarna använder gestaltningar för att föreslå performativa kroppar. Under operarepetitionerna orienterar sig sångarna mot regissören som en auktoritet i beslut om både repetitionernas agenda och de performativa kropparna (så kallad *deontisk auktoritet*, se Stevanovic & Peräkylä, 2012). Artikeln syftar till att undersöka hur sångarna konstruerar sina förslag under dessa deontiska omständigheter. I artikeln framkommer det att gestaltningar är multimodala och samskapade fenomen. Koordinationen av gestaltningarna deltagarna sinsemellan reflekterar fördelningen av deontiska rättigheter under de sceniska operarepetitionerna. För att visa för de andra deltagarna att de kommer att utföra en gestaltning som ett led i ett förslag om performativa kroppar gör sångarna rumsliga förflyttningar. Förflyttningarna innebär deontiska anspråk eftersom sångarna i och med dem begär att få göra en gestaltning tillsammans med de övriga deltagarna, och därmed att tillfälligt ta kontroll över den interaktionella agendan. Dessa deontiska anspråk kan båda accepteras och avvisas av de andra deltagarna, genom att de antingen väljer att positionera sig själva i enhet med den gestaltade scenen, antingen som medskådespelare i den, eller som åskådare till den, eller genom att de undviker en sådan positionering. I artikeln visas hur regissören möjliggör en skapandemiljö där deltagarförslag uppmuntras och tas i beaktande, även om regissören inte alltid finner att gestaltningar är relevanta på den lokala interaktionella agendan.

Avhandlingens artikel III fortsätter att undersöka hur sångare föreslår performativa kroppar och fokuserar på hur konstruktionen av förslag, med både gestaltningar och *beskrivningar* (eng. *descriptions*, Clark, 2016), förändras över interaktionella mikrohistorier av gemensamt beslutsfattande. Artikeln visar hur förslag går från att beskrivas till att gestaltas i takt med att de övriga deltagarna uttrycker, i allt högre grad, att de har tillgång till innehållet i ett förslag, att de går med på att en idé fortsätter att utvecklas samt att de håller med om att idén bakom förslaget passar den aktuella föreställningens estetik. Gestaltningar i operarepetitioner ökar alltså i komplexitet över tid. Detta kontrasterar med tidigare forskning om gestaltningars utveckling över tid (Gerwing & Bavelas, 2004), vilket har att göra med gestaltningarnas speciella natur under operarepetitioner som både interaktionella praktiker för att skapa performativ konst, och den konkreta manifestationen av den performativa konsten i sin nuvarande form. Artikeln visar på den dialogiska metod som denna ensemble använder sig av i skapandeprocessen genom att belysa hur varje gestaltning är ett resultat av föreställningens estetik såsom den har manifesterat sig i interaktionen, samtidigt som den förnyar denna estetik. Gestaltningarna stödjer således dubbla sociala handlingar (cf. Rossi, 2018) i sina egenskaper av både kontextkänsliga- och förnyande (jmf. Heritage, 1984).

Avhandlingen visar hur gestaltningar är tidsmässigt heterogena multimodala och samskapade gestalter. De kräver att deltagarna koordinerar multimodala resurser både intrapersonellt och interpersonellt (se Deppermann, 2014). En deltagare som vill gestalta ett förslag utför interaktionellt arbete för att öka sannolikheten för att det gestaltade förslaget ska få positiv återkoppling och för att se till att de andra deltagarna bidrar till gestaltningen på ett sådant sätt att den kan utföras. Vidare åstadkommer deltagarna gestaltningarna stegvis genom att först

mobilisera visuellt rumsliga resurser mot den gestaltade scenen och därmed ge de andra deltagarna möjligheten att samarbeta med den interaktionella konfiguration som gestaltningen innebär. Först därefter adresseras vokala/auditiva resurser till gestaltningen. På motsvarande sätt är gestaltningar i tidiga skeden av beslutsfattandeprocessen ofta formmässigt reducerade och åtföljs av samtidiga beskrivningar. Mot slutet av beslutsfattandeprocessen tenderar de dock att nå sina fulla former med både visuellt rumsliga och vokala/auditiva resurser. I förlängningen visar avhandlingen att gestaltningarnas multimodala samskapade konstruktion är ett utrymme för förhandlingar om deontiska rättigheter. På så vis bidrar studien till den framväxande kunskapen om hur visuellt rumsliga resurser kan användas för att förhandla om deontiska rättigheter (se Barske, 2009; Asmuß & Oshima, 2012; Tuncer, 2015, Stevanovic & Monzoni, 2016; Clifton et al., 2018; Van de Mieroop et al., 2020; Tranekjær et al., 2022; Kuroshima, 2023).

Språkets roll för gestaltningarnas konstruktion och förståelse kan delas in i lexikal semantik, grammatik och sekvensorganisation. Avhandlingen visar hur beskrivningar, det vill säga lexikal semantik, används för att öka intersubjektivitet kring idéer innan mer tvetydiga resurser, nämligen gestaltningar, används (se även Reber, 2012; Stevanovic & Frick, 2014). Grammatik projicerar gestaltningar och möjliggör gemensamt fokus på dem (Auer, 2005, 2009, Keevallik, 2013, 2018) medan det faktum att gestaltningarna inkorporeras i språkliga samtalturer ökar chansen för att de ska få en återkoppling, genom den interaktionella principen för villkorlig relevans (se Sacks et al., 1974; Schegloff, 1984; Deppermann & Streeck, 2018). Utöver semantikens roll fångas inte dessa aspekter av begreppet *beskrivning* (cf. Clark, 1996, 2016) som icke desto mindre är ett användbart begrepp för att förklara en interaktionell praktik som bygger på just lexikal semantik. Vidare argumenterar avhandlingen för ett

holistiskt, integrerat och *emiskt* (cf. Garfinkel, 1967) synsätt på *beskrivning* och *gestaltning*, såväl som på *språk* och *kropp*. Avhandlingen understryker behovet av att undersöka dessa fenomen i situerad social interaktion, såsom de framträder mellan människor, millisekund per millisekund.

Sammanfattningsvis belyser avhandlingen gestaltningars roll i den gemensamma beslutsfattandeprocessen där performativa kroppar till en operaföreställning skapas. Gestaltningar är oundgängliga för beslutsfattandeprocessen, vilket bland annat yttrar sig i regissörens frekventa önskemål om att sångarna ska visa henne vad de föreslår (se särskilt avhandlingens artikel III). Gestaltningarna är tillfälliga gemensamma aktiviteter, performativ konst som kortfattat framförs under diskussioner, för att åstadkomma förslag om hur föreställningen kan spelas. Målet med operarepetitionerna är att frambringa en föreställning av hög kvalitet. Deltagarna samarbetar på olika sätt för att uppnå detta mål och denna avhandling visar hur samarbetet tar sig uttryck i mikrosociologiska detaljer – hur deltagarna positionerar sig själva i interaktionsutrymmet för att hjälpa varandra att skapa scener som de sedan förhandlar och diskuterar i relation till opera som konstform, andra föreställningar och konstupplevelser de upplevt, samt sin intuitiva förståelse av mänskligt handlande. I detta arbete ger operaensemblen avsevärd betydelse till de interaktionella resurser som inom vissa lingvistiska discipliner har betraktats som ”perifera” till det ”centrala” språkliga meddelandet. Därmed befäster avhandlingen synen på social interaktion som multimodal och holistisk (se också Mondada, 2014a; Nevile, 2015; Keevallik, 2018), där inga resurser, varken de som traditionellt betraktas som språkliga, eller de som betraktats som icke-språkliga, på förhand betraktas som väsentligare än andra.

References

- Agawu, V. K. (1991). *Playing with signs: A semiotic interpretation of classical music*. Princeton University Press.
- Albert, S., & vom Lehn, D. (2022). Non-lexical vocalizations help novices learn joint embodied actions. *Language & Communication*, 88, 1–13. <https://doi.org/10.1016/j.langcom.2022.10.001>
- Aristotle. (1994). *Om diktkonsten* (J. Stolpe, Trans.) Alfabeta.
- Aristotle. (1997). *Poetics* (G. Whalley, Trans.). McGill-Queen's University Press.
- Arminen, I. (2005). *Institutional interaction: Studies of talk at work*. Ashgate.
- Ashmore, M., & Reed, D. (2000). Innocence and nostalgia in conversation analysis: The dynamic relations of tape and transcript. *Forum Qualitative Sozialforschung/ Forum Qualitative Research*, 1(3). <https://doi.org/10.17169/fqs-1.3.1020>
- Asmuß, B., & Oshima, S. (2012). Negotiation of entitlement in proposal sequences. *Discourse Studies*, 14(1), 67–86. <https://doi.org/10.1177/1461445611427215>
- Asmuß, B., & Svennevig, J. (2009). Meeting talk: An introduction. *Journal of Business Communication*, 46(1), 3–22. <https://doi.org/10.1177/0021943608326761>
- Atkinson, P. (2006). *Everyday arias: An operatic ethnography*. Rowman Altamira.
- Auer, P. (2005). Projection in interaction and projection in grammar. *Text & Talk*, 25(1), 7–36. <https://doi.org/10.1515/text.2005.25.1.7>
- Auer, P. (2009). On-line Syntax: Thoughts on the temporality of spoken language. *Language Sciences*, 31(1), 1–13. <https://doi.org/10.1016/j.langsci.2007.10.004>
- Balantani, A. (2022). Non-lexical vocalisations + “so_was” as a multimodal package in establishing joint decisions in music rehearsals. *Language & Communication*, 87, 147–160. <https://doi.org/10.1016/j.langcom.2022.07.007>
- Barnes, R., & Moss, D. (2007). Communicating a feeling: The social organization of “private thoughts.” *Discourse Studies*, 9(2), 123–148. <https://doi.org/10.1177/1461445607075339>
- Barske, T. (2009). Same token, different actions: A conversation analytic study of social roles, embodied actions, and ok in German business meetings. *Journal of Business Communication*, 46(1), 120–149. <https://doi.org/10.1177/0021943608325748>
- Berger, E., & Pekarek Doehler, S. (2015). Direct reported speech in storytellings: enacting and negotiating epistemic entitlements. *Text & Talk*, 35(6), 789–813. <https://doi.org/10.1515/text-2015-0023>
- Betz, E., Taleghani-Nikazm, C., & Golato, P. (2020). Mobilizing others: An introduction. In C. Taleghani-Nikazm, E. Betz & P. Golato (Eds.), *Mobilizing others: Grammar and lexis within larger activities* (pp. 1–18). John Benjamins. <https://doi.org/10.1075/slsi.33.01bet>

- Bogen, D. (1999). *Order without rules: Critical theory and the logic of conversation*. SUNY Press.
- Bolden, G. B. (2004). The quote and beyond: Defining boundaries of reported speech in conversational Russian. *Journal of Pragmatics*, 36(6), 1071–1118. <https://doi.org/10.1016/j.pragma.2003.10.015>
- Brandenberger, C., & Hottiger, C. (2018). Sharing perception when using hands-on exhibits in science centres: The case of vocal depiction. *Revue Tranel*, 68, 59–68. https://www.unine.ch/files/live/sites/tranel/files/Tranel/68/59-68_Brandenberger_Hottiger.pdf
- Broth, M. (2011). The performance as interaction between actors and their audience. *Nottingham French Studies*, 50(2), 113–133.
- Broth, M., Cromdal, J., & Levin, L. (2017). Starting Out as a Driver: Progression in Instructed Pedal Work. In Å. Mäkitalo, P. Linell & R. Säljö (Eds.), *Memory Practices and Learning: Interactional, Institutional and Sociocultural Perspectives* (pp. 113–143). Information Age Publishing.
- Broth, M., & Keevalik, L. (2014). Getting ready to move as a couple: Accomplishing mobile formations in a dance class. *Space and Culture*, 17(2), 107–121. <https://doi.org/10.1177/1206331213508483>
- Broth, M., & Keevalik, L. (2020). *Multimodal interaktionsanalys*. Studentlitteratur.
- Broth, M., & Mondada, L. (2013). Walking away: The embodied achievement of activity closings in mobile interaction. *Journal of Pragmatics*, 47(1), 41–58. <https://doi.org/10.1016/j.pragma.2012.11.016>
- Broth, M., & Mondada, L. (2019). Delaying moving away: Place, mobility, and the multimodal organization of activities. *Journal of Pragmatics*, 148, 44–70. <https://doi.org/10.1016/j.pragma.2019.05.013>
- Bühler, K. (1934/2011). *Theory of language: The representational function of language* (D. F. Goodwin, Trans.). John Benjamins.
- Burdelski, M. (2015). Reported speech as a cultural gloss and directive: Socializing norms of speaking and acting in Japanese caregiver-child triadic interaction. *Text & Talk*, 35(5), 575–595. <https://doi.org/10.1515/text-2015-0017>
- Calabria, V. (2023). Co-constructing and other-extending collaborative reported speech in Italian. *Journal of Pragmatics*, 214, 1–17. <https://doi.org/10.1016/j.pragma.2023.04.011>
- Cantarutti, M. N. (2020). *The Multimodal and sequential design of co-animation as a practice for association in English interaction*. PhD thesis, University of York.
- Cantarutti, M. N. (2021). Co-animation and the multimodal management of contextualization problems when jointly “doing being” others. *Journal of Social Interaction. Video Studies of Human Sociality*, 4(4). <https://doi.org/10.7146/si.v4i4.128166>

- Cantarutti, M. N. (2022a). Co-animation in troubles-talk. *Research on Language and Social Interaction*, 55(1), 37–58.
<https://doi.org/10.1080/08351813.2022.2026169>
- Cantarutti, M. N. (2022b). Responsive animation and the negotiation of (shared) self-deprecating attributes and experiences. *Language and Communication*, 87, 205–220.
<https://doi.org/10.1016/j.langcom.2022.08.002>
- Cekaite, A., & Mondada, L. (Eds.). (2020). *Touch in social Interaction: Touch, language, and body*. Routledge.
<https://doi.org/10.4324/9781003026631>
- Chandler, D. (2017). *Semiotics: The basics* (3rd edition). Routledge.
- Clark, H. H. (1996). *Using Language*. Cambridge University Press.
- Clark, H. H. (2016). Depicting as a method of communication. *Psychological Review*, 123(3), 324–347. <https://doi.org/10.1037/rev0000026>
- Clark, H. H. (2019). Depicting in communication. In P. Hagoort (Ed.), *Human Language: From Genes and Brains to Behaviour* (pp. 235–247). MIT Press.
- Clark, H. H., & Gerrig, R. J. (1990). Quotations as demonstrations. *Language*, 66(4), 764–805. <https://doi.org/10.2307/414729>
- Clayman, S. E., & Gill, V. (2004). Conversation analysis. In A. Bryman, & M. A. Hardy (Eds.), *Handbook of data analysis* (pp. 589–606). Sage.
- Clifton, J., Van De Mierop, D., Sehgal, P. & Aneet. (2018). The multimodal enactment of deontic and epistemic authority in Indian meetings. *Pragmatics*, 28(3), 333–360. <https://doi.org/10.1075/prag.17011.cli>
- Couper-Kuhlen, E. (1996). The prosody of repetition: On quoting and mimicry. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in conversation: Interactional studies* (pp. 366–405). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511597862.011>
- Couper-Kuhlen, E. (2007). Assessing and accounting. In Holt, E. & Clift, R. (Eds.), *Reporting Talk: Reported Speech in Interaction* (pp. 81–119). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511486654.005>
- Couper-Kuhlen, E. (2014). What does grammar tell us about social action? *Pragmatics*, 24(3), 623–647.
<https://doi.org/10.1075/prag.24.3.08cou>
- Couper-Kuhlen, E., & Selting, M. (2017) *Interactional linguistics: Studying language in social interaction*. Cambridge University Press.
<https://doi.org/10.1017/9781139507318>
- Deleuze, G. (1969). *La logique du sens*. Minuit.
- Deppermann, A., (2013). Multimodal interaction from a conversation analytic perspective. *Journal of Pragmatics*, 46, 1–7.
<https://doi.org/10.1016/j.pragma.2012.11.014>
- Deppermann, A. (2014). Multimodal participation in simultaneous joint projects: Interpersonal and intrapersonal coordination in paramedic emergency drills. In P. Haddington, T. Keisanen, L. Mondada & M.

- Nevile (Eds.), *Multiactivity in social interaction: Beyond multitasking* (pp. 247–281). John Benjamins. <https://doi.org/10.1075/z.187.09dep>
- Deppermann, A. (2018a). Changes in turn-design over interactional histories: The case of instructions in driving school lessons. In A. Deppermann & J. Streeck (Eds.), *Time in embodied interaction: Synchronicity and sequentiality of multimodal resources* (pp. 293–324). John Benjamins. <https://doi.org/10.1075/pbns.293.09dep>
- Deppermann, A. (2018b). Inferential practices in social interaction: A conversation-analytic account. *Open Linguistics*, 4(1), 35–55. <https://doi.org/10.1515/opli-2018-0003>
- Deppermann, A. & Haugh, M. (2022). Action ascription in social interaction. In A. Deppermann & M. Haugh (Eds.), *Action ascription in social interaction* (pp. 3–28). Cambridge University Press. <https://doi.org/10.1017/9781108673419.001>
- Deppermann, A., Mondada, L. & Pekarek Doehler, S. (2021). Early responses: An introduction. *Discourse Processes*, 58(4), 293–307. <https://doi.org/10.1080/0163853X.2021.1877516>
- Deppermann, A. & Pekarek Doehler (2021). Longitudinal conversation analysis – Introduction to the special issue. *Research on Language and Social Interaction*, 54(2), 127–141. <https://doi.org/10.1080/08351813.2021.1899707>
- Deppermann, A. & Schmidt, A. (2021a). Micro-sequential coordination in early responses. *Discourse Processes*, 58(4), 372–396. <https://doi.org/10.1080/0163853X.2020.1842630>
- Deppermann, A. & Schmidt, A. (2021b). How shared meanings and uses emerge over an interactional history: *Wabi Sabi* in a series of theater rehearsals. *Research on Language and Social Interaction*, 54(2), 203–224.
- Deppermann, A. & Streeck, J. (2018). The body in interaction: Its multiple modalities and temporalities. In A. Deppermann & J. Streeck (Eds.), *Time in embodied interaction: Synchronicity and sequentiality of multimodal resources* (pp. 1–29). John Benjamins. <https://doi.org/10.1075/pbns.293.intro>
- Dietrich, S., Ingo, H., Alter, K., Ischebek, A. & Ackermann, H. (2008). Understanding the emotional expression of verbal interjections: A functional MRI study. *NeuroReport*, 19(18), 1751–1755. <https://doi.org/10.1097/WNR.0b013e3283193e9e>
- Dingemanse, M. (2011). Ideophones and the aesthetics of everyday language in a West-African society. *The Senses and Society*, 6(1), 77–85. <https://doi.org/10.2752/174589311X12893982233830>
- Dingemanse, M. (2020). Between sound and speech: Liminal signs in interaction. *Research on Language and Social Interaction*, 53(1), 188–196. <https://doi.org/10.1080/08351813.2020.1712967>
- Due, B. L. (2016). Co-construction of imagination spaces: A multimodal analysis of the interactional accomplishment of imagination.

- CoDesign: International Journal of CoCreation in Design and the Arts*, 14(3), 1–17. <https://doi.org/10.1080/15710882.2016.1263668>
- Due, B. L. & Lange, S. (2020). Body part highlighting: exploring two types of embodied practices in two sub-types of showing sequences in video-mediated consultations. *Social Interaction: Video-based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.122250>
- Ekberg, K. & LeCouteur, A. (2015). Clients' resistance to therapists' proposals: Managing epistemic and deontic status. *Journal of Pragmatics*, 90, 12–25. <https://doi.org/10.1016/j.pragma.2015.10.004>
- Ekberg, S. (2011). *Making arrangements: Remote proposal sequences and attendant structure phenomena in social interaction*. PhD thesis, University of Adelaide.
- Ehmer, O. (2021). Synchronization in demonstrations: Multimodal practices for instructing body knowledge. *Linguistics Vanguard*, 7. <https://doi.org/10.1515/lingvan-2020-0038>
- Ehmer, O. & Brône, G. (2021). Instructing embodied knowledge: Multimodal approaches to interactive practices for knowledge constitution. *Linguistics Vanguard*, 7. <https://doi.org/10.1515/lingvan-2021-0012>
- Ehmer, O. & Mandel, D. (2021). Projecting action spaces: On the interactional relevance of cesural areas in co-enactments. *Open Linguistics*, 7, 638–665. <https://doi.org/10.1515/opli-2020-0172>
- Emerson, K., Williamson, V. & Wilkinson, R. (2019). Once more, with feeling: Conductors' use of assessments and directives to provide feedback in choir rehearsals. *Musicae Scientiae*, 23(3), 362–382. <https://doi.org/10.1177/1029864919844810>
- Enfield, N. J. (2005). The body as a cognitive artifact in kinship representations. Hand gesture diagrams by speakers of Lao. *Current Anthropology*, 46(1), 1–26. <https://doi.org/10.1086/425661>
- Enfield, N. J. (2009). *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511576737>
- Eriksson, M. (1995). A case of grammaticalization in modern Swedish: The use of *ba* in adolescent speech. *Language Sciences*, 17(1), 19–48. [https://doi.org/10.1016/0388-0001\(94\)00015-A](https://doi.org/10.1016/0388-0001(94)00015-A)
- European Data Protection Board. (2020). *Guidelines 05/2020 on consent under Regulation 2016/679*. Version 1.1. https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf
- Evans, B. & Reynolds, E. (2016). The organization of corrective demonstrations using embodied action in sports coaching feedback. *Symbolic Interaction*, 39(4), 525–556. <https://doi.org/10.1002/symb.255>
- Evans, B. & Lindwall, O. (2020). Show them or involve them? Two organizations of embodied instruction. *Research on Language and Social Interaction*, 53(2), 223–246. <https://doi.org/10.1080/08351813.2020.1741290>

- Fatigante, M., Liberati, V. & Pontecorvo, C. (2010). Transitions in and out of games: How parents and children bracket game episodes at home. *Research on Language and Social Interaction*, 43(3), 346–371. <https://doi.org/10.1080/08351813.2010.497631>
- Ford, C. E. (2004). Contingency and units in interaction. *Discourse Studies*, 6, 27–52. <https://doi.org/10.1177/1461445604039438>
- Ford, C. E., Fox, B. & Thompson, S. A. (1996). Practices in the construction of turns: The ‘TCU’ revisited. *Pragmatics*, 6(3), 427–454. <https://doi.org/10.1075/prag.6.3.07for>
- Fox, B. A. & Robles, J. S. (2010). It’s like mmm: Enactments with it’s like. *Discourse Studies*, 12(6), 715–738. <https://doi.org/10.1177/1461445610381862>
- Fox, B. A., Thompson, S. A. & Ford, C. E. (2013). Conversation analysis and Linguistics. In J. Sidnell & T. Stivers (Eds.) *The Handbook of Conversation Analysis* (pp. 728–740). Wiley-Blackwell. <https://doi.org/10.1002/9781118325001.ch36>
- Frick, M. (2013). Singing and code-switching in sequence closings. *Pragmatics*, 23(3), 243–273. <https://doi.org/10.1075/prag.23.2.03fri>
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Prentice-Hall.
- Garfinkel, H. & Sacks, H. (1970). On formal structures of practical action. In J. C. McKinney & E. A. Tiryakian (Eds.), *Theoretical sociology: Perspectives and developments* (pp. 338–366). Appleton-Century-Crofts.
- Gerwing, J. & Bavelas, J. (2004). Linguistic influences on gesture’s form. *Gesture*, 4(2), 157–195. <https://doi.org/10.1075/gest.4.2.04ger>
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harper and Row.
- Goffman, E. (1981) *Forms of talk*. University of Pennsylvania Press.
- Good, J. S. (2015). Reported and enacted actions: Moving beyond reported speech and related concepts. *Discourse Studies*, 17(6), 663–681. <https://doi.org/10.1177/1461445615602349>
- Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In J. Psathas (Ed.) *Everyday language: Studies in ethnomethodology* (pp. 97–121). Irvington.
- Goodwin, C. (1980). Restarts, pauses and the achievement of a state of mutual gaze at turn-beginning. *Sociological Inquiry*, 50(3-4), 272–302. <https://doi.org/10.1111/j.1475-682X.1980.tb00023.x>
- Goodwin, C. (1981). *Conversational organization. Interaction between speakers and hearers*. Academic Press.
- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96(3), 606–633. <https://doi.org/10.1525/aa.1994.96.3.02a00100>
- Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32(10), 1489–1522. [https://doi.org/10.1016/S0378-2166\(99\)00096-X](https://doi.org/10.1016/S0378-2166(99)00096-X)

- Goodwin, C. (2007a). Participation, stance, and affect in the organization of activities. *Discourse and Society*, 18(1), 53–73.
<https://doi.org/10.1177/0957926507069457>
- Goodwin, C. (2007b). Interactive footing. In E. Holt & R. Clift (Eds.), *Reporting talk: Reported speech in interaction* (pp. 16–46). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511486654.003>
- Goodwin, C. (2013). The co-operative, transformative organization of human action and knowledge. *Journal of Pragmatics*, 46(1), 8–23.
<https://doi.org/10.1016/j.pragma.2012.09.003>
- Goodwin, C. (2018). *Co-operative action*. Cambridge University Press.
<https://doi.org/10.1017/9781139016735>
- Goodwin, C. & Goodwin, M. H. (2004). Participation. In A. Duranti (Ed.), *A companion to linguistic anthropology* (pp. 222–244). Blackwell.
<https://doi.org/10.1002/9780470996522.ch10>
- Goodwin, C., Goodwin, M. H. & Osher, D. (2002). Producing sense with nonsense syllables: Turn and sequence in conversations with a man with severe aphasia. In C. E. Ford, B. A. Fox & S. A. Thompson (Eds.), *The language of turn and sequence* (pp. 56–80). Oxford University Press.
- Goodwin, M. H. (1980). Processes of mutual monitoring implicated in the production of description sequences. *Sociological Inquiry*, 50(3–4), 303–317. <https://doi.org/10.1111/j.1475-682X.1980.tb00024.x>
- Goodwin, M. H. (1990). *He-said-she-said: Talk as social organization among black children*. Indiana University Press.
- Goodwin, M. H. (1997). By-play: Negotiating evaluation in story-telling. In G. R. Guy, C. Feagin, D. Schiffrin & J. Baugh (Eds.), *Towards a social science of language: Papers in honor of William Labov 2: Social interaction and discourse structures* (pp. 77–102). John Benjamins.
<https://doi.org/10.1075/cilt.128.08goo>
- Greer, T. (2013). Word search sequences in bilingual interaction: Codeswitching and embodied orientation toward shifting participant constellation. *Journal of Pragmatics*, 57, 100–117.
<https://doi.org/10.1016/j.pragma.2013.08.002>
- Gudmundsen, J. & Svennevig, J. (2020). Multimodal displays of understanding in vocabulary-oriented sequences. *Social Interaction: Video-Based Studies of Human Sociality*, 3(2).
<https://doi.org/10.7146/si.v3i2.114992>
- Günthner, S. (1999). Polyphony and the ‘layering of voices’ in reported dialogues: An analysis of the use of prosodic devices in everyday reported speech. *Journal of Pragmatics*, 31(5), 685–708.
[https://doi.org/10.1016/S0378-2166\(98\)00093-9](https://doi.org/10.1016/S0378-2166(98)00093-9)
- Haddington, P., Keisanen, T., Mondada, L. & Nevile, M. (Eds.). (2014). *Multiactivity in social interaction*. John Benjamins.
<https://doi.org/10.1075/z.187>

- Harjunpää, K., Deppermann, A. & Sorjonen, M.-L. (2023). Displaying inner experience through language and body in community theater rehearsals. *Human Studies*, 46, 247–271.
<https://doi.org/10.1007/s10746-023-09674-6>
- Haviland, J. B. (2007) Master speakers, master gesturers: A string quartet master class. In S. D. Duncan, E. T. Levy & J. Cassel (Eds.), *Gesture and the dynamic dimension of language: Essays in honor of David McNeill* (pp. 147–172). John Benjamins.
<https://doi.org/10.1075/gs.1.16hav>
- Hayashi, M. (2003). Language and the body as resources for collaborative action: A study of word searches in Japanese conversation. *Research on Language and Social Interaction*, 36(2), 109–141.
https://doi.org/10.1207/S15327973RLSI3602_2
- Hazel, S. (2015). Acting, interacting, enacting: Representing medical practice in theatre performance. In M. Kjaer, J. L. Jørgensen, A. H. Riis (Eds.), *Thematic issues on 'health', Akademisk kvarter/Academic quarter 12* (pp. 44–64). Aalborg University.
<https://doi.org/10.5278/ojs.academicquarter.voi12.2727>
- Hazel, S. (2017). Doing notes: A practice for coordinating ideas development in theatre settings. In D. Day & K. Mortensen (Eds.). Special issue: Inscribed objects in professional practice. *Applied Linguistics and Professional Practice*, 14(2), 149–171.
- Hazel, S. (2018). Discovering interactional authenticity: Tracking theatre practitioners across rehearsals. In S. Pekarek Doehler, J. Wagner & E. González-Martínez (Eds.) *Longitudinal studies on the organization of social interaction* (pp. 255–283). Springer.
https://doi.org/10.1057/978-1-137-57007-9_9
- Heath, C. (1986). *Body movement and speech in medical interaction*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511628221>
- Heath, C. (2002). Demonstrative suffering: The gestural (re)embodiment of symptoms. *Journal of Communication*, 52(3), pp. 597–616.
<https://doi.org/10.1111/j.1460-2466.2002.tb02564.x>
- Heritage, J. (1984a). *Garfinkel and ethnomethodology*. Polity Press.
- Heritage, J. (1984b). A change-of-state token and aspects of its sequential placement. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 299–345). Cambridge University Press.
- Heritage, J. (1998). Oh-prefaced responses to inquiry. *Language in Society*, 27(3), 291–334. <https://doi.org/10.1017/S0047404500019990>
- Heritage, J. (2002). Oh-prefaced responses to assessments: A method of modifying agreement/disagreement. In C. E. Ford, B. A. Fox & S. A. Thompson (Eds.), *The language of turn and sequence* (pp. 196–224). Oxford University Press.
- Heritage, J. (2010). Conversation analysis: Practices and methods. In D. Silverman (Ed.), *Qualitative sociology*, 3rd ed. (pp. 208–230). Sage.

- Heritage, J. & Sorjonen, M.-L. (1994). Constituting and maintaining activities across sequences: And-prefacing as a feature of question design. *Language in Society*, 23(1), 1–29.
<https://doi.org/10.1017/S0047404500017656>
- Hofstetter, E. (2020a). Nonlexical “moans”: Response cries in board game interactions. *Research on Language and Social Interaction*, 53(1), 42–65. <https://doi.org/10.1080/08351813.2020.1712964>
- Hofstetter, E. (2020b). Thinking with the body: Embodying thinking as a practice in board games. In S. Wiggins & K. O. Cromdal (Eds.), *Discursive psychology and embodiment: Beyond subject-object binaries* (pp. 247–273). Palgrave Macmillan.
https://doi.org/10.1007/978-3-030-53709-8_10
- Hofstetter, E. (2021). Achieving preallocation: Turn transition practices in board games. *Discourse Processes*, 58(2), 113–133.
<https://doi.org/10.1080/0163853X.2020.1816401>
- Hofstetter, E. & Keevallik, L. (2023). Prosody is used for real-time exercising of other bodies. *Language & Communication*, 88, 52–77.
<https://doi.org/10.1016/j.langcom.2022.11.002>
- Hofstetter, E., Keevallik, L. & Löfgren, A. (2021). Suspending syntax: Bodily strain and progressivity in talk. *Frontiers in Communication*, 6.
<https://doi.org/10.3389/fcomm.2021.663307>
- Holt, E. (2007). ‘I’m eyeing your chop up mind’: Reporting and enacting. In E. Holt & R. Clift (Eds), *Reporting talk: Reported speech in interaction*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511486654.004>
- Holt, E. & Clift R. (2007). *Reporting talk: Reported speech in interaction*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511486654>
- Hopper, R. (2005). A cognitive agnostic in conversation analysis: When do strategies affect spoken interaction. In H. te Molder & J. Potter (Eds.), *Conversation and cognition* (pp. 134–160). Cambridge University Press. <https://doi.org/10.1017/CBO9780511489990.007>
- Houtkoop, H. (1987). *Establishing agreement: An analysis of proposal-acceptance sequences*. De Gruyter.
<https://doi.org/10.1515/9783110849172>
- Hsu, H.-C., Brône, G. & Feyaerts, K. (2020). In other gestures: Multimodal iteration in cello master classes. *Linguistics Vanguard*, 7.
<https://doi.org/10.1515/lingvan-2020-0086>
- Huisman, M. (2001). Decision-making in meetings as talk-in-interaction. *International Studies of Management & Organization*, 31(3), 69–90.
<https://doi.org/10.1080/00208825.2001.11656821>
- Hutchby, I. & Wooffitt, R. (2008). *Conversation analysis*. Polity Press.
- Ivaldi, A. (2016). Students’ and teachers’ orientation to learning and performing in music conservatoire lesson interactions. *Psychology of Music*, 44(2), 202–218. <https://doi.org/10.1177/0305735614562226>

- Iversen, C. & Flinkfeldt, M. (2020). Kroppen som objekt och resurs för interaktion. In M. Broth & L. Keevallik (Eds.), *Multimodal interaktionsanalys* (pp. 269–285). Studentlitteratur.
- Jakobson, R. (1971) *Language in relation to other communication systems. Selected Writings Volume II: Word and Language*. De Gruyter Mouton. <https://doi.org/10.1515/9783110873269.697>
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp.13–31). John Benjamins. <https://doi.org/10.1075/pbns.125.02jef>
- Jokipohja, A.-K., Lilja, N. (2022). Depictive hand gestures as candidate understandings. *Research on Language and Social Interaction*, 55(2), 123-145. <https://doi.org/10.1080/08351813.2022.2067425>
- Jönsson, L., (2005) Tal i tal. In J. Anward & B. Nordberg (Eds.), *Samtal och grammatik: Studier i svenskt samtalspråk* (pp. 89–108). Studentlitteratur.
- Kaukoma, T., Peräkylä, A. & Ruusuvuori, J. (2015). How listeners use facial expression to shift the emotional stance of the speaker's utterance. *Research on Language and Social Interaction*, 48(3), pp. 319–341. <https://doi.org/10.1080/08351813.2015.1058607>
- Keevallik, L. (2001). Tracing grammaticalization of oota “wait” in Estonian conversation. In I. Tragel (Ed.), *Papers in Estonian cognitive linguistics* (pp. 119–144). Tartu University Press.
- Keevallik, L., (2010). Bodily quoting in dance correction. *Research on Language and Social Interaction*, 43(4), 401–426. <https://doi.org/10.1080/08351813.2010.518065>
- Keevallik, L. (2013a). The interdependence of bodily demonstrations and clausal syntax. *Research on Language and Social Interaction*, 46(1), 1–21. <https://doi.org/10.1080/08351813.2013.753710>
- Keevallik, L. (2013b). Here in time and space: Decomposing movement in dance instruction. In P. Haddington, L. Mondada & M. Nevile (Eds.), *Interaction and mobility: Language and the body in motion* (pp. 343–370) . De Gruyter. <https://doi.org/10.1515/9783110291278.345>
- Keevallik, L. (2014a). Turn organization and bodily-vocal demonstrations. *Journal of Pragmatics*, 65, 103–120. <https://doi.org/10.1016/j.pragma.2014.01.008>
- Keevallik, L. (2014b). Having a ball: Immaterial objects in dance instruction. In M. Nevile, P. Haddington, T. Heinemann & M. Rauniomaa (Eds.). *Interacting with objects: Language, materiality, and social activity* (pp. 249–268). John Benjamins. <https://doi.org/10.1075/z.186.11kee>
- Keevallik, L. (2015). Coordinating the temporalities of talk and dance. In A. Deppermann & S. Günthner (Eds.), *Temporality in interaction* (pp. 309-336). John Benjamins. <https://doi.org/10.1075/slsi.27.10kee>
- Keevallik, L. (2017). Linking performances: The temporality of contrastive grammar. In R. Laury, M. Etelämäki & E. Couper-Kuhlen (Eds.),

- Linking clauses and actions in social interaction* (pp. 54–72). Finnish Literature Society.
- Keevallik, L. (2018). What does embodied interaction tell us about grammar? *Research on Language and Social Interaction*, 51(1), pp. 1–21. <https://doi.org/10.1080/08351813.2018.1413887>
- Keevallik, L., (2020a). Grammatik. In M. Broth & L. Keevallik (Eds), *Multimodal interaktionsanalys* (pp. 77–95). Studentlitteratur.
- Keevallik, L. (2020b). Linguistic structures emerging in the synchronization of a Pilates class. In C. Taleghani-Nikazm, E. Betz & P. Golato, *Mobilizing others: Grammar and lexis within larger activities* (pp. 147–173). John Benjamins. <https://doi.org/10.1075/slsi.33.06kee>
- Keevallik, L. (2021). Vocalizations in dance classes teach body knowledge. *Linguistics Vanguard*, 7(4). <https://doi.org/10.1515/lingvan-2020-0098>
- Keevallik, L. & Ekström A. (2019). How to take the floor as a couple: Turn-taking in lindy hop jam circles. *Visual Anthropology*, 32(5), 423–444. <https://doi.org/10.1080.089449468.2019.1671750>
- Keevallik, L., Ogden, R. (2020). Sounds on the margins of language at the heart of interaction. *Research on Language & Social Interaction*, 53(1), 1–18. <https://doi.org/10.1080/08351813.2020.1712961>
- Keisanen, T., Rauniomaa, M. & Haddington, P. (2014). Suspending action: From simultaneous to consecutive ordering of multiple courses of action. In P. Haddington, T. Keisanen, L. Mondada & M. Nevile (Eds.), *Multiactivity in social interaction: Beyond multitasking* (pp. 109–133). John Benjamins. <https://doi.org/10.1075/z.187.04kei>
- Kendon, A. (1990). *Conducting Interaction: Patterns of behaviour in focused encounters*. Cambridge University Press.
- Kendrick, K. & Drew, P. (2016). Recruitment: Offers, requests, and the organization of assistance in interaction. *Research on Language and Social Interaction*, 49(1), 1–19. <https://doi.org/10.1080/08351813.2016.1126436>
- Kent, A. (2012). Compliance, resistance and incipient compliance when responding to directives. *Discourse Studies*, 14(6), 711–730. <https://doi.org/10.1177/146144561245748>
- Kirsh, D. (2011) How marking in dance constitutes thinking with the body. *Versus: Quaderni di Studi Semiotici Vol 113–115* (pp. 179–210). Bompiani.
- Kress, G. & van Leeuwen, T. (2001). *Multimodal discourse*. Hodder Arnold.
- Krug, M. (2022). Temporal procedures of mutual alignment and synchronization in collaborative meaning-making activities in a dance rehearsal. *Frontiers in Communication*, 7. <https://doi.org/10.3389/fcomm.2022.957894>
- Kuroshima, S. (2023). When a request turn is segmented: Managing the deontic authority via early compliance. *Discourse Studies*, 25(1), 114–135. <https://doi.org/10.1177/14614456221136975>

- Lefebvre, A. (2018). Reading and embodying the script during the theatrical rehearsal. *Language and Dialogue*, 8(2), 261–288.
<https://doi.org/10.1075/ld.00015.lef>
- Lefebvre, A. & Mondada, L. (2023). Interactional contingencies in rehearsing a theatre scene: The consequentiality of body arrangements as action unfolds. *Human Studies*, 46, 303–335.
<https://doi.org/10.1007/s10746-023-09669-3>
- Levinson, S. C. (1979). Activity types and language. *Linguistics*, 17(5–6), 365–399. <https://doi.org/10.1515/ling.1979.17.5-6.365>
- Levinson, S. C. (2013). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 103–130). Wiley Blackwell. <https://doi.org/10.1002/9781118325001.ch6>
- Lindholm, C., Stevanovic, M., Valkeapää, T. & Weiste, E. (2020). Writing: A versatile resource in the treatment of the clients' proposals. In C. Lindholm, M. Stevanovic & E. Weiste (Eds.), *Joint decision making in mental health: An interactional approach* (pp. 187–210). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-43531-8_8
- Lindström, A. (2017). Accepting remote proposals. In G. Raymond, G. H. Lerner & J. Heritage (Eds.), *Enabling human conduct: Studies of talk-in-interaction in honor of Emanuel A. Schegloff* (pp. 125–142). John Benjamins. <https://doi.org/10.1075/pbns.273.07lin>
- Linell, P. (2005). *The written language bias in Linguistics: Its nature, origins, and transformations*. Routledge.
- Linell, P. (2009). *Rethinking language, mind and world dialogically: Interactional and contextual theories of human sense-making*. Information Age Publishing.
- Linell, P. (2020). Epilog: Multimodalitet och språk. In M. Broth & L. Keevallik (Eds.), *Multimodal interaktionsanalys* (pp. 417–432). Studentlitteratur.
- Lodge, D. (1990). *After Bakhtin: Essays on fiction and criticism*. Routledge.
- Magnusson, S. (2021). Establishing jointness in proximal multiparty decision-making: The case of collaborative writing. *Journal of Pragmatics*, 181, 32–48. <https://doi.org/10.1016/j.pragma.2021.05.003>
- Magnusson, S. (2022). *Boosting young citizens' deontic status: Interactional allocation of rights to decide in participatory democracy meetings*. PhD thesis, Södertörn University.
- Maynard, D. W. (2013). Everyone and no one to turn to: Intellectual roots and contexts for conversation analysis. In J. Sidnell & T. Stivers (Eds.) *The Handbook of Conversation Analysis* (pp. 11–30). Wiley-Blackwell. <https://doi.org/10.1002/9781118325001.ch2>
- McGurk, H. & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 764–748. <https://doi.org/10.1038/264746a0>
- Melander, H. (2012). Knowing how to play the game of jump rope: Participation and stancetaking in a material environment. *Journal of Pragmatics*, 44(11), pp. 1434–1456.
<https://doi.org/10.1016/j.pragma.2012.06.018>

- Menke, C. (2004). The presence of tragedy. *Critical Horizons*, 5(1), 201–225.
<https://doi.org/10.1163/1568516042653521>
- Messner, M. (2020). Gesangliche Demonstrationen als Interaktion Praktik in der Orchesterprobe. *Gesprächsforschung: Online-Zeitschrift zur verbalen Interaktion*, 21, 309–345. <http://www.gespraechsforschung-online.de/fileadmin/dateien/heft2020/si-messner.pdf>
- Messner, M. (2022). Abbiamo detto con te no che tu hai ta da di da dim (Moves right hand on the beat): The interplay of semiotic modes in chamber music lessons under a multimodal and interactional perspective. *Frontiers in Communication*, 7.
<https://doi.org/10.3389/fcomm.2022.877184>
- Mondada, L. (2005). La constitution de l'origo déictique comme travail interactionnel des participants: Une approach praxéologique de la spatialité. *Intellectica*, 41-42(2–3), 75–100.
<https://doi.org/10.3406/intel.2005.1722>
- Mondada, L. (2006). Video recording as the reflexive preservation-configuration of phenomenal features for analysis. In H. Knoblauch, J. Raab, H.-G. Soeffner & B. Schnettler (Eds.), *Video Analysis* (pp. 51–68). Lang.
- Mondada, L. (2009). Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics*, 41(10), 1977–1997.
<https://doi.org/10.1016/j.pragma.2008.09.019>
- Mondada, L. (2012). Talking and driving: Multiactivity in the car. *Semiotica*, 19(1), 223–256. <https://doi.org/10.1515/sem-2012-0062>
- Mondada, L. (2013). The conversation analytic approach to data collection. In J. Sidnell & T. Stivers (Eds.) *The Handbook of Conversation Analysis* (pp. 32–55). Wiley-Blackwell.
<https://doi.org/10.1002/9781118325001.ch3>
- Mondada, L. (2014a). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*. 65, 137–156.
<https://doi.org/10.1016/j.pragma.2014.04.004>
- Mondada, L. (2014b). The temporal orders of multiactivity: Operating and demonstrating in the surgical theatre. In P. Haddington, T. Keisanen, L. Mondada & M. Nevile (Eds.), *Multiactivity in social interaction: Beyond multitasking* (pp. 33–75). John Benjamins.
<https://doi.org/10.1075/z.187.02mon>
- Mondada, L. (2015). The facilitator's task of formulating citizens' proposals in political meetings: Orchestrating multiple embodied orientations to recipients. *Gesprächsforschung*, 16, 1–62,
<http://gespraechsforschung-online.de/fileadmin/dateien/heft2015/gadeppermann.pdf>
- Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*, 20(3), 336–366.
https://doi.org/10.1111/josl.1_12177

- Mondada, L. (2018a). Questions on the move: The ecology of question-answer sequences in mobility settings. In A. Deppermann & J. Streeck (Eds.), *Time in embodied interaction: Synchronicity and sequentiality of multimodal resources* (pp. 161–202). John Benjamins. <https://doi.org/10.1075/pbns.293.05mon>
- Mondada, L. (2018b). Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. *Research on Language and Social Interaction*, 51(1), 85–106. <https://doi.org/10.1080/08351813.2018.1413878>
- Mondada, L. (2018c). Controversial issues in participatory urban planning: An ethnomethodological and conversation analytic historical study. In S. Pekarek Doehler, S. Wagner & E. González-Martínez (Eds.), *Longitudinal Studies on the Organization of Social Interaction* (pp. 287–328). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-57007-9_10
- Mondada, L. (2021). How early can embodied responses be? Issues in time and sequentiality. *Discourse Processes*, 58(4), 397–418. <https://doi.org/10.1080/0163853X.2020.1871561>
- Mondada, L. (2022b). Adjusting step-by-step trajectories in public space: The micro-sequentiality of approaching and refusing to be approached. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion*, 23, 36–65.
- Nevile, M., (2015). *The embodied turn* in research on language and social interaction. *Research on Language and Social Interaction*, 48(2), 121–151. <https://doi.org/10.1080/08351813.2015.1025499>
- Niemelä, M. (2010). The reporting space in conversational storytelling: Orchestrating all semiotic channels for taking a stance. *Journal of Pragmatics*, 42(12), 3258–3270. <https://doi.org/10.1016/j.pragma.2010.06.015>
- Nishizaka, A. (2017). The perceived body and embodied vision in interaction. *Mind, Culture and Activity*, 24(2), 110–128. <https://doi.org/10.1080/10749039.2017.1296465>
- Nissi, R. (2015). From entry proposals to a joint statement: Practices of shared text production in multiparty meeting interaction. *Journal of Pragmatics*, 79, 1–21. <https://doi.org/10.1016/j.pragma.2015.01.002>
- Nissi, R. & Stevanovic, M. (2021). Interactional means of teaming up: Enacting the features of contemporary working life in a theater performance. *Multimodal communication*, 10(2). <https://doi.org/10.1515/mc-2020-0004>
- Norris, S. (2004). *Analyzing multimodal interaction*. Routledge.
- Norrthon, S. (2019). To stage an overlap: The longitudinal, collaborative and embodied process of staging eight lines in a professional theatre rehearsal process. *Journal of Pragmatics*, 142, 171–184. <https://doi.org/10.1016/j.pragma.2019.01.015>
- Norrthon, S. (2020). *Teaterrepetitionens interaktion*. PhD thesis, Stockholm University.

- Norrthon, S. (2021). Framing in theater rehearsals. *Journal of Applied Linguistics and Professional Practice*, 15(2), 187–214.
<https://doi.org/10.1558/jalpp.20370>
- Norrthon, S. & Schmidt, A. (2023). Knowledge accumulation in theatre rehearsals: The emergence of a gesture as a solution for embodying a certain aesthetic concept. *Human Studies*, 46, 337–369.
<https://doi.org/10.1007/s10746-022-09654-2>
- Ochs, E. (1979). Transcription as theory. In E. Ochs & B. B. Shieffelin (Eds.), *Developmental Pragmatics* (pp. 43–72). Academic Press.
- Park, S. H. & Hepburn, A. (2022). The benefits of a Jeffersonian transcript. *Frontiers in Communication*, 7.
<https://doi.org/10.3389/fcomm.2022.779434>
- Peirce C.S. (1955) *Logic as semiotic: The theory of signs*. In Buchler (Ed.), *Philosophical writings of Peirce*. Dover Publications.
- Persson, R. (2020). Fonetisk form. In M. Broth & L. Keevallik (Eds.), *Multimodal interaktionsanalys*. Studentlitteratur.
- Pfeiffer, M. & Weiss, C. (2022). Reenactments during tellings: Using gaze for initiating reenactments, switching roles and representing events. *Journal of Pragmatics*, 189, 92–113.
<https://doi.org/10.1016/j.pragma.2021.11.017>
- Plato (1992). *Republic* (M. J. Levett, Trans.). Hackett Publishing.
- Pomerantz, A. & Denvir, P. (2007). Enacting the institutional role of chairperson in upper management meetings: The interactional realization of provisional authority. In F. Cooren (Ed.), *Interacting and organizing: Analyses of a management meeting* (pp. 31–51). Lawrence Erlbaum.
- Potter, J. & Edwards, D. (2013). Conversation analysis and psychology. In J. Sidnell & T. Stivers (Eds.) *The Handbook of Conversation Analysis* (pp. 701–724). Wiley-Blackwell.
<https://doi.org/10.1002/9781118325001.ch35>
- Profazi, N. & Miecznikowski, J. (2023). Social interaction is among people. Legal, technical, and ethical explorations about personal information and its removal in talk-in-interaction as data. Chord-talk-in-interaction. <https://www.chord-talk-in-interaction.usi.ch/news/feeds/36387>
- Proust, M. (1919/2003). *In the shadow of young girls in flower* (J. Grieve, Trans.). The Penguin Press.
- Råman, J. (2019). *The multimodal and collaborative aspects of demonstrations in teaching budo sports*. PhD thesis, University of Oulu.
- Råman, J. & Haddington, P. (2018). Demonstrations in sports training: Communicating a technique through parsing the return-practice in the budo class. *Multimodal communication*, 7, 1–19.
<https://doi.org/10.1515/mc-2018-0001>
- Reber, E. (2012). *Affectivity in interaction: Sound objects in English*. John Benjamins. <https://doi.org/10.1075/pbns.215>

- Reber, E. (2021). *Quoting in parliamentary question time: Exploring recent change*. Cambridge University Press.
<https://doi.org/10.1017/9781108869898>
- Reed, D. J., (2015). Relinquishing in musical masterclasses: Embodied action in interactional projects. *Journal of Pragmatics*, 89, 31–49.
<https://doi.org/10.1016/j.pragma.2015.09.006>
- Reed, D. J. (2019). Touch and talk: Detailing embodied experience in the music masterclass. *Social Semiotics*, 30(5), 625–645.
<https://doi.org/10.1080/10350330.2019.1631431>
- Reed, D. J. & Szczepek Reed, B. (2013). Building an instructional project: Actions as components of music masterclasses. In B. Szczepek Reed & G. Raymond (Eds.), *Units of talk – Units of action* (pp. 313–342). John Benjamins. <https://doi.org/10.1075/slsi.25.10ree>
- Reed, D. J. & Szczepek Reed, B. (2014). The emergence of learnables in music masterclasses. *Social Semiotics*, 24(4), 446–467.
<https://doi.org/10.1080/10350330.2014.929391>
- Robinson, J. D. (2013). Overall structural organization. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 257–280). Wiley-Blackwell. <https://doi.org/10.1002/9781118325001.ch13>
- Robinson, J. D. & Stivers, T. (2001). Achieving activity transitions in primary-care encounters: From history taking to physical examination. *Human Communication Research*, 27(2), 253–298.
<https://doi.org/10.1111/j.1468-2958.2001.tb00782.x>
- Rohman Roth, A.-C. (2022). “Här ska kraften vara på!” *Interaktion vid körövningar: En studie om luft och andning*. Licentiate thesis, Linnaeus University.
- Rossano, F. (2013). Gaze in conversation. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 308–329). Wiley-Blackwell.
<https://doi.org/10.1002/9781118325001.ch15>
- Rossi, G. (2018). Composite social actions: The case of factual declaratives in everyday interaction. *Research on Language and Social Interaction*, 51(4), 379–397. <https://doi.org/10.1080/08351813.2018.1524562>
- Rönqvist, S. & Lindström J. (2021). Turn continuation and gesture: “And then”-prefacing in multi-party conversations. *Frontiers in Communication*, 6. <https://doi.org/10.3389/fcomm.2021.670173>
- Sacks, H. (1975). Everyone has to lie. In M. Sanches & B. G. Blount (Eds.), *Sociocultural dimensions of language use* (pp. 57–80). Academic Press.
- Sacks, H. (1978). Some technical considerations of a dirty joke. In J. Schenkein, (Ed.), *Studies in the organization of conversational interaction* (pp. 249–269). Academic Press.
- Sacks, H. (1984a). On doing “being ordinary”. In J. M Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in Conversation analysis* (pp. 413–429). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511665868.024>

- Sacks, H. (1984b). Notes on methodology. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in Conversation analysis* (pp. 21–27). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511665868.005>
- Sacks, H. (1992). *Lectures on conversation Volume I & II (Fall 1964-Spring 1972)*. Blackwell.
- Sacks, H. & Schegloff, E. A. (2002). Home position. *Gesture*, 2(2), 133–146.
<https://doi.org/10.1075/gest.2.2.02sac>
- Sacks, H., Schegloff, E. A. & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735. <http://dx.doi.org/10.2307/412243>
- Sambre, P. & Feyaerts, K. (2017). Embodied musical meaning-making and multimodal viewpoints in a trumpet master class. *Journal of Pragmatics*, 122, 10–23.
<https://doi.org/10.1016/j.pragma.2017.09.004>
- Sandlund, E. (2014). Prescribing conduct: Enactments of talk or thought in advice-giving sequences. *Discourse Studies* 16(5), 645–666.
<https://doi.org/10.1177/1461445614539065>
- Savijärvi, M. & Ihalainen, L. (2021). How an improvised scene emerges in theatre rehearsal. In J. Lindström, R. Laury, A. Peräkylä & M.-L. Sorjonen (Eds.) *Intersubjectivity in action: Studies in language and social interaction* (pp. 373–395). John Benjamins.
<https://doi.org/10.1075/pbns.326.17sav>
- Schegloff, E. A. (1968). Sequencing in conversational openings. *American Anthropologist*, 70(6), 1075–1095.
<https://doi.org/10.1525/aa.1968.70.6.02a00030>
- Schegloff, E. A. (1984). On some gestures' relation to talk. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 266–295). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511665868.018>
- Schegloff, E. A. (1996a). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. A. Schegloff & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 52–133). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511620874.002>
- Schegloff, E. A. (1996b). Confirming allusions: Toward and empirical account of action. *American Journal of Sociology*, 104(1), 161–216.
<https://www.jstor.org/stable/2782190>
- Schegloff, E. A. (1997). Practices and actions: Boundary cases of other-initiated repair. *Discourse Processes*, 23, 499–545.
<https://doi.org/10.1080/01638539709545001>
- Schegloff, E.A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511791208>
- Schmidt, A. (2018). Prefiguring the future: Projections and preparations within theatrical rehearsals. In A. Depperman & J. Streeck (Eds.), *Time in embodied interaction: Synchronicity and sequentiality of*

- multimodal resources (pp. 231–260). John Benjamins.
<https://doi.org/10.1075/pbns.293.07sch>
- Schmidt, A. & Deppermann, A. (2022). Showing and telling: How directors combine embodied demonstrations and verbal descriptions to instruct in theater rehearsals. *Frontiers in Communication*, 7.
<https://doi.org/10.3389/fcomm.2022.955583>
- Schmidt, A. & Deppermann, A. (2023). On the emergence of routines: An interactional micro-history of rehearsing a scene. *Human Studies*, 46, 273–302. <https://doi.org/10.1007/s10746-022-09655-1>
- Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J., Bergmann, P., Birkner, K., Couper-Kuhlen, E., Deppermann, A., Gilles, P., Günthner, S., Hartung, M., Kern, F., Mertzlufft, C., Meyer, C., Morek, M., Oberzaucher, F., Peters, J., Quasthoff, U., Schütte, W., Stukenbrock, A. & Uhmann, S. (2011). A system for transcribing talk-in-interaction: GAT 2, *Gesprächsforschung: Online-Zeitschrift zur verbalen Interaktion*, 12, 1–51. <http://www.gespraechsforschung-online.de/fileadmin/dateien/heft2011/px-gat2-englisch.pdf>
- Sidnell, J. (2006) Coordinating gesture, talk and gaze in reenactments. *Research on Language and Social Interaction*, 39(4), 377–409.
https://doi.org/10.1207/s15327973rlsi3904_2
- Sidnell, J. (2010). *Conversation analysis: An introduction*. Wiley Blackwell.
- Sidnell J. & Enfield, N. J. (2014). The ontology of action, in interaction. In N. J. Enfield, P. Kockelman, & J. Sidnell (Eds.), *Cambridge handbook of linguistic anthropology* (pp. 423–446). Cambridge University Press.
<https://doi.org/10.1017/CBO9781139342872.020>
- Sidnell, J. & Stivers, T. (Eds.) (2013). *The handbook of conversation analysis*. Wiley-Blackwell. <https://doi.org/10.1002/9781118325001>
- Soulaimani, D. (2018). Talk, voice and gestures in reported speech: Toward an integrated approach. *Discourse Studies*, 20(3), 361–376.
<https://doi.org/10.1177/1461445618754419>
- Smith, D. W. (2006). The concept of the simulacrum: Deleuze and the overturning of Platonism. *Continental Philosophy Review*, 38, 89–123. <https://doi.org/10.1007/s11007-006-3305-8>
- Speer, S. A. & Stokoe, E. (2012). Ethics in action: Consent-gaining interactions and implications for research practice. *British Journal of Social Psychology*, 53, 54–73. <https://doi.org/10.1111/bjso.12009>
- Stevanovic, M. (2012). Establishing joint decisions in a dyad. *Discourse Studies*, 14(6), 779–803. <https://doi.org/10.1177/1461445612456654>
- Stevanovic, M. (2013a). *Deontic rights in interaction: A conversation analytic study on authority and cooperation*. PhD thesis, University of Helsinki.
- Stevanovic, M. (2013b). Constructing a proposal as a thought: A way to manage problems in the initiation of joint decision-making in Finnish workplace interaction. *Pragmatics*, 23(3), 519–544.
<https://doi.org/10.1075/prag.23.3.07ste>

- Stevanovic, M. (2015). Displays of uncertainty and proximal deontic claims: The case of proposal sequences. *Journal of Pragmatics*, 78, 84–97. <https://doi.org/10.1016/j.pragma.2014.12.002>
- Stevanovic, M. (2018). Social deontics: A nano-level approach to human power play. *Journal for the Theory of Social Behaviour*, 48(3), 369–389. <https://doi.org/10.1111/jtsb.12175>
- Stevanovic, M. (2021). Deontic authority and the maintenance of lay and expert identities during joint decision making: Balancing resistance and compliance. *Discourse Studies*, 23(5), 670–689. <https://doi.org/10.1177/14614456211016821>
- Stevanovic, M., Frick M. (2014). Singing in interaction. *Social Semiotics*, 24(4), 495–513. <https://doi.org/10.1080/10350330.2014.929394>
- Stevanovic, M. & Monzoni, C. (2016). On the hierarchy of interactional resources: Embodied and verbal behaviour in the management of joint activities with material objects. *Journal of Pragmatics*. 103, 15–32. <https://doi.org/10.1016/j.pragma.2016.07.004>
- Stevanovic, M. & Peräkylä, A. (2012). Deontic authority in interaction: The right to announce, propose, and decide. *Research on Language and Social Interaction*, 45(3), 297–321. <https://doi.org/10.1080/08351813.2012.699260>
- Stevanovic, M. & Peräkylä, A. (2014). Three orders in the organization of human action: On the interface between knowledge, power, and emotion in interaction and social relations. *Language in Society*, 43(2), 185–207. <https://doi.org/10.1017/S0047404514000037>
- Stevanovic, M. & Svennevig, J. (2015). Introduction: Epistemics and deontics in conversational directives. *Journal of Pragmatics*, 78, 1–6. <https://doi.org/10.1016/j.pragma.2015.01.008>
- Stivers, T. (2008). Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language and Social Interaction*, 41(1), 31–57. <https://doi.org/10.1080/08351810701691123>
- Stivers, T., Mondada, L., Steensig, J. (2011). Knowledge, morality and affiliation. In T. Stivers, L. Mondada & J. Steensig (Eds.), *Knowledge, morality and affiliation in social interaction* (pp. 3–24). Cambridge University Press. <https://doi.org/10.1017/CBO9780511921674.002>
- Stivers, T. & Rossano, F. (2010). Mobilizing response. *Research on Language and Social Interaction*, 43(1), 3–31. <https://doi.org/10.1080/08351810903471258>
- Stivers, T. & Sidnell, J. (2005). Introduction: Multimodal interaction. *Semiotica*, 2005(156), 1–20. <https://doi.org/10.1515/semi.2005.2005.156.1>
- Stivers, T. & Sidnell, J. (2013). Introduction. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 1–8). Wiley-Blackwell. <https://doi.org/10.1002/9781118325001.ch1>

- Stivers, T. & Sidnell, J. (2016). Proposals for activity collaboration. *Research on Language and Social Interaction*, 49(2), 148–166.
<https://doi.org/10.1080/08351813.2016.1164409>
- Stoeckl, H. & Messner, M. (2021). Tam pam pam pam and mi – fa – sol: Constituting musical instructions through multimodal interaction in orchestra rehearsals. *Multimodal Communication*, 10(3), 193–209.
<https://doi.org/10.1515/mc-2021-0003>
- Streeck, J. (1993). Gesture as communication. I: Its coordination with gaze and speech. *Communication Monographs*, 60(4), 275–29.
<https://doi.org/10.1080/03637759309376314>
- Streeck, J. (2002). Grammars, words and embodied meanings: On the uses and evolution of so and like. *Journal of Communication*, 52(3), 581–596. <https://doi.org/10.1111/j.1460-2466.2002.tb02563.x>
- Streeck, J. (2008). Depicting by gesture. *Gesture*, 8(3), 285–301.
<https://doi.org/10.1075/gest.8.3.02str>
- Streeck, J. (2009). Gesturecraft: The manu-facture of meaning. *Gesture studies 2*. Amsterdam: John Benjamins Publishing Company.
- Streeck, J., (2013) Interaction and the living body. *Journal of Pragmatics*, 46, pp. 69–90. <https://doi.org/10.1016/j.pragma.2012.10.010>
- Streeck, J. (2018). Times of rest: Temporalities of some communicative postures. In A. Deppermann & J. Streeck (Eds.), *Time in embodied interaction: Synchronicity and sequentiality of multimodal resources* (pp. 325–350). John Benjamins.
<https://doi.org/10.1075/pbns.293.10str>
- Streeck, J. & Hartge, U. (1992). Previews: Gestures at the transition place. In P. Auer & A. di Luzio (Eds.), *The contextualization of language* (pp. 135–158). John Benjamins. <https://doi.org/10.1075/pbns.22.10str>
- Stukenbrock, A. (2012). Imagined spaces as a resource in interaction. In E. De Stefani, A.-D. Gazin & A.-C. Ticca (Eds.), *Space in Social Interaction*. Université de Neuchâtel: *Bulletin Suisse de Linguistique Appliqué*.
<https://libra.unine.ch/handle/123456789/9351>
- Stukenbrock, A. (2014a). Pointing to an ‘empty’ space: Deixis-am-phantasma in face-to-face interaction. *Journal of Pragmatics* 74, 70–93.
<https://doi.org/10.1016/j.pragma.2014.08.001>
- Stukenbrock, A. (2014b). Take the words out of my mouth: Verbal instructions as embodied practices. *Journal of Pragmatics*, 65, 80–102.
<https://doi.org/10.1016/j.pragma.2013.08.017>
- Stukenbrock, A. (2017). Intercorporeal phantasms: Kinesthetic alignment with imagined bodies in self-defense training. In C. Meyer, J. Streeck & S. Jordan (Eds.), *Intercorporeality. Emerging socialities in interaction* (pp. 237–263). Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780190210465.003.0009>
- Stukenbrock, A. (2018). Forward-looking: Where do we go with multimodal projections? In A. Deppermann & J. Streeck (Eds.), *Time in embodied interaction: Synchronicity and sequentiality of multimodal resources*

- (pp. 32–68). John Benjamins.
<https://doi.org/10.1075/pbns.293.01stu>
- Stukenbrock, A. (2021). Multimodal gestalts and their change over time: Is routinization also grammaticalization. *Frontiers in Communication*, 6.
<https://doi.org/10.3389/fcomm.2021.662240>
- Sunakawa, C. (2018). Bodily shadowing: Learning to be an orchestral conductor. In A. Deppermann & J. Streeck (Eds.), *Time in embodied interaction: Synchronicity and sequentiality of multimodal resources* (pp. 203–230). John Benjamins.
<https://doi.org/10.1075/pbns.293.06sun>
- Svennevig, J. (2011). Leadership style in managers' feedback in meetings. In J. Angouri & M. Marra (Eds.), *Constructing identities at work* (pp. 17–39). Palgrave Macmillan. https://doi.org/10.1057/9780230360051_2
- Swedish Research Council (2017). *Good Research Practice*.
https://www.vr.se/download/18.5639980c162791bbfe697882/1555334908942/Good-Research-Practice_VR_2017.pdf
- Szczepek Reed, B. (2010). *Analyzing conversation: An introduction to prosody*. Palgrave Macmillan.
- Szczepek Reed, B. (2012). Beyond the particular: Prosody and the coordination of actions. *Language and Speech*, 55(1), 13–34.
<https://doi.org/10.1177/0023830911428871>
- Szczepek Reed, B. (2021). Singing and the body: Body-focused and concept-focused vocal instruction. *Linguistics Vanguard*, 7(4).
<https://doi.org/10.1515/lingvan-2020-0071>
- Szczepek Reed, B. (2023). 'Go on keep going': The instruction of sustained embodied activities. *Discourse Studies*, 25(5), 692–717.
<https://doi.org/10.1177/14614456231153578>
- Szczepek Reed, B., Reed D. J. and Haddon, E., (2013). NOW or NOT NOW: Coordinating restarts in the pursuit of learnables in vocal master classes. *Research on Language and Social Interaction*, 46(1), 22–46.
<https://doi.org/10.1080/08351813.2013.753714>
- Ten Have, P. (2004). Reflections on transcription. Text based on a talk, given at the Colloque "l'Oral en contexte: des objectifs aux méthodes", Université Paul Valéry – Montpellier III, 17 November 2001.
<http://www.paultenhavenl/Transcription-rv2.pdf>
- Thompson, S. A. & Suzuki, R. (2014). Reenactments in conversation: gaze and reciprocity. *Discourse Studies*, 16(6), pp. 816–846.
<https://doi.org/10.1177/1461445614546259>
- Thompson, S. Fox, B. A., Raymond, C. W. (2021). The grammar of proposals for joint activities. *Interactional Linguistics*, 1(1), 123–151.
<https://doi.org/10.1075/il.20011.tho>
- Ticca, A. C. (2014). Managing multiactivity in a travel agency: Making phone calls while interacting with customers. In P. Haddington, T. Keisanen, L. Mondada & M. Nevile (Eds.), *Multiactivity in social interaction: Beyond multitasking* (pp. 191–223). John Benjamins.
<https://doi.org/10.1075/z.187.07tic>

- Tolins, J. (2013). Assessment and direction through nonlexical vocalizations in music instruction. *Research on Language and Social Interaction*, 46(1), 47–64. <https://doi.org/10.1080/08351813.2013.753721>
- Tolins, J. & Samermit, P. (2016). GIFs as embodied enactments in text-mediated conversation. *Research on Language and Social Interaction*, 49(2), 75–91. <https://doi.org/10.1080/08351813.2016.1164391>
- Trager, G. L. (1958). Paralanguage: A first approximation. *Studies in Linguistics*, 13, 1–12.
- Tranekjær, L., Due, B. L. & Nielsen, M. F. (2022). Constructing and negotiating the professional identity of ‘leader’ by suggesting and challenging improvement of professional practices: Deontics in a four-part sequential structure. *Discourse Studies*, 24(5), 640–661. <https://doi.org/10.1177/14614456221108604>
- Traxler, M. J. (2012). *Introduction to psycholinguistics: Understanding language science*. Wiley-Blackwell.
- Tuncer, S. (2015). Walking away: An embodied resource to close informal encounters in offices. *Journal of Pragmatics*, 76, 101–116. <https://doi.org/10.1016/j.pragma.2014.11.012>
- Tutt, D. & Hindmarsh, J. (2011). Reenactments at work: Demonstrating conduct in data sessions. *Research on Language and Social Interaction*, 44(3), 211–236. <https://doi.org/10.1080/08351813.2011.591765>
- University of Toronto. Social Sciences and Humanities Research Ethics Board (SSH REB). (2005). Guidelines for ethical conduct in participant observation. <https://www.mcgill.ca/mqhrq/files/mqhrq/participant-observation-guidelines.pdf>
- Urbanik, P. & Svennevig, J. (2021). Action-depicting gestures and morphosyntax: The function of gesture-speech alignment in the conversational turn. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.689292>
- Usatch, E. P. (2000). *A conversation analytic study of social actors’ methods for accomplishing activities in interaction*. PhD thesis, University of Southern Illinois.
- Van De Mierop, D., Clifton, J. & Verhelst, A. (2020). Investigating the interplay between formal and informal leaders in a shared leadership configuration: A multimodal conversation analytical study. *Human Relations*, 73(4), 490–515. <https://doi.org/10.1177/0018726719895077>
- Versteeg, W. (2018). *“How do you know?”: Everyday negotiations of expert authority*. PhD thesis, University of Twente.
- Weeks, P., (1996). A rehearsal of a Beethoven passage: An analysis of correction talk. *Research on Language and Social Interaction*, 29(3), 247–290. https://doi.org/10.1207/s15327973rlsi2903_3
- Wilkinson, R., Beeke, S. & Maxim, J. (2010). Formulating actions and events with limited linguistic resources: Enactment and iconicity in agrammatic aphasic talk. *Research on Language and Social Interaction*

- Interaction*, 43(1), 57–84.
<https://doi.org/10.1080/08351810903471506>
- Wu, H. (2016). Modified resayings of reported speech in Mandarin conversation. *Chinese Language and Discourse*, 7(1), 105–135.
<https://doi.org/10.1075/cld.7.1.04wu>
- Wu, R.-J. R. (2022). Gestural repair in Mandarin conversation. *Discourse Studies*, 24(1), 65–93. <https://doi.org/10.1177/14614456211037>
- Yasui, E. (2013). Collaborative idea construction: The repetition of gestures and talk during brainstorming. *Journal of Pragmatics*, 46(1), 157–172.
<https://doi.org/10.1016/j.pragma.2012.10.002>
- Zinken, J. & Ogiermann E. (2011). How to propose an action as objectively necessary: the case of Polish *trzeba x* (“one needs to x”). *Research on Language and Social Interaction*. 44(3), 263–287.
<https://doi.org/10.1080/08351813.2011.591900>

Appendices

Appendix 1: Transcription conventions

The transcription conventions used in this work is based on Jefferson (2004), with developments for visuospatial conduct by Mondada (2018).

(.)	pause in the stream of speech <0,3 seconds.
(0.4)	pause in the stream of speech >0,3 seconds.
=speech	speech that latches onto the previous turn.
speech=	speech that latches onto the subsequent turn.
[speech]	overlapping speech.
speech:	prolonged speech sound (additional :: equals additional length).
<u>speech</u>	emphasis.
spee-	interrupted speech.
SPEECH	speech produced relatively louder.
°speech°	speech produced relatively quieter.
>speech	speech produced relatively faster.
<speech>	speech produced relatively slower.
speech?	rising intonation at the end of a turn constructional unit (TCU).
speech,	slightly rising intonation at the end of a TCU.
speech.	falling intonation at the end of a TCU.
↑speech	rising pitch amid a phrase.
↓speech	falling pitch amid a phrase.
<i>speech</i>	code-switching.
hhh.	audible outbreath.
.hhh.	audible inbreath.
Sphheech	aspiration amid speech.
sp(h)eech	laughter-infused speech.
(speech)	uncertain word.
(xx)	inaudible word, x = one syllable.
(())	analyst comments or descriptions.
speech	visuospatial conduct (VC) delimited using the same symbol.
*-->	VC continues across subsequent lines.

-->*	VC continues until the same symbol is reached.
par	participant performing VC is identified with small caps in margin.
im	image line.
#	identifies the moment in interaction an image has been captured.

Appendix 2: Consent information

Appendix 2.1 First e-mail to the ensemble

E-mail heading: Viktig information om videoinspelning [name of performance]/Important information concerning video recording [name of performance]

This e-mail was sent to: Conductor, director, scenographer, costume designer, light designer, all (14) solo singers, all (8) stand-ins, assistant director, stage manager, all (2) prompters, language coach, all (2) production assistants and all (2) pianists).

Viktig information om videoinspelning / Important information concerning video recording (information in English below)

Hej på er alla!

Jag är verkligen tacksam för att jag som forskare får följa er repetitionsprocess, och det är så spännande för mig att sitta och titta! Jag skulle behöva göra videoinspelningar av några repetitioner för att kunna titta närmare på hur ni kommunicerar kring föreställningen med hjälp av röst, kropp och tal. Videoinspelningarna skulle användas i min forskning och undervisning men de går att göras anonyma med olika tekniker om ni vill det. Inspelningarna kommer att hanteras på ett säkert sätt så att inga obehöriga får se dem.

Jag hoppas verkligen att det ska vara OK med er att jag gör detta eftersom att det vore så värdefullt för min forskning. Om man inte vill vara med på en inspelning går det jättebra att säga till mig innan eller efter, så tar jag bort den inspelningen. Det går även jättebra att höra av sig till mig om man har frågor, antingen på plats, maila till agnes.lofgren@liu.se, eller ringa till mig på 073-xxxxxxx.

Vi ses på repetitionerna! Vänliga hälsningar, Agnes Löfgren

Dear all!

I am really grateful to have the opportunity to observe your repetitions, thank you for that! In order to proceed with my analysis, I need to make video recordings of some of the repetitions. This is to be able to look more closely at how you use your voice, body and speech to communicate in the working process. The recordings will be used in my research and education, but they can be anonymized with the aid of different techniques if you wish. The recordings will be dealt with in a confidential manner and will not be accessed by any unauthorized individual.

I sincerely hope that this will be OK with everybody involved, since it would be of such value for my research. If, for however reason, you do not wish to be recorded, please contact me before or after a recording, and I will delete it. Of course, you are all very welcome to get in touch with me in case you have any questions. You can just come up to me at the site, e-mail me at: agnes.lofgren@liu.se or call me +46 xxxxxxxx.

See you all at rehearsals, Kindest wishes,
Agnes Löfgren

Appendix 2.2 E-mail to the choir

Hej!

Tack så mycket för att jag fick vara med och spela in på måndagens repetition av NN! Här kommer lite skriftlig information om det som jag gör.

Jag gör videoinspelningar för att jag ska kunna titta på hur man kommunicerar kring föreställningen med hjälp av röst, kropp och tal. Jag fokuserar främst på regissör [NN]. Videoinspelningarna används i min forskning och undervisning och de går att göras helt anonyma med olika tekniker om ni vill det. Inspelningarna hanteras på ett säkert sätt så att inga obehöriga får se dem.

Jag hoppas verkligen att det går bra för er, eftersom att det vore så värdefullt för min forskning. **Men det är jätteviktigt för mig att alla känner sig bekväma med inspelningarna och ni har självklart full rätt att inte vilja vara med. Om man inte vill vara med på en inspelning går det jättebra att säga till mig innan eller efter, så tar jag bort de delar av inspelningen där man är med.** Det går även jättebra att höra av sig till mig om man har frågor, antingen på plats, maila till agnes.lofgren@liu.se, eller ringa till mig på 073-xxxxxxx.

Tusen tack igen!

Vänligen,

Agnes Löfgren, doktorand i språkvetenskap, Linköpings universitet

ENGLISH

Dear all!

Thank you for letting me observe the repetition of NN this Monday!
This e-mail contains some written information about my work.

I am video recording some of the [NN] repetitions to be able to look at how people use voice, body and speech to communicate in the working process, focusing on mainly director [NN]. The recordings will be used in my research and education, but they can be anonymized with the aid of different techniques if you wish. The recordings will be dealt with in a confidential manner and will not be accessed by any unauthorized individual.

I sincerely hope that this will be OK with everybody involved, since it would be of such value for my research. **If, for however reason, you do not wish to be recorded, please contact me before or after a recording, and I will of course delete the parts that you are in.** Of course, you are all very welcome to get in touch with me in case you have any questions. You can just come up to me at the site, e-mail me at: agnes.lofgren@liu.se or call me +46 73-xxxxxxx.

Thank you again!

Kindest wishes,

Agnes Löfgren, PhD student in linguistics, Linköping university

Appendix 2.3 Final e-mail to the ensemble

Hej alla!

Tack för att jag har fått följa er repetitionsprocess med NN! Det har så givande för mig och det ska bli jättespännande att gå igenom materialet för att så småningom skriva om det.

I det här mailet kommer lite mer detaljerad information om hur videoinspelningarna ska användas. Materialet kommer att hanteras och

användas i enlighet med Vetenskapsrådets etiska riktlinjer för forskning (som går att finna på: <http://www.codex.vr.se/forskningshumsam.shtml>). Videomaterialet kommer att sparas på externa lösenordskyddade hårddiskar och inga obehöriga kommer att se det. Det som är i fokus för den vetenskapliga analysen är samtalet kring hur ni gör under repetitionerna och övergångarna mellan samtal och repetition.

Ni har självklart fortfarande möjlighet att säga att ni inte vill att jag ska använda de delar av inspelningarna där man är med, och då är det bara att höra av sig till mig. Likaså går det bra att säga att ni inte går med på en av användningarna nedan, men däremot de andra. Om man sedan tidigare har en specifik överenskommelse med mig om hur de delar där man själv är med ska användas så gäller den självklart fortfarande.

Videospelningarna kommer att användas enligt följande:

1. Transkriptioner (skriftlig form) av videospelningarna kan publiceras i vetenskapliga texter. Dessa anonymiseras. Exempel på hur det kan se ut finns i bifogad fil.
2. Stillbilder från videofilmerna kan publiceras i vetenskapliga texter. Dessa kan anonymiseras vid önskemål. Exempel på hur det kan se ut finns i bifogad fil.
3. Några sekunders anonymiserade avsnitt från videospelningarna kan visas tillsammans med vetenskapliga publikationer.
4. Korta klipp om några minuter ur videospelningarna tillsammans med transkriptionen kan visas i undervisning- och i forskningssammanhang (konferenser, seminarium, datasessioner, vetenskapligt samarbete). Videomaterialet kommer garanterat inte att spridas.

Hör gärna av er om ni har frågor eller på annat sätt vill diskutera mitt arbete, det gör jag gärna!

Återigen, tusen tack för att jag har fått vara med. Det har varit så kul att få träffa er alla! Stort grattis till de superfina recensionerna, och lycka till med alla framtida föreställningar!

Allt gott,
Agnes Löfgren 073-xxxxxxx agnes.lofgren@liu.se

Dear all,

I want to thank you for letting me be a part of the rehearsal procedure of NN. It has been very rewarding for me and I'm looking forward to commencing the analyzes.

This e-mail contains elaborated details on how the video recordings will be used. The material will be handled and used according to the ethical codes of conduct of the Swedish research council (which can be consulted on: <http://www.codex.vr.se/forskninghumsam.shtml>). The video recordings will be stored on password-protected external hard drives and will not be accessed by unauthorized individuals. Central to my analysis will be your conversations about how to proceed during rehearsals, and the transitions between talking and rehearsing.

Of course, you still have the possibility of withdrawing your participation and in that case, you can just get in touch with me. You may also disagree on one or several of the ways in which I will use the material (as stated below) but agree on the others. If you already have an agreement with me on how I shall use the parts that you are in, then that agreement is still valid.

The video recordings will be used accordingly:

1. Transcriptions (written form) of the video recordings may be published in scientific writings. They will be anonymized. An example of what this may look like can be found attached to this e-mail.
2. Still images (frame grabs) from the video recordings may be published in scientific writings. They can be anonymized if you wish. An example of what this may look like can be found attached to this e-mail.
3. A few seconds of anonymized video clips from the recordings may be shown together with scientific publications.
4. Short clips of a few minutes from the video recordings may be shown in education and research (conferences, seminars, data sessions, scientific collaboration). The video material is guaranteed not to be spread outside of this context.

Please get in touch with me if you have any questions or if you want to discuss my work in any way. I am always open for that.

Thank you all again, it's been a pleasure to meet you all! Congratulations on the excellent reviews, and the best of luck for all the future shows!

Best regards,
Agnes Löfgren
0046 73-xxxxxxxxx agnes.lofgren@liu.se

Articles

The articles associated with this thesis have been removed for copyright reasons. For more details about these see:

<https://doi.org/10.3384/9789180754026>

FACULTY OF ARTS AND SCIENCES

Linköping Studies in Arts and Sciences No. 863
Studies in Language and Culture no. 37
Department of Culture and Society

Linköping University
SE-581 83 Linköping, Sweden

www.liu.se

At the Faculty of Arts and Sciences at Linköping University, research and doctoral studies are carried out within interdisciplinary research environments, often addressing broad problem areas. Linköping Studies in Arts and Sciences is the Faculty's own series for publishing research. This thesis comes from the graduate school in Language and Culture, Division of Language, Culture and Interaction at the Department of Culture and Society.