Radboud University

The Co-Construction of Duplo

A Multimodal Interaction Analysis of a Multiparty
Collaborative Goal-Oriented Replication-Building Task

Bachelor Thesis
English Language and Culture – Radboud University Nijmegen
15 August 2018

Joey Slijpen

Supervisor: Dr. J. Geenen

Teacher who will receive this document: Dr. J.G. Geenen

Title of document: The Co-Construction of Duplo

Name of course: Bachelor Thesis English Linguistics

Date of submission: 15 August 2018

The work submitted here is the sole responsibility of the undersigned, who has neither committed plagiarism nor colluded in its production.

Signed

Name of student: Joey Slijpen

Abstract

This thesis uncovers how social actors make use of meaning-making signs to collaborate towards a shared goal by applying a Multimodal Interaction Analysis (Norris, 2004, 2011). This thesis investigates the interactions of a visually-complex replication task in the context of social interdependence theory and research on spatial cognition, which allows for the incorporation of multiple modes and collaborative concepts to contextualise the collaborative replication task. The groups consisted of one builder and two assistants that had to work collaboratively to replicate a Duplo structure through communication and the assistants' building tools. The focus is on the communication during the problem sequences. This thesis is relevant to interaction in which goal-shared collaboration is prevalent. As such, the thesis argues that deixis plays a leading role in the reorientation of attention during the task. Another observation is that during problem sequences the priority of gaze allocation indicates that completing the task takes scope over social and cultural norms of gaze usage. The thesis also found that the interactions were triadically oriented, which indicated by the way the social actors divided and executed their roles. The findings create an insight into the co-construction of meaning within the problem sequence through deixis, gaze and their triadically-oriented social role division, which showcase that there are patterns, albeit with individual variation, to allocate attention within these interactions.

Keywords: Multimodal Interaction Analysis, multiparty collaboration, social interdependence, deixis, mediated action, spatial complexity.

Table of Contents

1. Introduction

Communication and interaction are central to our lives. This is reflected in our everyday lives, which happen in the real world through the material actions we employ to interact. Little is known about the material specifics of goal-directed collaborative tasks that are so prevalent within several aspects of everyday life. This is important within multiple sectors such as education, business, medicine and every other sector that requires people to interact collaboratively. It has been shown that effectively collaborating on problems can have a positive effect on everyone involved (Johnson, Johnson, & Smith, 2007). It is therefore meaningful to learn more about the real-time manifestations of collaboration by observing collaborative interactions within the context of its real-time materiality.

In this thesis, I investigate the materiality of collaborative interaction within the framework of the Multimodal Interaction Analysis. This investigation is realised through a multiparty goal-directed collaborative spatially complex replication building task in which social actors simulate real-life collaborative interactions through a naturalistic task. The complexity of the task is sufficient to elicit communication without external factors. The Multimodal Interaction Analysis as a frame is vital to elucidate how social actors in real-time situations allocate, express and perceive actions to co-create meaning within interactions (Norris 2004, 2011). This analytical method and setting leads to the following question: in overcoming obstacles during multiparty goal-oriented communicative tasks, to what extent does the (co)-construction of multiple communicative modes contribute to the resolution of ambiguity for visual-kinesthetic and information-based problems?

The analysis has provided three salient observations/findings. The first is that social actors use deictic gestures to redirect attention during the problem sequences, which occur within the boundaries of the minimal cooperative principle. Another matter that has come to light in the analysis is that social actors prioritise the building process and information grounding over the upholding of interactional and cultural norms in their use of gaze. The final observation is that even though the interactions seemed to be dyadic, the meaning-making signs of the social actors indicate that there is a triadic interaction in which there was a difference in roles and quantity of modes used.

The first section of this thesis functions as a literature review to introduce previous research and identify the gaps to cover for the present research. It features an overview of multimodality and some previous research into gaze and gesture, followed by previous research and theoretical background theories and research revolving collaboration. The final sub-section will provide key notions regarding spatiality, which features concepts such as deixis. The reviewed material is vitally important to understand the reasoning and rationale behind the findings.

After the relevant literature is reviewed the following section provides the rationale of the study through the empirical methodology. This is required to understand the task's collaborative properties and assert how the task design logically follows from the previous research on collaboration. It will first feature the goal of the research and will end with a sub-section on the design properties.

When the task's collaborative and interactional properties are depicted, the analytical methodology is featured to clarify how and why a Multimodal Interaction Analysis is the most suitable form of analysis for the task and the corresponding research question. The first sub-section will elucidate what the unit of analysis is, followed by the corresponding analytical tools required to utilise this unit. These tools

are contextualized through the transcription protocol in the second sub-section and finally, the limitations that follow from this analytical method will be presented.

The analysis section presents the findings that follow from the methodological sections to utilise the previous sections' information to stipulate the findings through careful analysis. The use of deixis, gaze and the triadic nature of the interactions are divided into three sub-sections which all deal with a separate phenomenon through analysis of a different transcript for each section.

The conclusion section provides a small summary of the research and subsequently brings together the findings in the analysis section. These findings are used to provide some potential implications and final remarks.

The final section is the discussion section. This section provides some reflection on the process of the research, some speculation on salient observations that are not covered by the literature and suggestions for further research.

2. Literature Review

This section will elucidate previous research to provide the required information and context to understand the analysis of the present task. First, the general trend acknowledging the multimodality of all social actors will be discussed with a particular focus on the deployment of gaze and forms of deixis in real-time interactions. These resources can help direct the attention of others while simultaneously fulfilling other communicative functions. Consequently, research will be reviewed on collaboration and social interdependence to clarify what social interdependence is within collaboration and what is required to allow it to manifest. Then research revolving collaborative tasks through social interdependence will be reviewed, followed by a multimodal interaction analysis of a multiparty cooperation to display the saliency of multimodal interaction analysis concerning the materiality of multiparty interactions. Finally, a review of earlier research regarding spatial elements and attention's effect on problem solving will be provided. The research will be focused on mental transformation tasks and the role of gesture and deixis to facilitate the resolution of spatially complex tasks.

2.1 Multimodality

Over the last couple of decades communication among people has received considerable attention in academia. Originally, the focus was always on language as the superordinate mode through which communication occurs. The focus on language as *the* only important mode has since then decreased because of the notion that all language is multimodal (Norris, 2004). The notion that language is multimodal is embedded in the social, cultural and historical background that is present within all

communication and manifests through multiple forms of meaning-making signs (Norris, 2004; Kress & Van Leeuwen, 2001).

Kress and van Leeuwen (2001) conceptualise modes as a semiotic resource with known meanings and regularities attached to it. Their theory was a response to the shift from monomodal to multimodal within cultural disciplines. Within semiotics, Kress & van Leeuwen felt the need for a semiotic theory that reflects the multiplicity within all the different (inter)sections of the contemporary world with a theory that can account for all the meaning signs needed to reflect this world. They argue that the multiple modes that are required for every form of communication have only taken place when there has been an articulation and some form of interpretation. They construe the meaning-making signs as being 'imported' from other contexts into their own to make new signs, while at the same time meaning signifiers are also given meaning through the material context surrounding what the articulator does when producing certain signs.

Norris (2004) takes this notion of modes in a real-time setting to create a framework to analyse real-time interaction through the Multimodal (Inter)action Analysis. The switch to multimodality from language as *the* superordinate mode is reflected by the emphasis on the interplay of modes rather than the modes in isolation. The modes which constitute interaction are defined as heuristic units. Heuristic units are units of analysis that are based on experience, which makes them suitable to reflect the flexibility and materiality of real-time interaction. Norris explicitly acknowledges that spoken language is an important mode just like all other modes, but not necessarily the *most* important mode in all possible contexts. The acknowledgement of these flexible hierarchical positions of different modes is incremental to allow for an effective analysis.

The most studied mode aside from spoken language is the study of gesture. A gesture is what can be defined as a "deliberately expressive movement [that has] a sharp boundary of onset and that [is] seen as an excursion, rather than as a result in any sustained change of position" (cited in Norris, 2004, p. 28). McNeill (2005) argues for the binding of gesture in the form of imagery and speech as being synchronous and, in that sense, belonging together. This synchrony also depicts the interplay of modes on a smaller scale for the way social actors think and communicate, which displays the interconnectedness of modes during interaction. McNeill supports this synchrony through multiple examples to display the speech-gesture bond. One of these examples is the synchrony of stuttering people that while stuttering their gestural depiction similarly 'stutters' by pausing mid-stroke. The congenitally blind gesture as well, which supports the synchrony between gesture and spoken language. The congenitally blind also provide additional evidence; they have no visual components, nor have they any past with gestures, yet Iverson and Goldin-Meadow (1997) found that they gesture as frequently as sighted subjects.

Gesture is usually divided into iconic, metaphoric, deictic and beat gestures. Iconic gestures are gestures that depict pictorial content and are used to mimic what an individual communicates verbally through co-verbiage (Norris, 2004). Metaphoric gestures depict pictorial content as well, but then as abstract ideas or categories, which are given form through the imagery portrayed by the specific gesture. Deictic gestures are usually pointing movements, which can be to real-world entities, but also to ideas or notions as if given a physical reality. Beat gestures are, as the name implies, gestures that have to do with beat-like movements. During conversations gesture seems to be able to improve cohesion and can function as a way for the speaker and the listener to create a more coherent story and provide easier transitions (McNeill & Levy, 1993).

The task will require social actors to interact in ways that require attention shifts and a certain understanding of complex physical forms when they are explained. Gestures, and especially deictic gestures, are important to gain insight into allocation of attention during these interactions.

Gaze as a mode is somewhat difficult to establish due to its individual nature and sometimes unsystematic distribution (Norris, 2004). To gaze is, as is implied in its name, to stare at something with an intent or focus. Kendon (1967) discusses the function of gaze between speaker and hearer and how this interaction can influence the gaze pattern. He has found that, even though there are individual differences, the hearer gazes more at the speaker than the other way around. It also showed that glances back and forth between the recipient and something else were about equal length. This assumption only works for western countries as the gaze distribution is based on factors such as cultural background and even differs among individuals (Norris, 2004). For the use of gaze within an interactional setting there are, within the cultural and social boundaries of the social actors, some guidelines or regularities.

Rossano (2012) champions the view that the interactional use of gaze is also dependent on the ongoing course of action within their environment and the nature of the discourse subject. Rossano suggests that a social actors' turn-taking behaviour regarding gaze, can, for example, be delayed within specific situations. He stipulates that the important factor for gaze behaviour is not necessarily the 'competing environment', which is advocated by Goodwin (1984), but rather the 'sequential environment' within the regulatory process of gaze. The competing environment argues that social actors compete for turns whereas the sequential environment suggest a natural and responsive form of turn-taking. The competing environment that Goodwin stipulates was considered to have too many irregularities even though there are cases in

which it yields similar results to Rossano's (2012) theory. This means that extended reciprocal gaze of the recipient is dependent on whether the speaker is used a 'turn constructional unit' which ended with an 'adjacency-pair-based sequence' or an 'extended telling sequence'. An adjacency-pair-based sequence is a sequence that asks for a request or offer, which as a result asks of the hearer to elicit some form of response or action. The extended-telling sequence is a sequence that indicates that an extended narrative is coming in which no turn-switch is supposed to occur. The difference between the two is in its expectations of the hearer within the context of their conversation. Within a Multimodal Interaction Analysis, these insights might generate a new way to look at the distribution of gaze during the task and specifically within the context of the problem sequence.

This sub-section has provided an overview of multimodality and some additional information of specific modes in general. The previous research indicates that the use of deictic gesture and gaze has yet to be tested within the context of a multiparty collaborative effort.

2.2 Collaboration and Problem Solving

Johnson and Johnson champion a perspective on collaboration and experiential learning which is that cooperation can provide better results through specific forms of social interdependence (Johnson & Johnson, 1989). Their theory of social interdependence stipulates that cooperation can provide a better learning experience as well as in some senses better long-lasting results rather than the individual social Darwinism, which has been prevalent for a long time within the field of education. The focus on education, is, in this case, too narrow for the interpretation of social

interdependence. Social interdependence is when the accomplishment of every individual's goals is affected by the actions of others (Johnson & Johnson, 1989).

Social interdependence theory originates from Deutsch's (1949), research on the relationship between goals and type of groups, which Johnson and Johnson expands upon (1989). Social interdependence can be split between positive interdependence and negative interdependence (Johnson et al., 2007). Negative interdependence is the perception that a goal can only be obtained through the completive failure of other social actors involved, which results in the obstruction of each other's efforts towards their goals. Positive social interdependence is when individuals perceive that they can collaboratively reach their goal as the means towards reaching their individual goals, which as a result promotes intraparty effort to achieve both their collaborative and individual roles. This collaborative function is only relevant when there is a situation in which there is an actual dependence on the other social actors for the completion of the goal (Van der Vegt & Van de Vliert, 2002).

One of the forms in which the positive interdependence can manifest is through promotive interaction (Johnson & Johnson, 1974, 1989; Johnson et al., 2007). Promotive interaction is a diverse form of interaction which is encompassed by all forms of encouragement and facilitation to collaboratively complete, achieve or produce something which is in co-accordance with a shared goal. This type of interaction is exemplified through matters such as mutual help, exchange of resources, (positive) effective communication, mutual influence and trust (Johnson et al., 2007).

The promotive interaction does not only improve the quantity of interaction but also the quality through an expansion of the self-interest of individuals in a group (Johnson et al., 2007). The emotional investment social actors spend in achieving goals,

the notion of working together and the openness involved in sharing something facilitates the shift from self-interest to a mutual interest which as a result is shown to make joint efforts more effective (Johnson et al., 2007). This happens through three psychological processes, i.e., the substitutability, inducibility and cathexis (Deutsch, 1949). The substitutability is the degree to which actions of one social actor substitutes for the actions of another, irreducibility is the openness to being influenced and influencing others and the cathexis is the investment of energy in objects other than one's self. These three are positively influenced by the positive interdependence and therefore lend itself to a more fruitful collaboration. In the present research, the cooperative nature of the task allows for positive interdependence and promotive interaction, which therefore provides additional context for the interactions. Promotive interaction as a theory can help elucidate and contextualise how social actors work together to realise their mutual goal.

Buchs, Gilles, Antonietti, and Butera (2016) did a study on the effect of cooperation, which is embedded in the frame of social interdependence as cooperative notions. They made three groups: one group that learns individually, one group that receives dyadic instruction on cooperation as a concept and a final group that has cooperative interaction within dyads. The results indicate that from group one to three there was a progressive linear trend in the immediate as well as the post-test results. The effect sizes are relatively low, which they attribute to the low affinity with mathematical elements of the students. The study shows that not only awareness of cooperation can have a positive effect, but also that the promotion and application of these concepts can provide even better results for the individuals as well as the group in a whole, which indicates how pivotal the role of cooperation is when collaborating towards a mutual goal.

Bavelas, Coates, and Johnson (2002) conducted an experiment embedded in similar cooperative notions to investigate what the role of gaze is within an interaction between two social actors in a collaborative process. Twenty-four students were told to form dyads without previous knowledge of each other. The focus of the experiment is on the duration and direction of the gaze and the number of listener responses. The study was conducted to find out more about how 'back channels' in interaction are timed and utilised within a conversation. Back-channel responses refer to the brief signals social actors make to their interlocutor to indicate attention or awareness of the conversation. They propose that speaker gaze creates an opportunity for gaze, while the response itself then terminates that gaze, which is similar to Kendon's (1967) proposition for this interactive turn-taking form of gaze. In their research Bavelas et al. (2002) measure interactions through listener responses and the onset and offset of gaze within conversations. The results indicate that gaze does not only regulate turn-taking but also allows the social actors to seek and provide listener feedback. They also suggest that even with a particular role division the collaborative nature of gaze within speech is still active and requires no role switch. Furthermore, they assert that gaze is, in most occasions, used functionally in the sense that it is non-redundant. The shortcomings of the study lie in the measurement of gaze, which did not account for the randomness of gaze by taking it together with other communicative acts. The functional overview of gaze in a collaborative setting creates insight into how collaborative gaze is utilised within a task-based setting. The present task also requires social actors to redirect their gaze towards multiple social actors to allow for interaction and turn-taking while simultaneously requiring gaze for the observations on the structure.

Norris and Pirini (2016) have done a Multimodal Interaction Analysis of two social actors communicating via video-conferencing to analyse how the participants

shift attention and convey their knowledge within disagreements. They used two students to analyse their interactions within a set of tasks. Within this research, there are some notions which are especially relevant within the context of the research. First, it is the notion of interaction and knowledge is based on notions such as the assertion that all forms of human organisation are communicative by nature (Kastberg, 2007). This, of course, resonates with the beliefs of Multimodal Interaction Analysis (Norris 2004). The second is that knowledge can sometimes be recognised by social actors, which can then be transmitted through language (Heritage, 2012). The notion of knowledge transmission solely through language is a bit limited in the context of Multimodal Interaction Analysis. It nevertheless provides insightful concepts of knowledge negotiation. When the participants were asked how they experienced the tasks, they mentioned that it looked similar to real-life context such as collaborative school assignments. This experience indicates that the concepts of transmitting and negotiating knowledge can be considered in a universal setting. Norris and Pirini (2016) assert that the modal production has some variation in the context of disagreements even though the resolution has a consistent pattern. This study not only demonstrates the relevance of a multimodal approach to collaboration, but it also demonstrates that there might be patterns and more universal implications for the multimodal research of collaborative endeavours. These patterns of resolution can provide tools to locate the resolutions of the problem sequence.

Norris (2006) has also done a Multimodal Interaction Analysis of multiparty interaction within an office. Here she investigated the way in which a social actor interacts with multiple surroundings. She concluded that the Multimodal Interaction Analysis revealed a more intricate way of action construction. Based on language alone it would seem that the social actor within his interactions switches between different

interactions through dyadic conversations. The Multimodal Interaction Analysis indicated, however, that within the context of multiparty interaction multiple higher-level interactions are co-constructed through multiple channels. Here Norris argues that all interactions are co-constructed, although there can be differences in how the social actors are 'linked' in each other's mind. This co-construction can be analysed through the modal-density social actors employ and should be individually analysed for every social actor. Within the context of the goal-oriented task, this emphasises the importance of co-construction in interactions, which is imperative to contextualise the collaborative effort of the social actors. This research supports the notion that the linked higher-level actions within the multiparty collaborative setting are also relevant for the analysis of the individual lower-level actions that constitute the higher-level actions social actors undertake in the present research.

It is thus apparent that social interdependence theory and promotive interaction have a positive influence on collaborative tasks. This positive influence will provide context for the group dynamics of the participants. The shared goal allows the social actors to overcome obstacles together but does not yet provide an answer as to how they interact.

2.3 Spatial Elements and Attention

Spatial elements have received substantial academic substantial attention within the context of communication and goal-oriented settings (Louwerse & Bangerter, 2005; Beum & Cremers, 1998; Kraut et al., 2002; Chu & Kita). Research has shown that spatial elements have an important function within goal-based interaction. Beun and Cremers (1998) suggest that when working within a shared visual space, the social actors try to make their utterances and referents as short as possible by virtue of the

principle of minimal cooperative effort. This principle is based on the Gricean maxim of quantity to make your contribution as informative within your goal and to simultaneously not make it more informative then required (Grice, 1975). The minimal cooperation principle was elaborated on by Clark and Wilkes-Gibbs (cited in Beun & Cremers, 1998) by stating that references to objects are a collaborative process. The speaker initially has the option to create ambiguity and hope the other social actor can make an educated guess or ask for clarification.

Beun and Cremers (1998) suggest in their research that for the focus of attention, deixis is important. Within their experiment they researched how a social actor collaboratively builds a replica of an example structure when it is only visible to the instructor, which is the other social actor. The building consisted of multiple colours and shapes to create divergent options which are still simple and non-figurative. The specific focus of the experiment was on the referential acts of the social actors within the experiment. The referents hardly displayed cases of ambiguity or redundancy. The social actors tended towards functional information to resolve ambiguity and a focus on absolute features when describing objects. Their emphasis, however, lies mainly on focus within the context of dialogue, at least in how it was measured and analysed, which can also be broadened to interaction to remove the focus on the verbal elements. The referential analysis with replica building suits the nature of the building task very well due to its overlapping elements. Although there are some differences in the visual and interactional conditions, the gist of replica building and referents with colours and form create a nice frame of reference for the present task.

An example of how focus can be facilitated is the joint-attention hypothesis, which was argued for as a result of an eye-tracking experiment (Louwerse & Bangerter, 2005). Their results suggest that joint attention facilitates reference resolution and that

this joint attention is facilitated or can come to fruition by use of deictic gestures. Most signs to signify spatial elements happen through deixis. The word 'deixis' originates from the Greek word *deiktikos* and it means 'pointing' or 'indicative', which adequately describes its function within the spatial domain. The joint-attention hypothesis stipulates that pointing helps a joint focus of attention indirectly through redirection of gaze and directly through the cognitive processing that joint-attention facilitates. Bangerter (2004) has in earlier research studied the joint-attention hypothesis and argued that pointing can be used for more than the sole purpose of identifying referents, arguing for a more flexible view on deictic gestures. The ambiguity that follows from interaction in cohort with the minimal cooperative effort principle, can, in some cases be resolved by choosing the content that distinguishes the target object from the surrounding ones most effectively and this can be done through joint attention initiated by deictic gestures (Beum & Cremers 1998).

Kraut, Gergle, and Fussell (2002) also investigated collaborative goal-based interactions within a shared visual space. Their findings on the benefits of shared visual space are insightful and account for the effect of shared visual space in the context of its complexity and ambiguity when interacting. When a visual task is visually more complex, language becomes less adequate and shared visual space becomes more salient for the solution of problems. Kraut et al (2002) also seem to suggest that grounding, which is the finding of a common ground, is facilitated when there is a shared visual domain. The results suggest that the complexity of the task and the extent to which temporal accuracy is required positively influence the effectivity of shared visual space. When the visual space is compromised, however, the participants in their experiment seemed to adapt their language in order to create a new common ground. The general trend was that shared visual space had a positive effect on understanding

within the collaborative working group. The information grounding is particularly salient because the realisation of information grounding indicates that the ambiguity is resolved. The ambiguity that follows from the task thus requires information grounding, which can be seen as incremental for the process of problem resolutions.

When talking about spatial information co-speech gestures are more likely to be produced than when they talk about nonspatial information (Alibali, Heath, & Myers, 2001). Beyond that, Chu and Kita (2011) suggest that co-speech gestures are often spontaneously produced when they have to provide verbal descriptions of the ways in which the participant would solve spatial problems. They investigated if co-thought gestures enhanced performance in spatial task and if this benefit would last throughout multiple tasks. They also investigated if the beneficial effect would be problem-specific or problem-general. Their findings indicate that gesture-encouraged social actors produced almost seven times as many gestures and did significantly better than the non-encouraged group. The second block discussed the longevity of the earlier found effects of gesture on visuospatial performance. They found an effect on spatial memory but not on working memory in the later blocks. In the third experiment they researched if effects of gesture are problem-specific or problem-general. The gesture-encouraged group had a better accuracy and performed better than the gesture-allowed group in a similar spatial transformation task but did not do benefit from gesture in a visuospatial task that was not similar. The gesture production is also increased when there is difficulty in verbally describing spatial visualisation and when mental transformation is described (Chu & Kita, 2011).

The common trend is that visual elements elicit more forms of non-verbal communication to create meaning, which usually occurs in cohort with speech. This is interesting because the deictic gestures that the social actors present in visual context

are in multiple occasions in itself ambiguous. Within the building task the spatial information that is then presented is made unambiguous through the interplay of percept, utterance and gesture, which through their interplay lower the amount of specific meanings available to decrease ambiguity (Roth & Lawless, 2002). The use of gesture within temporal visuospatial tasks indicates that the analysis of interaction within such a task should implement a multimodal approach that incorporates and identifies such actions.

The positive effects of shared visual space on collaboration and the prevalence of deixis in the utilisation of this shared visual space have provided notions that are fundamental for the focus of the multimodal analysis. The prevalent role of deixis to create joint attention in shared visual spaces is shown to be pivotal in the resolution of problems. If these deictic gestures are self-sufficient for the redirection of attention, which often is the first step in problem resolution, is still unclear.

3. Empirical Methodology

This chapter will provide an overview of the circumstances surrounding the task's design. The goal of the task will first be explained within the context of social interdependence. This will be followed by the design which includes participants, materials and preparation, and finally, the procedure of the task will be discussed.

3.1 Goal of the Task

The task was designed to elicit situated multiparty communication where the communicative demands will be primarily spatial, visual and information based. In figure 1 [below] there is a visual representation of the task's set-up.
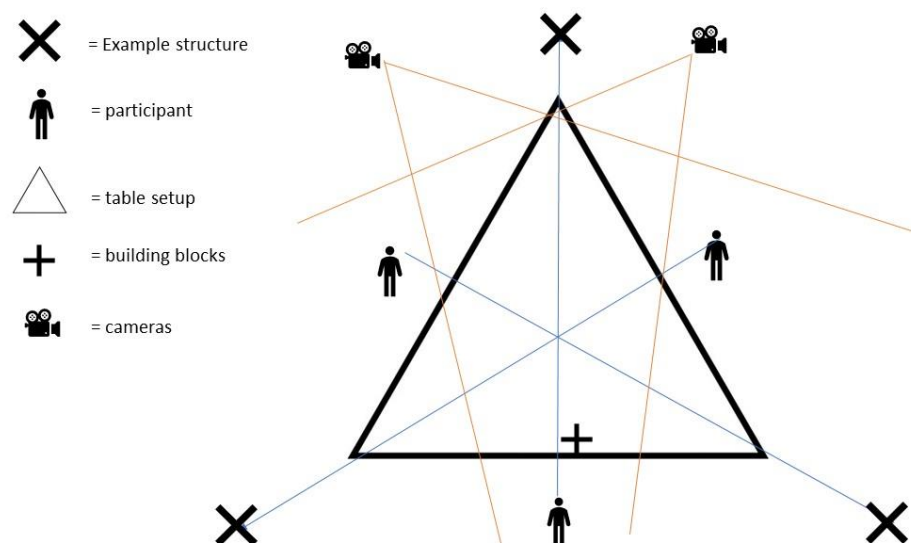

Figure 1.

The main goal is to generate social interaction within a naturalistically situated setting. This has been facilitated through the complex Duplo structure that had to be replicated. The Duplo duplication task has as an advantage that people enjoy doing this facilitates

interaction by creating a mutual interest and emotional investment (Johnson et al., 2007). When a group works towards one shared goal in a positive way, it creates promotive interaction through positive interdependence.

The exact focus had been determined after the raw data were collected. The assistant tools were intended to stimulate/necessitate interaction. The assistant tools were also instrumental in creating a form of social interdependence within the task. The visually-complex nature of the task made the interaction between the advantaged assistant and the builder a tool for dynamic, positive multiparty interaction towards the completion of the task (Van der Vegt & Van de Vliert, 2002). The triangular setting also had the added function that everyone could take the same vantage point through the structures. Within this setting, the participants were exposed to goal-directed multiparty communicative task where communicative demands are spatial, visual and information based.

3.2 Design

Data has been collected from 18 participants. All participants were self-proclaimed above C1 level on the CEFR scale, some native speakers but mostly L2 speakers. The recruitment occurred through snowballing as demographic particulars were not consequential to task design. Everyone was screened for colour blindness before the task, which in cohort with a basic understanding of English were the only exclusion criteria. All the participants gave their consent before filming them for this task. The data was videotaped with one or two video cameras per group and all the only materials required aside from the cameras was four bags of Duplo to allow for the construction of 4 identical structures. The benches with the three example structures were put at approximately four meters distance from the one directly aligned with that bench to

allow for visibility while still mainly revealing the front side. The participants' tables were located in a way that draws extra attention outside of the experimental elements. The exact amount of Duplo required to create a building identical to the example structures was placed on one of the desks. The cameras were placed in a manner that showcased the participants fully to partially from two different perspectives.

The participants were given an instruction sheet which explained the task procedure[1]. They were then allowed to read and discuss the contents of the instruction sheet. The participants could then decide on who becomes the builder and who the assistants. Both assistants were allowed to choose one cheat per person[2]. This could also, before they started the building process, be debated as a group. The participants all had to remain seated for the duration of the task except when physically utilising one of the assistant tools. After the assistant tools had been chosen the participants could begin the task. The assignment of the task was building a replica of the example structures. The participants had the opportunity to ask the proctor questions at any moment during and before the task. As soon as the building started the participants were left to their own devices unless they had any questions.

---

[1] See appendix, figure 5.
[2] See appendix, figure 6

4. Analytical Methodology

The previous section has accounted for the design of the task. This section will provide an overview of the considerations that underlie the analytical method of analysis and the boundaries in which this form of analysis will be set. The following section will display that Multimodal Interaction Analysis is most suitable for the task, which is followed by an extensive overview of Multimodal (Inter)action Analysis with a focus on the mediated action, which will be used as the unit of analysis, followed by the concepts of modal density and the foreground-background continuum. After this clarification, the protocol for data processing will be expanded upon. After the analytical method is clear, problem sequences as the focus of the Multimodal Interaction Analysis will be clarified. Finally, the foci within these problem sequences will be elucidated and the section will end with the limitations that follow from this analysis.

4.1 Multimodal (Inter)action Analysis

The analytical method will be based on the Multimodal (Inter)action Analysis championed by Sigrid Norris (2004, 2011). Multimodality is a theoretical framework that takes social semiotic ideas of communication and employs these notions of meaning-making signs to champion that all (inter)action is multimodal. Norris (2004) has created a form of analysis that allows for variability and the composition of modes which are usable in an analysis (Geenen, 2013). Norris (2013) defines a mode as "a system of mediated action that comes about through concrete lower-level actions that social actors take in the world" (p. 155). The changeability in the properties of these

analytical units is what makes Multimodal (Inter)action Analysis a fitting approach to capture the complexity that comes with everyday interaction.

The changeability of these analytical units marks the notion that multimodality is a certain orientation of focus, which can be applied across many fields and be given many different foci marking it as a trans- and interdisciplinary empirical means of study towards the use of modes and meaning-making in communication. The object of focus in this study is how people interact in a real-time setting and to elucidate the way in which everyday interactions provide meaning to our lives. This experience is subjective, which should be reflected within the framework and analysis. The possibilities this frame gives are well suited to the task due to the social and interactive nature of the task.

4.1.1 Mediated Action

Multimodal Mediated Theory is grounded in the Vygotskian notion that mediation is vital in understanding social actions, and that all actions, are mediated by mediational means or cultural tools (Vygotsky, 1978; Norris 2004, 2011). Norris and Jones (2005) state that "[a] mediated action focuses on two elements: the agent and the mediational means, emphasizing an inherent irreducible tension between the two" (p. 17), which denotes that every action is taken by an agent and, therefore, is also mediated through their irreducible tension.

This unit of analysis was first championed by Wertsch (1991). Wertsch wanted to provide a "coherent account of the human mind" by means of a unit of analysis (p.1). The aim of this unit was to provide a coherent unit that would exemplify the tension between the social actor and mediational means by introducing the mediated action.

The mediated action as unit of analysis keeps alive the complex processes in which cognition, action and communication are ingrained (Geenen, 2013). The complexity and relatedness of the social actor and the mediational means lie in that they cannot ever be treated in isolation, primarily because mediational means can only ever exist when it is employed by the social actor. All actions, however, are mediated and the social actor must therefore be viewed in cohort with the mediational means (Norris 2004). The aim of the mediated action is to create a unit that can change when actions require it to change (Norris, 2004).

Scollon (2001) identified an ambiguity of scope and concreteness within the concept of the mediated action. Scollon describes it as "[a] mediated action is carried out through material objects in the world including the materiality of the social actors – their bodies, dress, movements in dialectical interaction with structures of the habitus", which is one of the central concepts of the Mediated Discourse Analysis (p. 4). This links the mediated action to all the complexities surrounding the interaction by utilising the mediational means as a binding element. These actions and means are set within what Scollon calls the 'site of engagement'. The site of engagement in combination with what Scollon calls the practice and the Nexus of practice is what sets the focus on real-life interaction using microstructural analysis to create a view of the macrostructure which Scollon calls the nexus of practice. This notion allows the analyst to look at social actions and the role of discourse within real-time actions to apply and look at multiple forms of interaction, which is a pivotal notion for Multimodal (Inter)action Analysis. The combination of these principles indicates that Scollon moves away from the idea that discourse is a system of representation, thoughts and values (2001). Mediated Discourse Analysis is instead best conceived as a matter of social actions.

The notions embedded in Mediated Discourse Analysis form a big and integral part of the frame that makes up Multimodal (Inter)action Analysis. The mutability that Multimodal (Inter)action Analysis brings is in the definition of the communicative mode as a unit of analysis. Where a traditional multimodal semiotic system within functional linguistics would be incompatible with Multimodal Interaction Analysis (Geenen, 2013), Norris (2004) resolves this incompatibility by employing the communicative mode as a heuristic unit. Norris explicates the term heuristic unit as a unit that "highlights the plainly explanatory function, and also accentuates the constant tension and contradiction between the system of representation and the real-time interaction among social actors" (p. 12). The fact that the mediated action is defined as a heuristic unit underlines the possibilities in the analysis and simultaneously foci of Multimodal Interaction Analysis. This holistic view where the modes by themselves provide a unit to analyse the complexities that accompany communication, action and interaction in our everyday lives.

4.1.2 Lower-level and Higher-level Actions

The plethora of possibilities that the mediated action or communicative mode as the unit of analysis give is what Norris defines as lower-level actions within the analytical frame (2004). These lower-level actions are conceptualised by the interactional unit with the smallest pragmatic meaning. The lower-level actions within the scope of this thesis looks at the smallest pragmatic units for gesture, proxemics, gaze, posture and verbal utterances. For gesture this is the stroke, for proxemics this is a change in

proximity towards other relevant elements, for gaze it is a shift in gaze, for posture it is a shift in posture and for verbal utterances it is an utterance[3].

These units are suitable for analysis because they function as heuristic units. This allows for the researcher to capture the individual lower-level actions to utilise them as explanatory units, which can then be utilised to put the pieces together. lower-level actions do not carry inherent communicative value because they are the smallest pragmatic unit of meaning. Therefore, it is not the lower-level actions in isolation that are the focus of Multimodal Interaction Analysis, but rather the interplay of the lower-level action to construct meaning. Norris (2004) argues that "[i]ndividuals in interaction draw on systems of representation while at the same time constructing, adapting, and changing those systems through their actions. In turn, all actions that individuals perform are mediated by the systems of representation that they draw on" (p. 12). The notion that social actors are constantly changing, adopting and constructing, display the saliency of the lower-level action as heuristic units. In the present task, the focus is on capturing the complex multiparty interaction and analyse it through a manageable unit of analysis that allows us to investigate the coherent whole through the smallest systems of representation.

Meaning is constituted through the means with which social actors interrelate through a chain of simultaneous and interwoven interaction. The assimilation of the multitude of communicative modes is what constitutes a higher-level action. The shift of gaze, posture and the verbal utterance *could you give me food please* would constitute the higher-level action of asking for food. The higher-level and lower-level actions thus occur simultaneously. These higher-level actions are "bracketed by an

---

[3] Chafe (1994) points out that language is naturally segmented through intonation units, which is physiologically conditioned by the need to breathe.

opening/closing", which means that there are lower-level actions within the higher-level actions which indicate a shift within the higher-level actions (p. 13). Higher-level action can also be embedded in other higher-level actions, which allows for higher-level actions to account for the multi-layered complexity that interaction can represent (Norris 2004). If, for example, the asking of food is happening while reading a book, both higher-level actions would then simultaneously occur with their own lower-level action to serve as opening and closing brackets.

Norris argues that there are certain triggers which can function as an opening or the closing of these brackets. The shift in attention can be analysed through what she calls the semantic/pragmatic means. This is the way communicative devices are used to communicate the occurrence of a shift of higher-level action.

4.1.3 Modal Density & Foreground-background Continuum

Social actors have the possibility to focus on multiple things at once. The allocation of attention and awareness is, however, not equally divided. The foreground-background continuum accounts for this attentional inequality which occurs during multiple embedded higher-level actions. Originally, the concept of three levels of awareness was used in the music and art sector (Norris 2011). The notion of three hierarchical levels of awareness was eventually adopted by chafer and van Leeuwen (cited in Norris, 2004). This notion of hierarchy and levels of awareness was then adapted by Norris (2004, 2011) through the perception of utilising the hierarchy as a relational heuristic unit.

The manner of measurement for the foreground-background continuum lies within what Norris (2004, 2011) calls modal density. The higher the modal density, the more foregrounded a social actor's attention generally is within the context of that real-

time interaction. High modal density can come about in three different ways. High modal density can come about through the intensity of a mode, through the modal complexity of multiple modes that were intertwined and through the combination of intensity and complexity of multiple modes can account for a high modal density. The intensity and complexity of modes are relational and can in that sense not be quantified (Norris, 2011). This suits the interactional elements of the task well due to its complex interactional nature. The modal density and foreground-background together give us insight into what a social actor is focussing on and how and when there might be a shift in this focus or awareness.

4.1.4 Transcription Materials and Procedure

Norris outlines a detailed method to transcribe a Multimodal (Inter)action Analysis with a focus on imagery to create a holistic and accurate representation (2011). The transcriptions were originally created via atlasTI. Here the video sequences were indexed, segmented and transcribed through use of snapshots. These snapshots were later converted to 'Jpegs', which were utilised to create a multimodal transcript with timestamps, indication of the relevant modes and the utterances added in Wordart to create an overview which accounts for the complexities by being able to articulate the intricacies of modes in isolation and then compiling them to provide a context about the higher-level actions that are performed by the social actors. The participants were given pseudonyms for the analysis. The relevant transcripts are added to my appendix and utilised throughout the final analysis.

Figure 2 [below] is an example of a transcript, which will be used to illustrate the transcription protocol. The transcript's legend is designed as follows: The thick red arrows signify a postural shift, the thick blue arrows signify a head turn, the thin blue

arrows indicate gaze direction and the circles indicate a form of gesture or object handling. In frame 1, the double-pointed arrow indicates reciprocated gaze. The red arrow indicates a postural shift, which in this case indicates that the social actors bends forward. The textboxes are all colour-coded per person. For all the transcripts the most left social actor will be red, the one in the middle yellow and the most right social actor will have blue text. The thick blue arrows in frame 1 indicate head movement to the right. The circle around the most right-based social actor's hand signifies the deictic gesture.
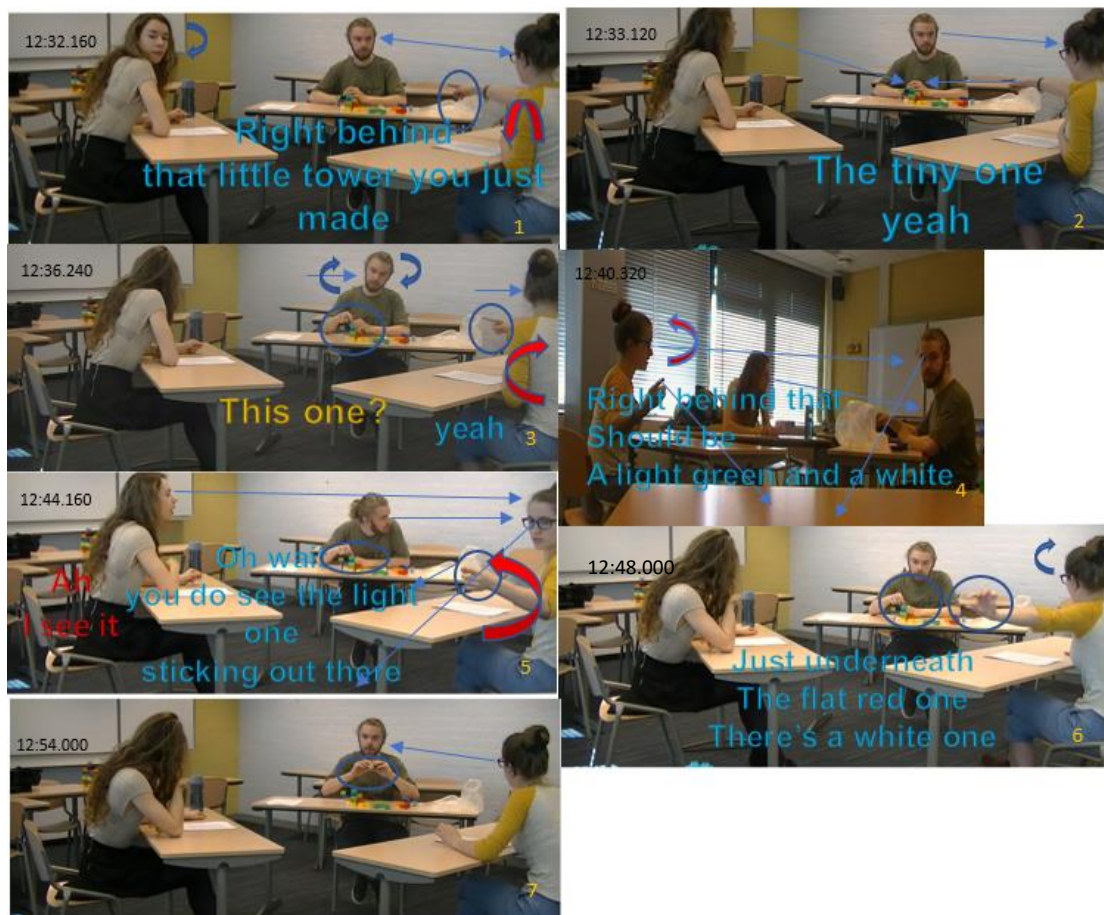

Figure 2.

4.2 The Problem Sequence

The focus of the Multimodal (Inter)action Analysis is on what will be called the problem sequence. Wehmeier defines problem solving as "the action of finding a way to deal with a problem" in the Oxford Advanced Learner's Dictionary (cited in Khoo, 2015). A problem sequence in this thesis is defined as a moment within the collaborative process where either the builder or the assistant identifies a problem of any nature at that site of engagement. Within this context, problem sequences are initiated by means of utterances, deixis or the use of head-movement. Problem sequences are also interesting as they require a means to put it to attention and usually go paired with problematic or ambiguous situations as there is a mismatch in the knowledge between the social actors. The problem sequence will be analysed within the context of the resolution of ambiguity, grounding of information and the construction of meaning as a collaborative endeavour.

The higher-level action of the problem sequence actions required for solving the problems has become the main focus of the Multimodal (Inter)action Analysis. The ambiguity that a mismatch of knowledge, perspective and attention potentially generates, provides a site of engagement well-suited to the holistic and complex possibilities that Multimodal (Inter)action Theory covers.

4.3 Limitations

The shortcomings of this analytical method are that even though different perspectives can be taken, there is still lack of inclusivity of all modes as a holistic method of analysis. Norris indicates that the use of a camera for imagery in itself already provides difficulty, which captures the difficulty to include all modes (2011). This already

determines the limited perspective that can be recorded. The specific view of what occurs during the interaction only gives a limited perspective, which also guides toward a specific viewpoint. The subjective nature of qualitative analysis and the setup of the task. Due to the density of a multimodal analysis there is also the issue that only a limited set of data can be analysed within a reasonable amount of time. The other limitation is the speculative nature of qualitative analysis. Theories can be applied to the Multimodal Interaction Analysis as a framework, yet the perception, reception and attention of the social actors only show the tip of the iceberg. This limitation is, however, omnipresent in all forms of research trying to elucidate the vast complexities of human cognition. The analysis of seconds can in discourse analyses, and especially in Multimodal (Inter)action Analysis, take hours a "full analysis of a short passage might take months and fill hundreds of pages" (Van Dijk, 2001, p. 99). The specific results of the task at hand can only, due to the culturally-diverse nature of meaning-making devices, be related to our own western socio-cultural environment.

5. Analysis

The analysis of interactions within the problem sequences uncovered several salient lower-level actions which are instrumental in our understanding of focus distribution and the resolution of ambiguity within problem sequences as a higher-level action. Within this section, deixis will be used to denote actions which have deictic elements in them due to the high complexity that some modes and, in particular, gestures bring concerning their classification. This section will provide three transcripts that are each expanded upon to illustrate a different point. In the first sub-section, the analysis will have a focus on deixis as a tool for reorientation. Afterwards, the seconds transcript's focus will be on the builder's use of gaze and its perception by the other social actors. Finally, there will be an analysis of the co-construction of the higher-level action with a specific emphasis on the what social actor plays a role in its co-construction[4].

5.1 Deixis as Means

Figure 2 exemplifies an instance where the location of a particular green block prompts the unfolding of a problem sequence. Mary wants to indicate the position after thinking about a potential resolution. In this problem sequence, the attention within the collaborative group is redirected by deictic gestures in interplay with additional modes such as posture, gaze, proxemics and spoken language. Due to their intensity, the deictic gestures are particularly salient for the redirection of social actors' attention. Its longevity is particularly salient because their materiality as a spatiotemporal element has a longer duration than minimally necessary for the sole function of attentional redirection. This will be called lower-level action residue to indicate that its

---

[4] For additional transcripts to support the findings, please refer to the appendix.

communicative function has passed, which could indicate that it could fulfil different functions such as the internal structuring of complex visual information, but it could also suggest some unresolved element of the lower-level action.

Within this problem sequence, the minimal cooperative effort principle has been adhered to by initially utilising minimal amounts of non-redundant information through their utterance. On every occasion, Mary's initial lower-level actions contain a minimal amount of non-redundant spoken language in cohort with the deictic gesture. Mary adds extra information only when, after a small pause, no confirmation of the information grounding takes place. The co-construction of information grounding as a higher-level action has a back-and-forth element of meaning-making signs to see if the social actors are on the same page, which is paramount to both giving and receiving instructions. The privileged assistant first confers information to the builder, to which the builder confirms that the given information is perceived and comprehended.
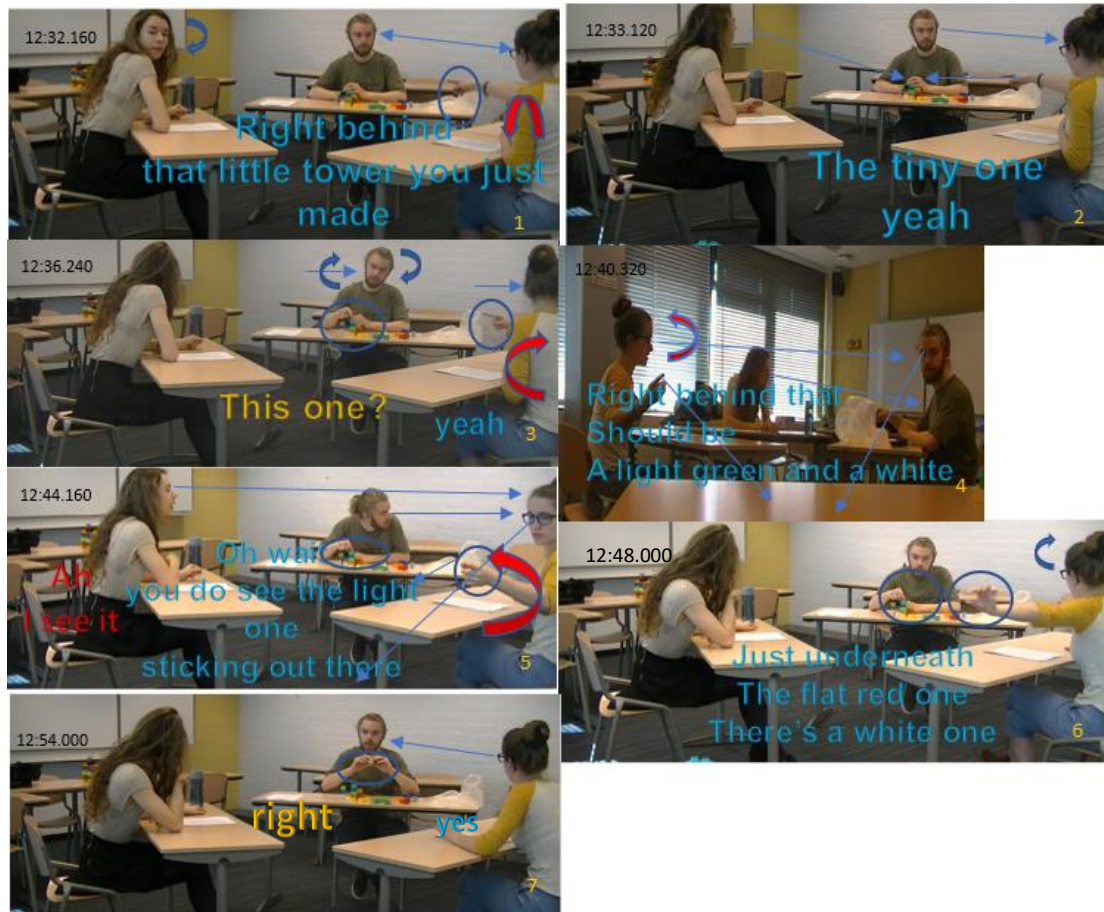
Figure 2.[5]

In frame 1, Mary undertakes multiple lower-level actions through the modes of gesture, proxemics, posture, spoken language and gaze. At 12:32:160 Mary says *Right behind that little tower you just made,* which is co-accompanied with a deictic gesture towards the main structure while simultaneously gazing and posturing towards John. John reorients his gaze and shifts his posture towards Mary. The modal complexity of multiple lower-level actions undertaken through posture, gaze and proxemics suggest that John is paying attention to Mary as a result of the earlier deictic gesture. In frame 2, Aria shifts her posture, bends forward and reorients her gaze towards John, which indicates through her lower-level actions that the higher-level action of examining the

---

[5] The assistant to the right is named left is named Aria and the assistant to the right is Mary. The builder is named John.

situation is being undertaken by Aria. At 12:33:120 Mary says *The tiny one yeah* while bending forward and shifting her gaze to the main structure. So, the co-construction of information grounding with John is now in the foreground of her awareness and her higher-level action of drawing his attention is resolved. Aria's reorientation in attention and Mary's actions match with the notion that deixis redirects the attention so that the social actors can co-construct the higher-level action of achieving a common ground within their shared visual domain to resolve the mismatch created by the ambiguity of tempo-spatial elements (Kraut et al., 2002; Beun & Cremers, 1998).

In frame 3, John undertakes multiple lower-level actions through gaze, posture, proxemics, gesture and spoken language. At 12:36.240 John asks *This one* and makes a deictic gesture towards the main structure[6]. Here the deictic gesture is mainly used for specificity as the attention was already on John moments before. In frame 3, John's actions together constitute the higher-level action of confirming the resolution of ambiguity, which expresses that he listened to Mary's instructions to together co-construct the higher-level action of the problem sequence. Mary's arm moves slightly after the postural shift, but her hand remains in the air and does not lose its deictic qualities[7], but does, however, not point towards the main structure anymore, nor is it still functional for John. So, the specificity and redirection of attention of the deictic gesture are lost, but the form remains. This lingering materiality suggests some form of lower-level action residue, which can have several implications. There could be some communicative function that is not yet completed or uncertainty of the completion which allows the residue.

---

[6] The arrows near John in frame 3 indicate that he, as indicated by his curved head form, was moving his head on that timeframe, which is because his gaze was on the main structure milliseconds earlier.
[7] In frame 3, Mary's hand is still in the air in a similar deictic form but changes direction, seemingly pointing at nothing relevant.

In frame 4, Mary undertakes multiple lower-level actions through the modes of gaze, gesture, spoken language, proximity and posture. At 12:40.320 she says *Right behind that should be a light green and white.* Her spoken language co-occurs with a postural shift and redirection of gaze towards John while her deictic gesture points towards the example structure to her right, lowering the distance between her finger and the example structure. The actions employed through the deictic gesture here is salient in the sense that it also displays qualities of lower-level action residue through its materiality after the redirection of attention is completed. John's redirection of gaze and postural shift are towards Mary's signified referent-object, indicating attentional allocation towards the example structure. John returns his hands towards their original position on the main structure, indicating that his building is still in the background due to the proximity of John's hands.

In frame 5, Mary undertakes multiple lower-level actions through the modes of posture, gesture, gaze, proxemics and spoken language. Mary's posture, gaze and deictic gesture reorient towards the example structure to her right to point out the direction and nature of her new information in co-occurrence with *Oh wait you do see the light one sticking out there* as spoken language to re-establish the grounding of information within the collaboration. Here Mary's role as privileged assistant enables her to take a leading role in the interaction, which results in her redirecting the attention with her additional knowledge through the use of deictic gesture. Up until now, Mary has only given the minimal amount of information required. In frame 5, however, she adds extra information because the information grounding did not work out with the absolute minimum non-redundant information. In frame 5, John's posture and gaze are reoriented to the example structure Mary referred to in frame 4. Aria says A*h I see it and* reorients her gaze, shifts her proxemics, and posture towards the example structure

ahead of her, which indicates through these lower-level actions that her attention shifts within the higher-level action of the problem sequence occur, albeit later than John's attention shift. The modal complexity of multiple lower-level actions undertaken through the modes of posture, gaze, proxemics and spoken language suggests that she now foregrounds and actively co-constructs John and Mary's collaborative higher-level action of information grounding.
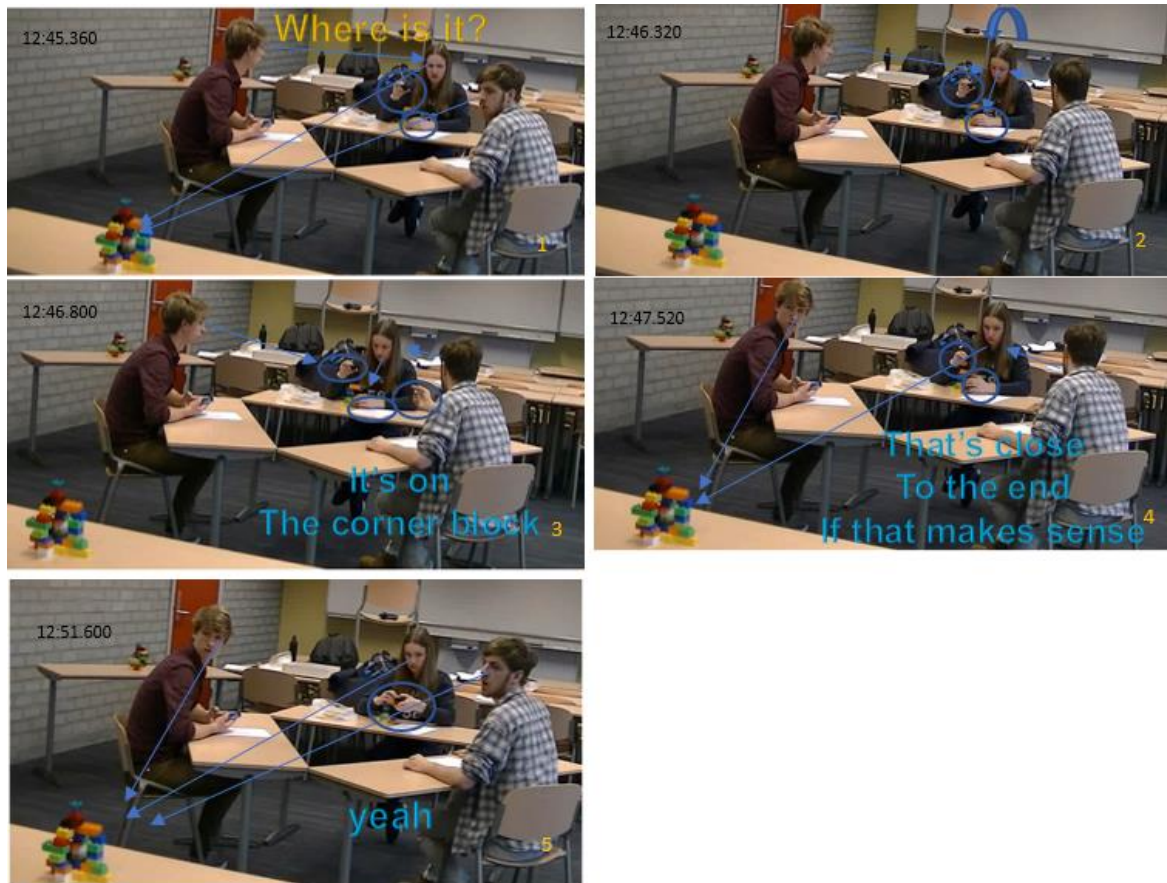
In frame 6, Mary undertakes multiple lower-level actions through the modes of gesture, gaze, proxemics, spoken language and posture. She reorients her gaze towards John and bends forward after a postural shift to be in closer proximity of John. At 12:48.000 Mary indicates through spoken language *Just underneath the red one there's a white one* that she can see a relevant object. This modally complex chain of lower-level actions acts as a temporal-spatial specification within the co-construction of the higher-level action of information grounding. Mary's deictic gesture towards the example structure to her left has a high (modal) intensity, which facilitates the redirection of attention. John's proxemics and postural shift closer to the example structure and the reorientation in gaze towards that structure indicate that John perceived Mary's higher-level action of redirecting attention. In frame 7, John undertakes multiple lower-level actions through spoken language, gaze, posture and proxemics. He straightens his back and changes the proxemics of the blocks by clicking them unto each other while gazing at the example structure straight ahead. At 12:54:000 John says *Right* to ask Mary's confirmation. A couple of milliseconds later, Mary says *yes* and confirms through spoken language, which happens in co-occurrence with multiple lower-level actions through the modes of gaze and posture. Mary's posture shifts towards John while simultaneously bending forward after which her gaze turns

towards John's hands. Mary's confirmation marks the resolution of the higher-level action of the problem sequence.

5.2 Gaze Priority

Figure 3 exemplifies a problem sequence in which Ned gives instructions regarding the position of a black block, after which Vera prompts a question. Ned consequently specifies his earlier enquiry regarding the proxemics of the black block. Within this sequence, the higher-level action of thinking and executing the building process takes scope for Vera over the western interactional social and cultural standards revolving gaze (Rossano, 2012). The use of gaze is salient here as social actors usually desire confirmation that they are being listened to. As observed in this case, however, the resolution of ambiguity is successful even without reciprocated gaze. Vera's gaze towards the example structure is particularly salient for two reasons. The first is the sanctionability of her gaze if used similarly in other situations. The second reason is that she does not reorient her gaze after Ned's deictic gesture. In frame 5, the collaborative effort of checking Vera's object handling also indicates that for the higher-level action of verifying Vera's execution it is natural to redirect the attention towards the example structure, as is indicated by Ned's gaze on the example building while uttering *yeah* as confirmation.

Figure 3.[8]

In frame 1, Vera undertakes multiple lower-level actions through the modes of gesture, proxemics, posture, spoken language and gaze. At 12:45:360 Vera says *Which is it,* which is co-accompanied with an iconic gesture while holding the black block in her hand. Vera shifts her posture to face the example structure, reorients her gaze towards it and changes the proximity of the black block concerning the main structure. The ambiguity of Vera's perception regarding the instructions, mark the opening of the problem sequence. Ned's posture shifts a bit to his left and his gaze reorients towards the same example structure as Vera. Here, Ned reorients to provide additional information to co-construct the higher-level action of information grounding with regard to Vera's perspective. In frame 2, Vera shifts her posture forward and bends a

---

[8] Assistant to the left is Bran, assistant to the right is Ned and the builder is Vera

little while simultaneously reorienting her gaze towards the main structure and reorients her position to the main structure. Ned's postural shift and gaze reorientation indicate that through these lower-level actions looking at Vera's construction process is now on Ned's foreground.

In frame, 3 Ned undertakes multiple lower-level actions through spoken language, gesture, gaze and posture. At 12:46.800 he says *It is on the corner block* giving a specification in co-occurrence with a deictic gesture, a postural shift forward and gaze orientation towards Vera. Through the undertaking of these lower-level actions, Ned expresses a desire to inform the other collaborators. Vera uses her gesture for the object handling of the black block which indicates that it is still within her field of attention. Her gaze orientation and posture stay directed towards the main structure. The gaze is especially salient here because it usually is sanctionable during interactions to not look at other social actors when spoken to in Western, and particularly European, culture. There are, however, cases in the sanctionability of mutual gaze is relatively flexible (Rossano, 2012). This indicates that within the goal-directed co-construction of the higher-level action of building the structure takes interaction priority over the social norm that the hearer should look at the speaker. Within this particular context, the social actors do not express any discontent about this particular phenomenon during the interactions, which indicates that the goal is perceived as more important than the socially and culturally coded norms of gaze which are normally adhered to.

In frame 4, Ned undertakes multiple lower-level actions through the modes of Gaze, gesture, posture and spoken language. He has a (slightly lower than earlier) deictic gesture and his posture changed as a consequence of a head-moving directional nudge. At 12:47.520 he says *That's close to the end if that makes sense* during which his gaze is still fixed on Vera. These actions constitute the higher-level action of

instructing Vera to co-construct the higher-level action of information grounding to resolve the problem sequence. The information grounding process seems to adhere to the minimal cooperation effort principle. In the same frame, Vera gazes towards the example structure after shifting her posture upwards and slightly bending forward, which changes her proxemics to the example structure and the main structure. Her iconic gesture on the black block remains to indicate her attention on the black block remains somewhere in her attention continuum. Vera is being spoken to for the second time during this interaction, yet she allocates her meaning-making signs back-and-forth between the example structure and the main structure.

In frame 5, Vera undertakes multiple lower-level actions through the modes of gaze, gesture, proxemics and posture. The postural shift is slightly backwards and co-occurs with a back-and-forth gaze reorientation between the example structure and her own hands while she is manipulating the proximity of the blocks object handling. Here the lower-level actions of Vera construct the higher-level action of putting the black block on the main structure, which indicates that she has understood Ned's earlier instructions. Ned shifts his posture from the main structure to the example structure while similarly reorienting his gaze in the same manner. He utters *Yeah* at 12:51.600 to indicate to the group that Vera connected the right blocks, which marks the end of the collaborative co-construction of the problem sequence.

5.3 Not Dyadic but Triadic

Figure 4 exemplifies a situation where there is ambiguity regarding the position of the blue block within the main structure. Mary instructs about the position while John reacts to solve the problem sequence. Within this problem sequence the way in which John and Mary's dyadically-oriented lower-level actions could be seen as the co-construction of the higher-level action of instructing John. This, however, would be an incomplete assessment. I argue that Aria, through her moderation also co-constructs meaning within the higher-level action of instructing John in her role as observer. John and Mary's actions are more modally intense and complex than Aria's, but that does not mean that Aria does not co-construct, it means that she merely requires or employs fewer modes than the other social actors within the context of her role within multiparty collaboration (Norris, 2006). The overlap in the lower-level actions undertaken in frame 6 exemplifies the way in which Mary and Aria take on similar positions to, at least to some extent, fulfil similar functions, which supports the idea that Aria has (actively) been co-constructing all along. The dyadic nature of the conversations seems to indicate that there is an understanding within the social actors that allows for the successful role distribution within the cooperative process, which does not require three social actors to converse at the same time. Talking at the same time would also negatively impact the amount of attention a social actor can distribute among the higher-level actions that require attention among their interlocutor.

Figure 4.

In frame 1, Mary undertakes multiple lower-level actions through the modes of gaze, spoken language and posture. At 13.09.600 Mary says *You have the blue brick* while simultaneously orienting her gaze towards John while bending forward. These lower-level actions together construct the higher-level action of instructing. John's posture and gaze are directed towards Mary to indicate that he is aware of her actions. The proxemics and gesture construct the higher-level action of holding the blue block, which materialises John's awareness of the block he has to manipulate. John and Mary co-construct the higher-level action of information grounding through their collaboratively oriented lower-level actions. Aria's posture and gaze are directed towards the structure to her left. These actions construct the higher-level action of

looking at an example structure. Mary and John seemingly co-construct the higher-level action of information grounding in a dyad. Aria's attentional shift, however, merely indicates that she is *now* allocating attention towards an example structure, not that she is not still co-constructing within the higher-level action of the problem sequence.

In frame 2, John undertakes multiple lower-level actions through the modes of gaze, posture, proxemics and spoken language. At 13:10.560 John says *Yeah* together with a gaze reorientation to the main structure and a shift towards a more centred posture, while his hand remains on the white block in hands retaining its proxemics towards the structure. Their actions co-construct the higher-level action of responding to Mary indicating that John is aware of Mary's earlier utterance. The high modal density of John's higher-level action suggests that it is in the foreground of John's attention. John's proximity to the white block indicates that he still has some attention allocated towards the white block in the background of the continuum while Mary is in the onset of a gesture stroke, during which her posture and gaze remain oriented towards John. The latter lower-level actions construct the higher-level action of checking John's process while the former is an onset for the action of instructing John. Aria redirects her posture towards John while simultaneously reorienting her gaze towards John. After the postural shift, she bends forward and positions herself within closer proximity of John. These actions construct the higher-level action of observing John. So, Aria, John and Mary now together co-construct the higher-level action of the problem sequence. The posture and gaze are salient to indicate that Aria is actively co-constructing within the higher-level action of the problem sequence.

In frame 3 to 5, Mary undertakes several lower-level actions through the modes of gesture, gaze, posture, proxemics and spoken language. At 13:12.000 she says *The two ones closest to me* followed by T*hose should not be covered* at 13:14.400. In frame

3, Mary's gaze shifts from the main structure to the builder while employing a postural shift backwards. These actions constitute the higher-level action of information grounding with John and Aria. In frame 4 and 5, John reorients his gaze and shifts his posture towards the main structure after he expressed his awareness of Mary's instructions. His deictic gesture shifts towards Mary's two referent objects and alternates between the two blue blocks[9]. Here they co-construct the higher-level action of instructing John. In frame 3, Aria's posture is oriented a bit more to the front in co-occurrence with a steady gaze towards the main structure. After her gaze is set upon the main structure, she uses a gesture to take a pose generally associated with thinking, which from a sociocultural perspective is perceived as an active pose. Aria constructs the higher-level action of moderating the process while the lower-level actions employed by John, Mary and Aria together co-construct the higher-level action of the problem sequence and embedded within that the higher-level action of information grounding.

In frame 6 and 7, John undertakes several lower-level actions through the modes of gaze, proxemics, posture and spoken language. At 13:16:320 John says Y*eah yeah it goes here*, which signifies that he perceived and, in his perception, understood Mary's earlier instructions. The spoken language invites a form of confirmation which indicates that he is still aware of the other social actors while executing the instructions. In frame 6, Mary reorients her gaze towards the main construction which is accompanied by a postural shift slightly forward. At 13:16.320 Mary says *Yes* as a confirmation of the earlier provided utterance by John. In frame 6 and 7, Aria undertakes the same lower-level actions as in frame 3 to 5, which indicates that she is still in her role as an observer.

---

[9] This is only visible in the video because its near-impossible to capture that in frames with videos of this quality

Mary's employment of lower-level actions in frame 7 marks the end of the problem sequence. In frame 6, Mary and Aria undertake similar actions, which indicates that they to some extent assign similar attention to the higher-level action of observing John's implementation of the instructions. The lower-level actions of John, Mary and Aria together co-construct the higher-level action of the problem sequence and thus mark the end of the problem sequence.

6. Conclusion

The present study analysed the interactions within a complex naturalistic collaborative multiparty goal-oriented task to observe which additional modes were employed to resolve the ambiguity that followed from the task. This was done in a controlled setting within a shared visual domain. The controlled setting supported authentic communication within the confines of the task due to the intrinsic motivation that the task promotes. The participants chose their own role division and assistant tool to make the task execution possible. This division of roles within a collaborative goal-directed task facilitated forms of communication which allowed for positive interdependence.

A Multimodal Interaction Analysis of the problem sequences reveals that several salient additional modes were employed by the social actors as a means to resolve the ambiguity that followed from the mismatch in information, one of which is the deictic gesture. When a problem sequence transpires, the ambiguity that follows requires that the social actors know where to look when interacting to have a similar referred object within their visual space. The social actors' attention is regulated and reoriented by deictic gestures as a pragmatic means when the higher-level action of information grounding requires the intraparty members' attention to be directed towards an object or space within the shared visual space. This use of deictic gestures as a means of redirection is primarily used by assistants, contrary to the builders, who typically use it for specification purposes rather than redirection, as the attention, in most cases, is already on their process. Within the problem sequences, the deictic gesture also indicates a start of the co-construction of information grounding as a higher-level action. The interaction that follows is generally back-and-forth co-construction of meaning-making signs to see if the social actors are on the same page, which is paramount to both giving and receiving instructions. Another salient mode is

gaze during dyadic conversations within interactions. When assistants were assisting the builder, the builder mostly oriented their gaze towards their structure building process. Under normal circumstances, looking away when being spoken to indicates a diminished engagement in the conversation, which makes it a conflicting use of gaze when situated in the Western cultural and social norms (Rossano, 2012). The speakers, in this case, the assistants, did not perceive the builder's allocation of gaze as rude or undesirable within their interactions. This indicates that within this particular context, the goal takes interactional priority over the social and cultural norms involved, which means that during collaborative goal-oriented higher-level action of completing the task as a co-construction, a curious employment of gaze is utilised, which usually would be considered a sanctionable way of interaction. This seems to suggest that a priori maxims for interaction might indeed not exist, which does not entail that there may be situation-specific conventions. Clearly the purpose of the exchange, material setting, and physical organization have a consequential effect on the norms of interaction. This indicates a form of flexibility to adapt to the needs and purposes of the social actors. This was reflected in the builder's use of gaze to focus on the material consequences of the instructions rather than what the participant might convey through other modes than spoken language. The dyadic conversations as a phenomenon were also particularly salient in the sense that the Multimodal Interaction Analysis displayed how, within a collaboration, the higher-level action is almost always co-constructed in one way or another. The triadic group, in every case, co-constructed meaning towards the goal even if the mode of spoken language indicated otherwise. Within the higher-level action of the problem sequence, there was a role-division which allowed for a different allocation of semiotic resources to, albeit with differences between their modal densities, co-create the collaborative solution through the interplay of the three social actors' mediated

actions. This follows Norris' (2006) notion that all interaction is co-created and by extension that within a collaborative endeavour, then, all interaction towards the goal in the form of resolution of problems, is also co-created. The lower-level actions undertaken by Mary and Aria in frame 6 of figure 4 indicate through their similarity that they fulfil a similar role at that point. Indicating that when Mary's role as instructor was over, both social actors had taken on a moderating role.

This generates new insights into the way in which we create and allocate meaning within a multiparty goal-oriented task. Communication is always mediated by the cultural, social and historical influences of our society. Within the context of the goal-directed action, we have established that some semiotic resources are used similarly when compared to general interaction, in the form of the deictic gesture, while other forms, such as gaze, take on different forms than expected within the general social-normative interactions of (Western) society. The multiparty interactions show that there are different ways in which multiple social actors can together construct meaning in a way that goes beyond the mode of spoken language, which is best reflected through a real-time analysis of the interplay of modes between the social actors. Gaze and deictic gestures have elucidated more of the intricacies and possibilities that a Multimodal Interaction Analysis can provide us with. The Multimodal Interaction Analysis has indicated through use of the mediated action that all (inter)action is mediated and should, within the context of interaction, be viewed holistically. This was reflected through the way that deictic gestures redirected attention, while simultaneously allowing for us to see how the other social actors respond to this through their use of other lower-level actions such as posture, gaze and proxemics. This showcases that deixis is a material manifestation of the first step in the process of information grounding, which is the first step towards the resolution of ambiguity. The

allocation of modes within some problem sequences indicates that the use of gaze might not have a specific set of rules. The analysis provides insights in the builders' use of meaning-making signs while simultaneously allowing for us to see if the group perceives these signs. This suggests that the rules are dependent on matters such as the site of engagement, material setting and the physical manifestation of the interaction which all perform a role that *then* plays an integral role for the conventions regarding gaze. Finally, the Multimodal Interaction Analysis showcases a form of co-construction that through spoken language alone would have been nothing more than a simple dyadic conversation. This demonstrates how the collaborative role division might be reflected within their use of multiple modes to co-construct a collaborative interaction towards the resolution of ambiguity within their goal-directed building task. These findings showcase that, within a collaborative endeavour, social actor can display their own form of meaning construction which is more complex and three-dimensional than first meets the eye.

7. Discussion

This section is divided into three parts. The first part is a reflection on the process of the research and the limitations that followed from the process, which will be followed by speculative theorising about phenomena outside of the analytical frame and will finally present suggestions for further research.

The study started with a non-specific directed research question and a task to elucidate how social actors interact within a collaborative goal-oriented setting. As a result, the composition of this thesis was a result of an organic process[10]. As the raw data were collected, the question progressed from a general interest in additional modes within a task to a focus on the problem sequence and how communicative modes contributed to the resolution of these problems. This new focus gave new challenges, perspectives, but primarily extra time-consuming work within the limited time available to construct this research. The biggest limitation in that sense is the amount of time the initial analysis of this task required to provide interesting input, this in cohort with the earlier mentioned detailed nature of the analytical method. This means the given transcripts only provide a limited, yet representative, sample of the entire bulk of video data[11]. Another aspect is that well-informed adjustments to the task were not possible because the reviewed literature and analysis was based on the gathered data rather than vice-versa. This does not mean the current version of the task is better or worse, just that some changes that would have been interesting could not have been followed up on.

---

[10] The task was of a naturalistic nature and therefore nothing was decided pertaining to exact focus and segmentation of the data.

[11] I call it video data rather than raw data because it is segmented, just not transcribed due to problems pertaining to time.

The analysis was set in the context of social interdependence theory (Johnson et al., 2007). The task was particularly suitable for positive social interdependence due to its interactive and cooperative nature. One of the requirements for individual interdependence among the groups, was, however, not entirely fulfilled and could have influenced the way in which the group dynamic worked. Collaborative problem-solving requires that, rather than individual ability, the group is the decisive factor in the resolution of the problem. The builder had a specific goal and task which only he/she could fulfil. These assistant tools, however, did not all allow the assistants to add their own unique input to the process as not all of the tools were set up in a way that privileges one assistant over the rest of the group with a certain piece of unique insight or information. This is a possible explanation for the dyadic nature of the conversations within the problem sequences. These ideas are partially based on post-test conversations with the participants and speculations based on observation of the data. One possibility is, that the information grounding process only required two social actors to engage with each other, or it is possible that the assistant was either conscious or unconsciously aware of this individual's 'lack' within the group dynamic and therefore assumed a more passive role within the collaboration like moderation or observation. This means that it could also be that the weakness of having no 'individual' interdependency resting on you as an individual can influence the collaborative process. The second interesting concept was an afterthought based on the lower-level actions residue and its potential functions within the problem-solving interactions aside from the function of attention redirection. The joint attention hypothesis supports the use of the deictic gesture for the establishment of joint attention. Gesture, however, is also shown to have positive effects on the internal structure of thoughts and general spatial awareness within visually complex tasks (McNeill & Levy, 1993; Louwerse &

Bangerter, 2005). This could also be another potential explanation for the lower-level action residue, to either support the own cognitive processing of the visually complex object or to retain/support the structure of the utterance based on earlier made observations within the communicative process. The residue in itself could be a salient subject of study due to its hard to determine nature and its, at surface level, seemingly functionless materiality.

The analysis in section 5.2 when put in perspective of multiple theories, gives way to some potential explanations for the phenomena revolving gaze, based on earlier research, which could explain this phenomenon[12]. One of them is a result of the workload division between builder and assistant being specific and interdependent, which allows for positive social interdependence, and even more specific, promotive interaction within this collaborative goal-oriented task. The (implied) distribution of role-based tasks within the co-constructed higher-level action of the problem sequence could then through substitution and division of labour simply not require reciprocal gaze to function as a group because the task at hand is self-evident. The collaborative goal of the group in that sense, takes precedence then, over the social norms of intragroup interactions between the social actors. This pertains particularly to Ned's task of information sharing and monitoring as assistant and Vera's task of building, which allow for both social actors to focus on their own task within the collaborative goal through substitutability (Deutsch, 1949). Both actors can, through substitutability, then focus on their own task without allocating attention to redundant actions. A different explanation would be that the structural and spatiotemporal complexity of the task requires a certain intensity of focus, which is then reflected in the intensity of the gaze orientation towards the example and main structure. This in cohort with the

---

[12] This is based on figure 3.

positive effect of shared visual space on complex tasks execution and information grounding, which means that the amount of attention required for the act of listening might be lower through the visibility of referent objects Kraut et al., 2002). Both possibilities, however, are grounded in the idea that the collaborative goal-oriented co-construction of problem solving as a higher-level action is indicated as the most prioritised higher-level action during the collaborative building process A third possibility is that Vera interprets Ned's actions as a form of an extended telling sequence (Rossano, 2012). This would put the focus on the perception of Ned's interaction rather than on the difficulties pertaining to the task at hand. This also raises questions with regard to the minimal cooperation effort principle because these two concepts seem to clash. The former implies a perception of a long narrative from the hearer, while the latter assumes a wish for a short-as-possible narrative from both speaker and hearer.

For further research, there are multiple suggestions regarding the setup of the task. An interesting idea would then be to see what happens if the assistant tools within the task *are* suitable for individual privileging and how this would influence the group dynamic. Would the implicit promotion of social interdependence make collaboration more actively triadic or would there still be a similar role distribution? There were some groups who had among their assistants, to some extent, more individual interdependence incorporated into their chosen assistant tools. They used a picture and a video of the structure. This specific combination of assistant tools, however, had its own problems because the video cheat was so much better than the other. Therefore, it could be argued that the video cheat overshadowed the other assistant's individual cheat with regard to its influence or possibilities within the social dynamic of the positive collaborative interaction. Another interesting aspect of the task could be to a focus at

social dominance within intraparty interaction starting from the initial role division in the procedure. The task could also be put in a different context, so rather than a visuospatial task, there could be a task based on the auditory channels or on the solution of a non-spatial complex problem. Social actors tended to look at the same structure when discussing problems. So, the blocks could be screened off in a way that allows every social actor to only see their own structure or even a screen which blocks only one direction. This would increase the spatial complexity and at the same time gives individual privilege among the collaborators through to three unique dyadic perspectives. The final suggestion for further research would be to see what happens if the collaborative problem solving is based on a task that requires a different intellectual domain to solve the task or even a competitive, rather than a collaborative task, which could elucidate how this situational context influences the use of actions within the multiparty environment.

8. References

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language, 44*, 169-188.

Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science, 15*(6), 415-419.

Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication, 52*(3), 566-580.

Beun, R. J., & Cremers, A. H. M. (1998). Object reference in a shared domain of conversation. *Pragmatics & Cognition, 6*(1-2), 121-152.

Buchs, C., Gilles, I., Antonietti, J.-P., & Butera F. (2016). Why students need to be prepared to cooperate: A cooperative nudge in statistics learning at university. *Educational Psychology, 36*(5), 956-974.

Chafe, W. L. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing.* Chicago: University of Chicago Press.

Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General, 140*(1), 102-116.

Deutsch, M. (1949). A theory of co-operation and competition. *Human Relations, 2*(2), 129-152.

Geenen, J. (2013). Kitesurfing: Action, (inter)action and mediation (doctoral dissertation). Auckland University of Technology, Auckland.

Goodwin, C. (1984). Notes on story structure and the organization of participation. In M. Atkinson, & J. Heritage (Eds.), *Structures of Social Action* (pp. 225-246). Cambridge: Cambridge University Press.

Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. Morgan (Eds.), *Syntax and Semantics.* 3: Speech Acts (pp. 41-58). New York: Academic Press.

Heritage, J. (2012). Epistemics in action: Action formation and territories of knowledge. *Research on Language & Social Interaction, 45*(1), 1-29.

Iverson, J. M., & Goldin-Meadow, S. (1997). What's communication got to do with it? Gesture in children blind from birth. *Developmental Psychology, 33*(3), 453-467.

Johnson, D. W., & Johnson, R. T. (1974). Instructional goal structure: Cooperative, competitive, or individualistic. *Review of Educational Research, 44*(2), 213-240.

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Minneapolis, MN: Interaction Book.

Johnson, D. W., Johnson, R. T., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review, 19*(1), 15-29.

Kastberg, P. (2007). Knowledge communication: The emergence of a third order discipline. In C. Villiger, & H. Gerzymisch-Arbogast (Eds.), *Kommunikation in Bewegung: Multimedialer und multilingualer Wissentransfer in der Experten-Laien-Kommunikation* (pp. 7-24). New York: Peter Lang.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica, 26*, 22-63.

Khoo, Y. Y. (2015). Collaborative problem solving promotes students' interest. *Journal of Economics and Economic Education Research, 16*(1), 158-167.

Kraut, R. E., Gergle, D., & Fussell, S. R. (2002). The use of visual information in shared visual spaces: Informing the development of virtual co-presence.

Proceedings of the 2002 ACM Conference on Computer Supported

Cooperative Work, 31-40.

Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.

Louwerse, M. M., & Bangerter, A. (2005). Focusing attention with deictic gestures and linguistic expressions. In B. Bara, L. Barsalou, & M. Bucciarelli (Eds.), P*roceedings of the Cognitive Science Society (pp. 1331-1336).* Mahwah, NJ: Lawrence Erlbaum Associates.

McNeill, D. (2005). How Gestures Carry Meaning. In *Gesture and Thought* (pp. 22-\ 59). Chicago: University of Chicago Press.

McNeill. D., & Levy, E. T. (1993). Cohesion and gesture. *Discourse Processes, 16*(4), 363-386.

Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. London: Routledge.

Norris, S. (2006). Multiparty interaction: A multimodal perspective on relevance. *Discourse Studies, 8*(3), 401-421.

Norris, S. (2011). *Identity in (inter)action: introducing multimodal (inter)action analysis.* Göttingen: De Gruyter Mouton.

Norris, S. (2013). What is a mode? Smell, olfactory perception, and the notion of mode in multimodal mediated theory. *Multimodal Communication, 2*(2), 155-169.

Norris, S., & Jones, R. H. (2005). Introducing mediated action. In S. Norris, & R. Jones (Eds.), *Discourse in action: Introducing mediated discourse analysis* (pp. 15-19). London: Routledge.

Norris, S., & Pirini, J. (2016). Communicating knowledge, getting attention, and

negotiating disagreement via video conferencing technology: A multimodal
analysis. *Journal of Organizational Knowledge Communication, 3*(1), 23-48.

Rossano, F. (2012). Gaze behaviour in face-to-face interaction (doctoral dissertation).
Radboud University, Nijmegen.

Roth, W.-M., & Lawless, D. V. (2002). Signs deixis, and the emergence of scientific
explanation. *Semiotica, 138,* 95-130.

Scollon, R. (2001). *Mediated discourse: The nexus of practice*. London: Routledge.

Van Dijk, T. A. (2001). Multidisciplinary CDA. In R. Wodak, & M. Meyer (Eds.),
Methods of Critical Discourse Analysis (pp. 95-120). London: Sage
Publications, Ltd.

Van der Vegt, G., & Van de Vliert, E. (2002). Intragroup interdependence and
effectiveness: Review and proposed directions for theory and practice. *Journal
Managerial Psychology, 17*(1), 50-67.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological
processes*. Cambridge, MA: Harvard University Press.

Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated
action.* Cambridge, MA: Harvard University Press.

9. Appendix

9.1 Design

Duplo Building Replication

Instructions

As a group, you have a very specific task to accomplish. You need to build an exact replica of a Duplo Lego structure which will be sitting on tables 1, 2 and 3. Each example structure is the exact same and you will all have the same view. You will be unable to see the back side of the three dimensional structure and this will makes it very difficult, if not impossible, to replicate the structure. So, there are a number of "cheats" available that will assist in the building task itself. Once you have read all the instructions, decided on roles and selected "cheats" the example structures will be placed on tables 1, 2 and 3 and you can begin the build.

First some ground rules:

1. You will have to divide the workload amongst the group members. 1 group member will be the Builder, while the other 2 group members will be able to assist the builder, but they cannot physically handle the building blocks themselves.

2. After you have made your decision about who will perform each role, you must collaboratively decide on a selection of 2 "cheats" which will help you during the build process. There are 5 available "cheats" and a selection of *any* two will provide enough help so that you can collaboratively accomplish the build process. However, some "cheats" may be more helpful than others.

3. Once you have come to a decision about your selection of "cheats", you will have to assign each cheat to the *construction assistants*. Each cheat however, cannot be physically shared between group members. You will need to help the builder create an exact replica drawing upon the information you have available from the "cheats".

Figure 5.

Cheats

Please note: if a cheat involves taking pictures, videos or turning or manipulating the structure, only one group member is able to physically perform the task. The others may help in any way they would like, but they must remain seated during the process.

1. One group member will be provided 2 pictures of the structure. These pictures will be of sides of the structure which you cannot currently see. You can use these pictures to assist in the build process but you cannot show them to group members.

2. One group member can use the tablet to take 1 picture of the structure from any vantage point they desire. They cannot however, move the structure in any way. The group member can discuss with the team which pictures to take, however, the other team members cannot leave their chairs during the picture taking. You will have 2 minutes to take the picture. After the picture is taken, *that example structure will be removed* from its table for the duration of the build process.

3. One group member can place a mirror in any position on any side of an example structure. Once the mirror is in position, that group member will have to return to their seat for the duration of the building process.

4. You can ask for one example structure to be broken in half, then the group member with a direct view of this replica can place each piece in any position they would like. Both pieces must be placed back on the table, and that construction assistant will have to return to their seat for the duration of the building process.

5. One group member can move one of the example structures in any way they would like (ie turn it around, rest it on its side etc). They can only move the structure once, and then they must return to their seat for the duration of the building process.
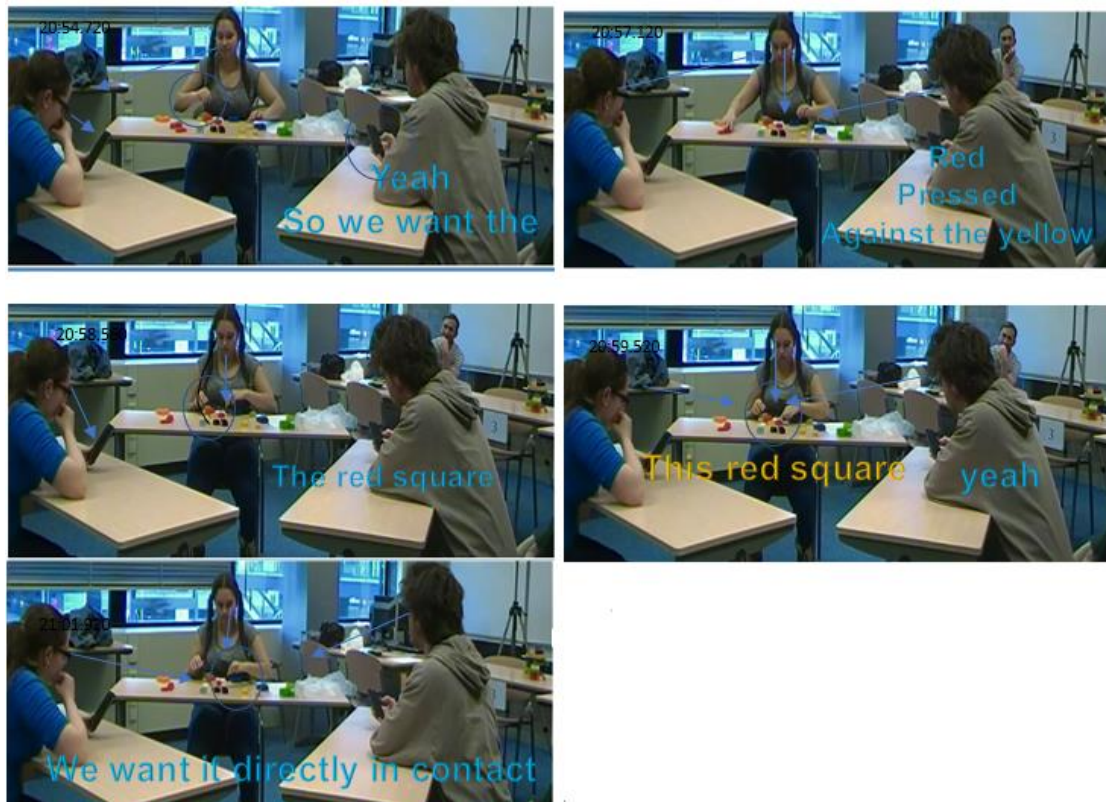
Figure 6.

9.2 Transcripts

Figure 7.

Figure 8.
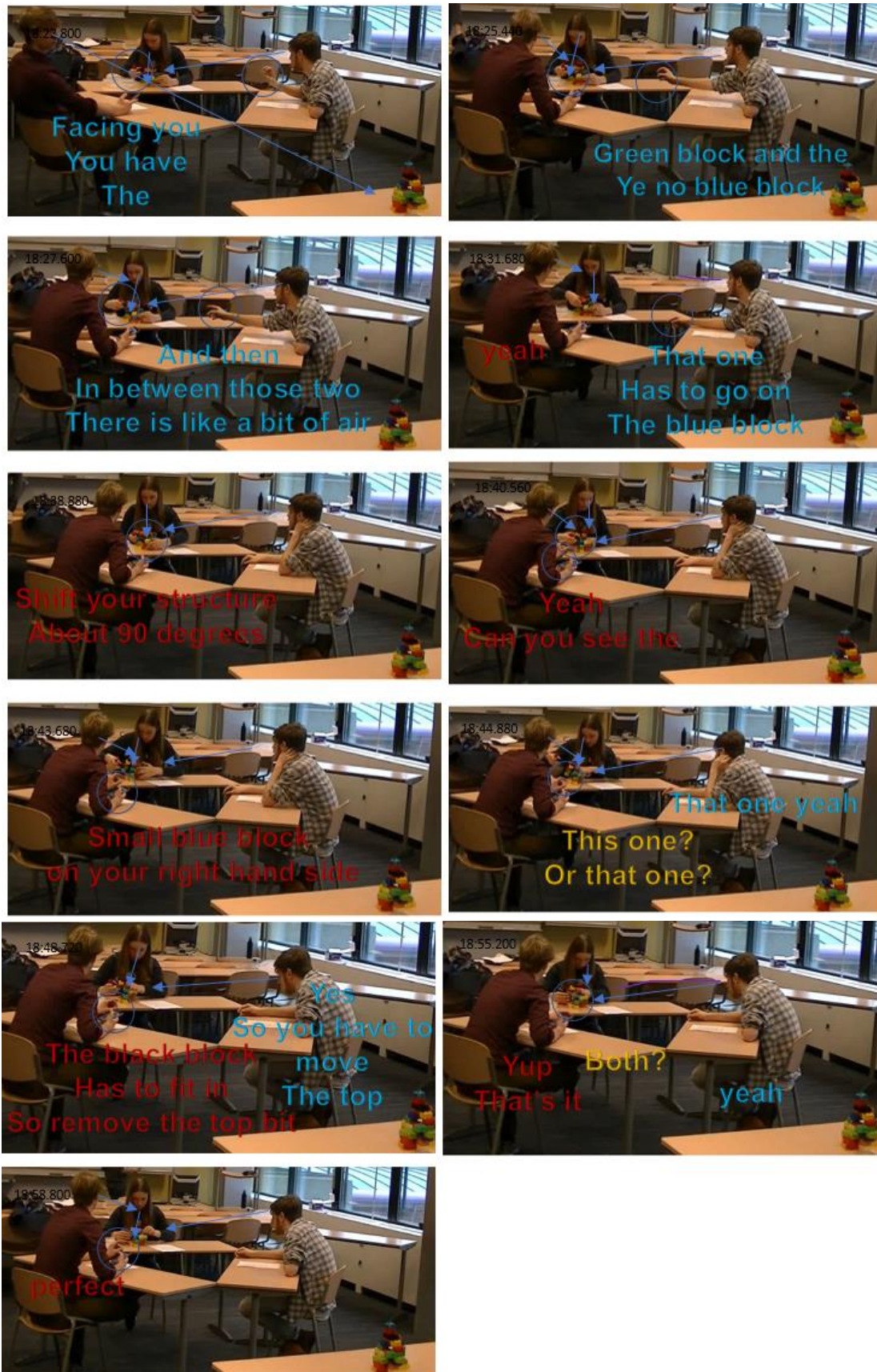
Figure 8 (continued).

Figure 9.

Figure 9 (continued).