Temporal alignment of hand gestures and speech within and between dyads of native and non-native speaking children

M.D Soederhuizen

Master Thesis - Developmental Psychology

S3639428
Januari 2024
Department of Psychology
University of Groningen
Examiner/Daily supervisor: Dr. Lisette De Jonge-Hoekstra
Second examiner: Dr. Gerrit Breeuwsma

Abstract

Within 3 studies, we investigated interpersonal temporal alignment within dyads of children speaking in their first (L1) or second language (L2). In study 1 we developed an experiment protocol called 'This-or-That debates phrases' to study spontaneous and non-verbal communication in children. This was done through creating five experiment tasks, that we evaluated through a list of observation points, and selecting the most successful protocol. In study 2 we analysed several quantitative and qualitative aspects within the video data of four Dutch and three Dutch-Frisian dyads that completed the protocol task. ELAN 6.4 and Microsoft Excel were used for this analysis. We concluded that the developed protocol is suitable to study differences and similarities in verbal communication and hand gesture use in dyads of children speaking in their L1 or L2 (Dutch). In study 3 we ran a first attempt to use OpenPose, a body movement tracking program, to analyse interpersonal temporal alignment of hand movements and speech in two dyads of children speaking in their L1 or L2. We compared the average temporal difference (ms) between peaks in the intensity of speech and gestures of both individuals within the dyads. We found that there might be a higher degree of interpersonal speech-gesture alignment in speech partners with a different L1 compared to a similar L1. We recommend future research to further investigate this relation. More insight in the verbal and non-verbal coordination between speakers across languages increases understanding of important communication processes in, for instance, educational settings.

Key concepts: Temporal alignment, hand gestures, second language, speech-gesture alignment, interpersonal synchronisation

Table of contents

Temporal alignment of hand gestures and speech within and between dyads of native and non-native speaking children

Hand gestures are increasingly viewed as an aspect of communication that entails several complex factors (Gullberg, De Bot, & Voltera, 2010). Gestures appear to be a part of a dynamic, intertwined system that connects verbal and non-verbal communication (Lowie & Verspoor, 2022). Speech and multiple other contextual factors have an interactional influence on people's use of gestures (Gullberg et al., 2010). At the same time, gestures have an influence on the perception, interpretation and memorization of speech. In this thesis, we will mainly focus on interpersonal coordination of gestures and speech as mediated by language. More specifically, we will focus on interpersonal differences in gesture and speech use in the context of second language use within and between primary-school children. In order to do so, a task measuring interpersonal coordination in temporal alignment of hand gestures and speech in children with a different first language will be first designed and piloted. Then the collected data will be analysed. This introduction provides context to this procedure and sets out the main findings within the fields of gesture development, speech-gesture interaction and bilingual influences on interpersonal communication.

Gesture development

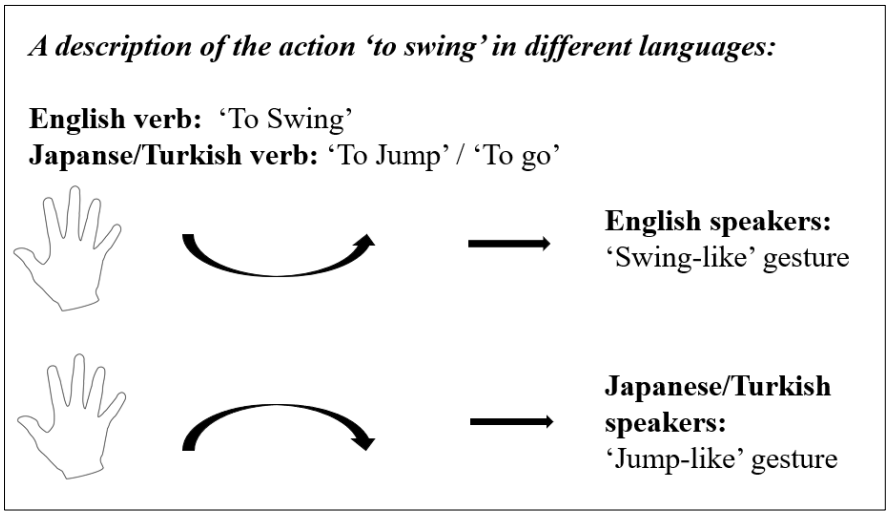
Although it is known that infants already start using non-verbal gestures before they start producing words, no exact answer has been provided to the question when exactly gesture and speech integrate into a synergetic system (Gullberg et al., 2010). Yet, it is clear that the context in which children are learning their first gestures and their first words is similar. Both their first words and gestures refer to simple objects in their environment. Gestures that are used to point at objects are called deictic gestures and infants usually start to express these at 9 to 10 months old (Rusiewicz & Esteve-Gibert, 2018). Within the period of 15 to 26 months old, infants start to use representational gestures. These gestures are iconic (concrete depictions) or metaphorical representations (abstract depictions) of objects or actions. Interestingly, children usually start referring to an object with deictic gestures before these same objects are firstly described with speech (Iverson & Goldin-Meadow, 2005; Özçalışkan & Goldin-Meadow, 2005 as cited in Rusiewicz & Esteve-Gibert, 2018). On the other hand, representational gestures usually occur when the same events have already been described with speech (Özçalışkan et

al., 2013 as cited in Rusiewicz & Esteve-Gibert, 2018). A specific type of iconic or representational gestures are gestures that have an established meaning but are a more abstract depiction of an action or attribute (Gullberg et al., 2010). These kind of gestures have been referred to with many different terms, yet here we will refer to them as conventional gestures. An example of a conventional gestures is the 'thumbs up' gesture (See also Wagner et al., 2014). When children are two to three years old they also start to use conventional gestures (Perrault et al., 2019 ; Hodges, Özçalışkan & Williamson, 2017). Later in development children start to use a third category of gestures which are referred to as beat or rhythmic gestures (Rusiewicz & Esteve-Gibert, 2018). Beat gestures usually do not convey meaning in themselves and are co-occurring with speech. As not many studies have focussed on these kinds of gestures, no consensus has been reached on the precise period during which children start to use them to accompany their expressions and narratives or on the order in which they integrate with rhythmic speech. Mayberry and Nicoladis (2000) stated that children from two years on already use beat gestures in a way similar to adults. On the contrary Colletta (2004) argues that children increasingly start to use abstract, metaphoric and beat gestures a couple of years later, at primary school-age, when they have further developed their conversational skills.

The amount and type of gestures children use is furthermore influenced by the people and cultures that are present in the proximal environment of children (Gullberg et al., 2010). Italian children, for example, use language with many gestures that refer to actions and attributes. This can be explained by the fact that they also see a lot of these gestures being used around them. In contrast American children mainly use more conventional gestures (Iverson et al., 2008 as cited in Gullberg et al., 2010). In a literature review Kita (2009) states that every culture further has their own specific conventional gestures, which entail general cultural meaning (for example, a thumbs up meaning success in many western cultures). People's use of gestures is also influenced by the lexical content of a language. In a study comparing English, Japanese and Turkish language, Kita and Özyürek (2003) found that the word used to describe the English verb 'to swing' does not exist in Japanese and Turkish. Instead, these languages use a verb more similar to the verbs 'jump' or 'go'. As a consequence the Japanese and Turkish speakers also use a different co-occurring gesture when they described the same events as English speakers. This example is illustrated in Figure 1.

Figure 1

Illustration of an example of semantical influences on gesture use provided by Kita and Özyürek (2003)



Integration of gesture and speech

In order to illustrate the exact way in which gestures co-occur with speech multiple language models have been designed by different researchers (Wagner et al., 2014). An important distinction can be made between researchers that state that gestures bear meaning in itself apart from verbal expressions (see Cienki & Müller, 2008) or researchers that view language and gestures as an integrated system (see McNeil, 1992). When gestures and speech are observed together there are different ways in which they organize. In many cases they do align in their pragmatic function and refer to the same thing (Wagner et al., 2014). In other cases, however, the synchronization of (hand)gestures and speech is not so clearly observable. When the meaning of a gesture does not align with the semantic expression it accompanies, it is referred to as a semantic mismatch (Goldin-Meadow, 2003 ; De Jonge-Hoekstra et al., 2021).

More recent and complex models of speech-gesture interaction do incorporate the exact temporal coordination and alignment of speech and hand gestures (Wagner et al., 2014). The specific way in which gestures are timed in relation to speech is referred to as temporal alignment (De Jonge-Hoekstra, 2021; see also Pouw & Dixon, 2019). Part of the discussion on temporal alignment in gesture and speech has been around the question if gestures occur simultaneously or non-simultaneously in comparison to verbal expressions, when one closely studies their rhythmic pattern. When gestures and

verbal expression do not align in timing, meaning that a hand gesture does not occur during the peak of its co-occurring speech, one could speak of a mismatch in timing (Leonard & Cummins, 2010, as cited in Wagner et al., 2014).

Both semantic and temporal mismatches have been studied in several contexts and linked to developmental processes (De Jonge-Hoekstra et al., 2021). Multiple researches have shown that mismatches of gesture and speech occur when children and adults are on the verge of learning something (De Jonge-Hoekstra et al., 2021 ; see also Goldin-Meadow, 2003). When answering questions that require insight (for example questions about the contents of differently shaped glasses) some children first showed non-verbal understanding through their hand gestures before verbally expressing their thoughts (e.g. Goldin-Meadow, 2003; Church & Goldin-Meadow, 1986). In a study following up on previous research, De Jonge-Hoekstra and colleagues (2016) found that when children engage in a scientific task, the temporal attunement between children's gestures and speech differed greatly. Variability in timing was seen between individual children, children of different ages and in different levels of difficulty and understanding of the task. They suggest that speech and gesture are organized in a suboptimal manner, when for example engaging in a novel or difficult task. However, De Jonge-Hoekstra and colleagues (2021) found that there was more temporal alignment of gesture and speech in adults engaging in an easy cognitive task compared to a difficult one. All in all, findings confirm that the relation between speech and gestures is quite dynamic, complex and above all context dependent. Yet better understanding of these interactional processes could provide important information about developmental and learning processes within individuals.

Interpersonal coordination of gestures

When people engage with each other they tend to coordinate their behaviour in similar patterns (Richardson et al., 2007). This is referred to as interpersonal synchronization (De Jonge-Hoekstra., 2021) or alignment. An example of this is when two people are both rocking in chairs and eventually rock in the same rhythm (Richardson et al., 2007). Or in another case, when two people are performing knock-knock jokes and then automatically adapt the timing of their sentences to the rhythm of the joke (Schmidt et al., 2014). Within conversations, there are also multiple levels in which people tend to synchronize their behaviour (Rasenberg et al., 2020). This includes the timing of phrases and words,

their use of specific words and structures and the tone of their conversations but also their use of gestures.

In a recent review Rasenberg and colleagues (2020) provide an overview of a broad scala of studies that have been conducted on behavioural alignment during conversations. They also provide an important aspect of the current theoretical background behind interpersonal coordination. In summary, they describe the difference between the theories of grounding and priming alignment. Grounding refers to a more considerate or purposeful way of aligning with a speech partner, where alignment might play a role in the outcome of a conversation. Priming, on the other hand, suggests that alignment occurs automatically and involves the purposeless copying of the other's verbal or nonverbal behaviour.

De Jonge-Hoekstra (2021) investigated interpersonal coordination of speech, head movements, and hand movements in school-aged children (aged 6 - 10 years). Her research appears to be the first to study interpersonal coordination in children within collaboration tasks. She found that, just like adults, children align their hand and head gestures when engaging in a discussion task although it appeared in a slightly different manner. Furthermore, within the task itself she found no relation between the degree of synchronisation of the two children and the outcome of the discussions. The study of De Jonge-Hoekstra is only a first attempt to study interpersonal gesture-speech synchronization in children and one should bear in mind that developmental differences between children and adults could play a role in differences in alignment between children and adults.

Bilingual influences on gesture-speech alignment

Another aspect that could possibly influence interpersonal synchronization in speech and gestures is the first language of an individual. As mentioned, earlier researches have pointed out differences in use of hand gestures between different languages (see for example Gullberg, 2011 or Kita and Özyürek, 2003). Recently researchers started to explore the cross lingual effects of these differences in the context of second language learning acquisition. Lowie and Verspoor (2022) describe second language acquisition as a complex dynamical process and point out that the context of a learners' first language (L1) could influence their use of gesture and speech in a second language (L2). Most of the cross-lingual differences that have been studied so far concentrate on language learning and fluency (Ordin and Polyanskaya, 2014 ; Kosmala, Candea and Morgenstern, 2019 ; Gullberg, 2011) or

participants response to certain prompts (e.g. Hoetjes, 2018). Only one study focussed on interpersonal coordination (Schneider et al., 2020).

Within this section we will summarize the most important findings from the literature mentioned above. Ordin and Polyanskaya (2014) suggest that children experience more difficulty learning the rhythm and intonation of a second language when these aspects deviate further from their first language. In relation to non-verbal communication Kosmala and colleagues (2019) argue that L2 speakers sometimes use hand gestures in relation to their disfluent speech and might therefore express a higher amount of gestures compared to L1 speakers. An example they provide is of a speaker who struggles to find a certain word, and instead makes a representational hand gesture of this same word, while suspending her speech and looking at her speech partner. This behaviour seems to be in line with the finding that hand gestures of L2 speakers can be related to moments of 'thinking' or meant to elicit help from a L1 conversation partner (Gullberg, 2011). A study of Hoetjes (2018) investigated the lexical content of languages. She asked Dutch speakers to describe 'placement events' in English to explore differences in their use of verbs and gestures that refer to the placement of objects. She found that even when the Dutch participants used similar verbs to English participants, they used different types of gestures. This suggests that even when L2 speakers understand the meaning and use of the verbal content they use, they still transfer the gestures of their L1 instead of adapting to common gestures of native L2 users. Schneider and colleagues (2020) were the first to explore interpersonal coordination in speech between L1 and L2 conversation partners. Specifically, they looked at convergence of speech, which refers to the similarity (or synchronicity) of the language use of two speakers. They asked speakers of both English and Spanish to engage in three conversations about movies, music and television. They compared the convergence of speech in three language combinations: English - English, Spanish-Spanish and English-Spanish. They found that similarity in the style and character of speech occurred in all three language combinations and was independent of linguistic content.

The researches that we presented above did mostly compare English with another language. In this thesis we will study cross-lingual influences of the Frisian language as transferred to the Dutch language. In order to understand potential differences in Frisian individuals it important to have some

background knowledge on previous studies that compared these two specific languages. Therefore, we will describe some specific features of the Frisian language within the next section.

Frisian language

Frisia is a bilingual province in the Netherlands. Children with two Frisian parents usually converse in Frisian within the family environment and get in touch with the Dutch language outside their home (Informatievoorziening, provincie Fryslân, 2020). In elementary schools in Frisia, Dutch is usually spoken as the dominant language (Bosma & Blom, 2020). Originally, the Frisian language is closer to the English language compared to Dutch. Yet, since Dutch and Frisian are used alongside, the languages have started to converge (Gooskens & Heeringa, 2004). In lexical content, Dutch and Frisian show extensive overlap, though there also exist various differences in, for instance, inflection rules (Bosma & Blom, 2020). Furthermore, intonation patterns in Dutch and Frisian mainly overlap (Nota et al., 2016). Yet, Nota and colleagues found slight differences in the alignment of the peak and pitch of speech in Frisian-Dutch bilinguals. In a study with 5- and 6 year old Frisian-bilingual children, Bosma and Blom (2019) found that children had fewer problems shifting from Frisian to Dutch than the other way around. They hypothesize that this is due to the fact that in Frisian, replacing words in Dutch is commonly accepted. They also found that younger children more often shifted to Frisian when speaking in Dutch, compared to older children.

Methodological challenges

As set out earlier in this section, there is only a small scope of studies that closely investigate interpersonal temporal alignment of gesture and speech. The limitation of studies on temporal gesture-speech alignment is partly due to methodological challenges (Wagner et al., 2014). For a long period technology used to analyse video recordings was not openly or easily accessible, which made researchers reliant on empirical and manual methods that are very time consuming. Yet, the amount of hardware and software to use in communicative research has been growing rapidly (for an overview see Wagner, 2014). Recently, Matahelumual (2022) looked into the possibility of using OpenPose to study temporal alignment of speech and hand gestures of bilingual participants. OpenPose is an open source software that enables users to track a wide range of body movements varying from facial features to head-, feet- and head-movements by either analysing 2D video recordings or live footage (Cao et al.,

2019). Matahelumual used the video material of Hoetjes (2018) and recordings of online Zoom meetings. She successfully used OpenPose to analyse participants' hand gestures in video recordings of videocalls. However, in her study she describes some key-features of the experimental set-up and material that must be met in order to use video data in OpenPose. Important is that the upper body of the participants' fits within the frame from their forehead until their middle. Furthermore, the camera frame rate must be sufficiently high to be able to analyse faster hand movements (i.e. 50 f/s). Lastly, the audio quality must be sufficient to detect peaks in the voice of participants in the presence of background noise.

A second challenge within the context of studying communication in a quantitative way regards the set-up of experiments. One of the requirements that is ideally met in order to study natural interactions is the evocation of spontaneous conversation between two participants. Though, forcing spontaneous speech is challenging (Roberts, 2014). Roberts gives some guidelines in developing experimental tasks to elicit natural conversations between two partners that are more generally applicable in research on non-verbal communication. She describes that experimental set-ups should be structured in a way that there is enough space for natural conversation. Furthermore she suggests that videorecording could be a valuable resource in respect of making detailed analyses afterwards. Yet she argues that the closer the experimental setting is to the real environment, the higher is the ecological validity of the results. Opposed to this, an experimental setting might benefit the interpretability of the material. To achieve a successful experimental task she mentions three basic characteristics; Firstly the setting of the experiment must closely resemble the home environment of the participants. Secondly, the proposed task must be familiar to the participants and thirdly the task must be engaging and should provoke emotions.

Current Research

In a review on the current progression in the field of convergence of language use between speech partners, Gullberg (2021) states that current data on the convergence of language users across languages demonstrates that gesture and speech should be assessed as a bimodal phenomenon. Whereas many researchers have looked into monolingual settings, she stresses that more data is needed to understand processes of bilingual convergence. Her theoretical perspective also extends to the use of

non-verbal communication such as hand gestures. So far previous research has confirmed that interpersonal temporal synchronization of gesture and speech occurs between children in the age range of 6-10 years (De Jonge-Hoekstra, 2021). Furthermore, multiple studies have shown interlingual effects on communication in speech and gesture use (e.g. Hoetjes, 2018 ; Schneider et al., 2020). To investigate a new analysis method Matahelumual (2022) successfully used OpenPose to study temporal alignment of gesture and speech within bilingual adults. Yet, it seems that no existing study has directly looked into the process of interpersonal temporal alignment of speech and gesture in bilingual adults or children. In order to build on previous research this thesis presents three consecutive studies that answer three related questions. Firstly, a pre-pilot will be conducted in order to design a protocol to investigate interpersonal temporal alignment within children. This protocol will be used in the second study, which will be piloting it as a revised version. The third and final study will use the data from two dyads of children from the pilot in order to investigate whether OpenPose could be used to investigate differences in interpersonal temporal alignment between dyads of native and non-native children. This third study will be performed in line with the analysation process that was developed within the thesis of Matahelumual (2022). Together these three studies will answer the following three research questions:

1. What type of task can be used to elicit spontaneous use of speech and hand gestures in dyads of school-aged children?

2. Can the most promising task from study 1 be used to investigate interpersonal temporal alignment of hand gestures and speech in dyads of native and non-native school-aged children?

3. Can the software Openpose be used to analyse video data of the task of study 2 to investigate interpersonal temporal alignment of hand gestures and speech in dyads of native and non-native school-aged children?

The development of a task that elicits spontaneous and non-verbal communication in school-aged children would be very useful to study processes of interpersonal communication in children. In addition, knowledge about the influences of bilingualism in communication is important, as the number of bilinguals worldwide has been estimated to be more than 50% of the entire population (Leikin et al., 2011). Thorough understanding of the influences of bilingualism on interaction and communication is

necessary to produce useful advice to implement in, for instance, educational practices. Finally, new evidence of similarities in synchronization between speakers with a same and different native language would contribute to the view of language as a multimodal system and contribute to future research.

Study 1: Pre-pilot

As a brief reminder, this pre-pilot was conducted in order to develop a protocol for an experimental task to elicit spontaneous and non-verbal communication in school-aged children (6 to 10 years old). Earlier it has already been mentioned that this is something quite challenging to achieve. Within the execution of this pre-pilot we used the recommendations made by Roberts (2014). In addition we formulated a list of prerequisites that should be met by the tasks at hand. Finally, the experimental task that we considered most successful was selected and prepared for further use.

Participants

In total, 10 dyads of pupils (N=20) participated in this pre-pilot. All children were 7 or 8 years old. All dyads consisted either of two girls (n =10) or two boys (n =10) and the children knew each other from their class. All dyads participated in one experimental task. For each of the five experimental tasks within this pre-pilot, one dyad of boys and one dyad of girls was selected. For the selection of participants we made use of samples from two different elementary schools in Groningen. In both schools we selected children from the fourth year of Dutch the elementary school (equivalent to the American 2nd grade). In both schools the teachers picked pairs of pupils based on whether they could miss out some of the instructions or schoolwork. The type of tasks in which the children participated was dependent on the order in which the children were selected by the teacher, whereas the experimenter followed a pre-ordered list of tasks.

Materials

For the pre-pilot study we made use of five experimental tasks that were designed to conduct in the Dutch language. These were developed beforehand. In all five tasks we used audio recordings, illustrations or videos to elicit conversation between children in a dyad. For each task a short description is provided below. The complete preliminary designs are included in Appendix A. All necessary material that was part of the protocols was shown to the participants via a laptop. The experimenter used a timer, paper and pencil to make observations.

Protocol 1: Dog story

The material we use in protocol 1: Dog story entails a picture story of two children that go for a picnic (Heaton, 1966). Their dog secretly crawls in their basket with food and eats the food they brought for the picnic. The children will have to find out what happened in the picture story by combining the visual input they both separately received. The picture story consists of 6 pictures. One of the children will receive the first and second picture and the other child the third and the fifth picture of the story (picture 4 depicts redundant information). They are both shown the last picture of the series, that reveals the clue of the story that leads from the pictures that the children already received.

Protocol 2: Snack Attack

The material we use in protocol 2: Snack Attack entails an animated video with unintelligible speech of an old lady that buys herself cookies from a vending machine (Verastegui, 2016, June). On the railway platform, it seems that a boy sitting next to her is eating her cookies. The lady gets angry at the boy. In the end, the whole scene is revealed as being a misconception whereas the boy actually had the same cookies with him. It turns out the lady was unconsciously stealing his cookies while having hers in her purse. The children have to discuss together what happened in the video and specifically with the cookies.

Protocol 3: Joy

The material we use in protocol 3: Joy entails an animated video of a little dog that goes fishing with his owner (JD.COM, 2018, February). There is no use of any speech in this video. A heron bird attempts to steal the bait worms of the fisherman. The dog tries to prevent this and scares the bird away. In the end it turns out that the heron wants to feed her babies with the worms, since the baby birds seem to dislike fish. After seeing this the dog offers the worms to the heron, while disobeying his owner. The children have to discuss if the dog did something bad or good.

Protocol 4: The mole who knew it was none of his business

The material we use in protocol 4: The mole who knew it was none of his business entails a video of the Dutch version of the story book 'The mole who knew it was none of his business' (Dutch title: Over een kleine mol die wil weten wie er op zijn kop gepoept heeft) (Lisette van der Beek, 2020, April 6). A woman reads out the book using her voice and filming the pictures. In this story an animal

pooped on the head of the mole after which he goes by different animals to ask if they were the ones who did this. All animals show their type of poop. Finally, it turns out that the dog pooped on the head of the mole. The mole poops on the head of the dog as a payback. The children have to focus on either the types of animals or the types of poop in the story, together they have to recall which poop belonged to which animal.

Protocol 5: This-or-That debate phrases

The material we use in protocol 5: This-or-That debate phrases entails 13 debate phrases which are given by the experimenter. The children have to reach consensus on these phrases by discussing together. An example is 'Would you rather be able to be invisible or be able to fly?'. All phrases are written out in Appendix A under protocol 5.

Procedure

All five protocols were tested with a pair of boys and a pair of girls. The trials were performed during class- or break time within the school environment of the children. During the run of each protocol the experimenter observed the behaviour of the participants. The focus of these observations was around several aspects that were deemed important for an experimental tasks that has the goal of eliciting spontaneous dialogue within children: the degree in which they found the instruction complex, the length of their conversation, the share of each child in the conversation, their enthusiasm for the experiment, the degree to which the task seems appropriate for their age and the amount of hand gestures they used. In Table 1 the specific points of the observation are listed together with a series of questions that further specify them. Based on these observations, the experimenter chose the protocol that met the most requirements to use in the follow up study. Finally, it was written out in a standardized experiment protocol for the use in the pilot study (See Appendix B). A couple of adaptations were made to the initial task.

Table 1

Observation points pre-pilot

Observation points	**Observation questions**
1 Complexity of the instruction	- Do the children understand the task without many extra instructions?
2 Length of the conversation	- How long does it take to complete the task? - How much of a conversation do the children have before providing their final answer?
3 Share of each child in the conversation	- Do both children speak as frequently? - Is one child taking a dominant role in the conversation?
4 Enthusiasm of the children	- Do the children spontaneously engage in the task? - Do the children need a lot of encouragement? - Do the children like the material used in the experiment?
5 Age appropriateness of the task	- Are the children cognitively ready to engage in the task? - Do the children understand the material presented in the task?
6 Amount of hand movements the children used	- Do the children use hand gestures accompanying their speech?

Results

The first protocol (Dog story) lasted five minutes or more for both dyads, yet a lot of encouragement was needed to keep the conversation going and to keep both children engaged. The children did engage with and understand the instructions of task. The picture story itself seemed to be too complex for the children. It seemed that they did not recognize the old cartoon drawing as familiar. The marmalade jar and sandwich were not recognized as food and the children did not get to the clue of the story until they received a lot of tips. The girls were using more hand gestures than the boys, yet only towards the end of the discussion.

The second protocol (Snack Attack) lasted over five minutes for the girls and around three minutes for the boys. All children were very excited about the content of the video. Yet, they did not understand the instructions and the content of the video well enough to engage in a dialogue. The boys did use hand gesture and one stood up to mimic a character. Yet, one of the boys was talking a lot more than the other boy. The girls were quiet and needed a lot of encouragement.

The third protocol (Joy) took about one and a half minutes for the girls and half a minute for the boys. The children did engage in watching the video. Yet, the question they were asked did not spark a lot of conversation within the children. The girls seemed not sure about what to answer and needed the questions to be repeated. Within the boys' dyad, one of the two boys was very dominant and gave the correct answer almost right away. There was little sign of use of hand gestures.

The fourth protocol (The mole who knew it was none of his business) lasted just under five minutes for the girls and about one and a half minute for the boys. The children were enthusiastic about the content and understood the instructions. The answers that were given were a sum up of what they heard in the story. There was little dialogue. The girls were counting on their fingers to keep on with their answers. Other than that, little hand gestures were used.

The fifth protocol (This-or-That debate phrases) took approximately five minutes for both dyads. Both the boys and girls that participated in the experiment followed the instructions and the further procedure of the task well and did both equally and enthusiastically engage. In addition, they did both express non-verbal communication using their hands. Yet, some statements within the tasks worked better than others. The phrase, 'would you rather have a dog or a cat' did not spark elaborative conversation in the children whereas the phrase 'would you rather be able to fly or to be invisible' did evoke a lot of verbal and non-verbal speech in all four children. In Table 2, an overview of all observation points and protocols is given.

Table 2

Evaluation of observation points

Observation points	**Protocol 1 (Dog story)**	**Protocol 2 (Snack Attack)**	**Protocol 3 (Joy)**	**Protocol 4 (The mole who knew it was none of his business)**	**Protocol 5 (This-or-That debate phrases)**
1 Complexity of the instruction	✓	**X**	**X**	✓	✓
2 Length of the conversation	✓	✓	**X**	✓	✓
3 Share of each child in the conversation	✓	**X**	**X**	**X**	✓
4 Enthusiasm of the children	✓	✓	**X**	✓	✓
5 Age appropriateness of the task	**X**	**X**	**X**	✓	✓
6 Amount of hand movements the children used	**X**	✓	**X**	**X**	✓

Discussion

In this study we piloted five experimental tasks that are supposed to elicit spontaneous and non-verbal communication in school-aged children. The material used in the tasks existed of either auditive, video or imagery cues. The results point out that the fifth protocol (This-or-That debate phrases) was the most successful since it was the only protocol that met all the requirements stated in Table 1 (see also Table 2). Interestingly, within this task the more abstract statements were more successful than the concrete statements. This might be due to the level of imagination that is involved in the answers that the children provide. These findings are also coherent with the recommendation made by Roberts (2014), who stresses the need of a task that provokes engagement and emotion in order to stimulate natural conversation in participants. Especially the more abstract statements provoked emotions in the children and might therefore have encouraged them to answer in a natural and elaborative manner. The adapted protocol that will be used in the pilot, should therefore also include many abstract statements.

From the pre-pilot we further learned that tasks with a single or pre-existing answer are more sensitive to a situation in which one of the children takes the lead and talks more than the other. Which was the case for the boys that engaged in protocol two and three (Snack Attack and Joy). We also learned from the fourth protocol (The mole who knew it was none of his business) that children might use finger counting within tasks that include answers that must be given in a certain order, which might prevent them from using other spontaneous hand gestures. Finally, we learned that when instructions are broad and refer, for example, to an entire video or story (i.e. protocol 1, 2 and 3) children need more cues to be able to provide elaborative answers. These findings might be an indirect explanation for the fact that protocol 5, which provides many smaller cues and requires equal participation from each child, works well.

Study 2: Pilot

In this study we used an improved protocol of the This-or-That debate phrases (see Appendix B) task to investigate if this task could be used to research interpersonal temporal alignment of hand gestures and speech in dyads of school-aged children speaking in a first or second language. In order to do so we performed a small pilot study using four pairs of Dutch children as a control group and three dyads of Dutch and Frisian children as an experiment group (more details will follow below). The first goal of this study is to explore whether the revised This-or-That debate phrases task is a suitable task to elicit spontaneous communication and hand gestures in dyads of children. The second goal is to evaluate specific aspects of the experiment to explore whether improvements to the design should be made. The final goal is to explore differences within the individual children or between the groups, which could provide useful information for follow up studies.

Participants

For the selection of the research participants for the pilot study, we made use of a convenience sample from an elementary school in Frisia and children living in a small village in Noord-Holland. We recruited 14 children, including girls (n = 8) and boys (n = 6). The selection of the children was based on both their own and their parents' consent to participate in the experiment. The age of the children varied from 6 to 9 years (M_{age} = 7.88, SD_{age} = 0.95). Eleven of the children spoke Dutch as their first language (L1) and three of the children spoke Dutch as their second language (L2) and Frisian as a first language. Children were divided into dyads based on whether they had the same sex, a similar age and a similar availability. Prior to the experiment the children already knew each other. In Table 3 characteristics of the dyads are described into more detail. The dyads in Noord-Holland were part of the control group (n_{dyads}= 4) (L1-L1) and the dyads in Frisia formed the experimental group (n_{dyads}=3) (L1-L2). The age of the children in the experiment group was between 7 and 8 years (M_{age}=8.0, SD_{age}= 0.30). Within the experiment group the age of the children was between 6 and 9 years (M_{age}=7.77, SD_{age}=1.26).

Table 3

Characteristics of dyads of children

Dyad	Age child L - child R (years)	Sex child L - child R	L1 child L - child R	Group
1	6.6 - 6.3	female - female	Dutch - Dutch	Control
2	6.6 - 7.4	male - male	Dutch - Dutch	Control
3	9.3 - 9.4	female - female	Dutch - Dutch	Control
4	8.8 - 7.8	male - male	Dutch - Dutch	Control
5	8.3 - 7.6	female - female	Frisian - Dutch	Experiment
6	8.1 - 7.9	male - male	Frisian - Dutch	Experiment
7	7.6 - 8.3	female - female	Frisian - Dutch	Experiment

Materials

For this study we made use of an adapted version of the This-or-That debate phrases protocol (Appendix B), which was developed in the previous study. This protocol entails a list of 12 funny and imaginative statements (Figure 2). The protocol will be discussed in detail in the procedure.

We made recordings of both the speech and the hand gestures of the participants. To record their gestures we made use of a GoPro Hero 9, placed on a table in front of the participants. To record their speech we made use of two wearable TasCam DR-10L microphones which were attached to the cheeks of the participants with a sticking plaster for children (with prints of dragons/princesses). During the experiment both children were seated on a chair that was placed diagonally towards the camera, either left or right shifted. In this way, the children were able to face each other while both of their hands were also visible in the camera frame. The distance between the participants and the camera was long enough to capture at least the full length of their torso. Additionally, a live connection was set up using Google Meets on a laptop placed on the table next to the GoPro. This allowed the experimenter to observe the children and enabled communication with the children without their presence in the room. On the laptop screen no signs of a live connection were visible, instead the children were shown a timer running from 10:00 minutes to 00:00 minutes. This was done to prevent them from being distracted by the fact that

they were being watched. The audio of the video call on the laptop was connected to a JBL 3 speaker. A second laptop was used to share the statements of the experiment with the children. They were able to hear the voice of the experimenter through the JBL 3 speaker, using the video call. The set-up of the experiment is also visible in Figure 3.

To analyse the recorded material we made use of the software ELAN 6.4. This software is used to make annotations for audio and/or video recordings. We also used an adapted version of Table 1, that was developed within the procedure of the previous study. Lastly, we used Microsoft Excel to calculate means, standard deviations and create scatterplots and trendlines.

Figure 2

Statements of the This-or-That debate phrases protocol

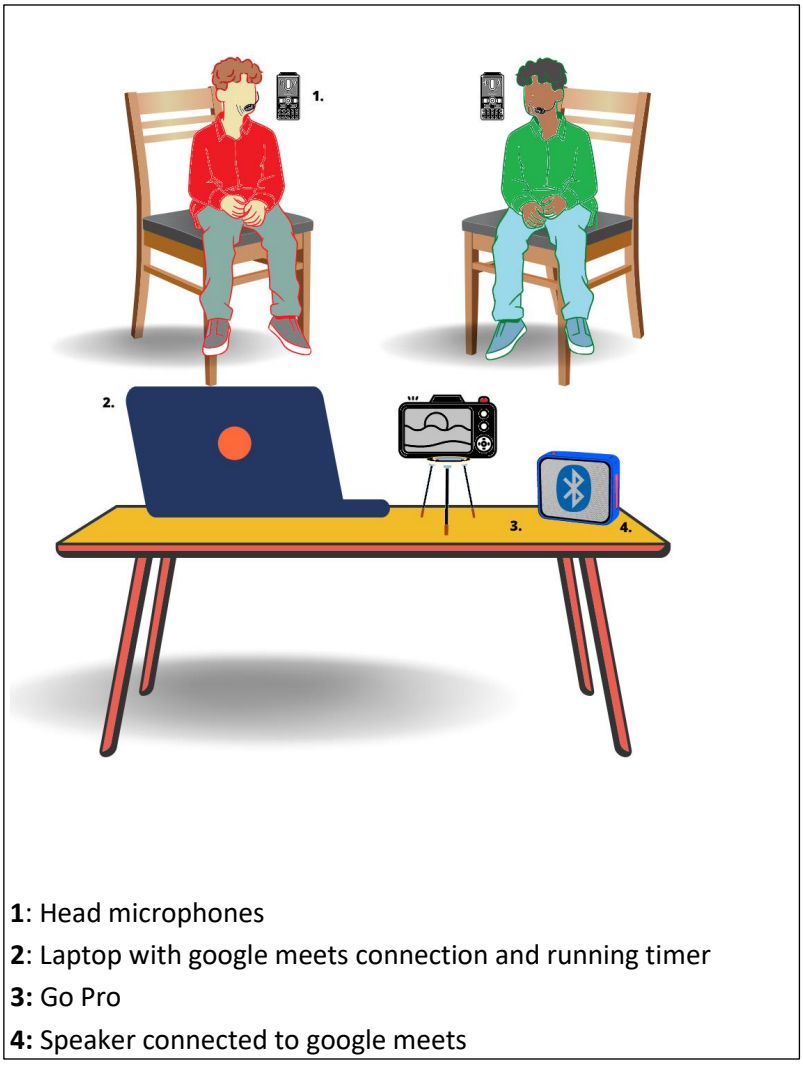
This-or-That Debate phrases

1. Zouden jullie liever een koning of een superheld willen zijn?
   Would you rather be a king or a superhero?
2. Zouden jullie liever zo klein als een kabouter of zo groot als een reus willen zijn?
   Would you rather be as small as a gnome or as big as a giant?
3. Zouden jullie liever vier handen of vier benen willen hebben?
   Would you rather have four hands or four legs?
4. Zouden jullie liever kunnen vliegen of onzichtbaar kunnen zijn?
   Would you rather be able to fly or to be invisible?
5. Zouden jullie liever reusachtige handen of reusachtige voeten willen hebben?
   Would you rather have gigantic hands or gigantic feet?
6. Wat zou beter zijn, nooit meer kunnen ruiken of nooit meer kunnen proeven?
   What would be better, never being able to smell or never being able to taste?
7. Wat zou beter zijn, de grappigste of de slimste persoon op de wereld zijn?
   What would be better, to be the funniest or the smartest person in the world?
8. Wat zou beter zijn, in de zee of in de ruimte wonen?
   What would be better, to live in the sea or in space?
9. Zouden jullie liever kunnen toveren of de toekomst voorspellen?
   Would you rather be able to do magic or to forecast the future?
10. Zouden jullie liever kunnen tijdreizen of de tijd stil kunnen zetten?
    Would you rather be able to travel in time or to freeze the time?
11. Zouden jullie liever met dieren kunnen praten of elkaars gedachten kunnen lezen?
    Would you rather be able to talk with animals or be able to read each other's mind
12. Wat was er als eerste; De kip of het ei?
    What excisted first; The chicken or the egg?

Bonus: Wie denken jullie dat er gewonnen heeft (Jullie of de andere groepjes)?
Who do you think won the game (You or the other groups)?

Figure 3

Set-up of the experiment This-or-That debate phrases protocol



1: Head microphones
2: Laptop with google meets connection and running timer
3: Go Pro
4: Speaker connected to google meets

Procedure

The experiments for the control group were conducted in the living room of one of the participants. The two children were welcomed together with at least one of their legal guardians accompanying them. Before we conducted the experiment, the legal guardian of each of the participants completed a consent form. The age of the child and their relation to the other child in the dyad was registered. Afterwards the children were left with the experimenter and the mother who lived in the house where the experiment took place. The experiments for the experiment group were conducted in an office within the school of the children. The legal guardian signed a form beforehand and the children were taken out of the classroom by the experimenter. The classroom teacher registered their age.

At the start of the experiment the two children were asked to sit down in the chairs. Before

starting with the actual experiment the experimenter introduced herself and asked the children to do the same. Afterwards the experiment was explained to the children. Firstly they were told that they were going to play a game together. The presence of the microphone and the GoPro was shortly explained and included as a tool in the game 'to check if they played it by the rules'. After the microphones were attached they were told to further ignore the equipment and not to touch it. Afterwards, the game itself was explained. The experimenter told the children that they had to work together to pass as many statements as possible, to beat the other teams. Emphasis was put on the fact that the children needed to agree with each other and that they both needed to provide their own reason for each common choice. This was done in order to stimulate them to talk frequently. This format was practiced with two example statements. The experimenter was present in the room during the first example and provided the second example via the speaker, using the google meet connection. In this way the children had a chance to get accustomed to the set-up. Furthermore they were instructed to keep seated on their chairs. This was done to ensure that their torso endured in the camera frame of the GoPro. Finally, they were told that they had 10 minutes to play the game. The experimenter emphasized that they needed to be as fast as possible in order to win a surprise. This was also done in order to stimulate ongoing conversation and enthusiasm of the children.

During the experiment the experimenter was sitting outside the room to read out the statements to the children. When the children both provided a reason that was related to the statement at stake, the next statement was given. Additionally, the experimenter had the opportunity to correct or stimulate the children when needed. For example, when they put their legs on the chair (*'Can you put down your leg?'*) or when they forgot some details of the instructions ('*Now I still need a reason from [child L]*'). When all statements were provided, the experimenter returned to the room to close off the experiment. After the experiment, all children were rewarded with a small bubble blower or a bouncy ball. An example of the explanation of the experiment and some examples of intervening statements are included in the protocol (Appendix B).

Analyses

Whereas the goal of this pilot is to investigate whether the task in the protocol is suitable to use in the context of interpersonal alignment we will firstly closely define when we would deem the protocol to be successful. In order to do this, we elaborated the observation points that we defined in the pre-pilot (See Table 1). Part of these points were quantified (this will be further explained below). In this way, the findings can also be used to assess differences between the separate statements of the protocol as well as individual and group differences. These could be differences related to the age or the first language of the children.

Firstly, the This-or-That debate phrases task would be defined as successful when all dyads were at least able to perform the task. Furthermore, the average conversation time used by all dyads for each statement as well as the average minutes used for each statement in each dyad should at least be half-a-minute. We chose this limit since the total amount of conversation time of the children should ideally be between 8 and 10 minutes. A longer conversation time would namely provide more data to draw conclusions on interpersonal alignment. Earlier studies have shown that some patterns of interpersonal coordination only occurred after a period of, for example, 20 to 30 seconds (Jaffe et al., 2001 as cited De Jonge-Hoekstra, 2021). In addition, it has been suggested that the longer people speak, the more gestures they produce (Nagpal et al. 2011).

With regards to the deviation in speaking time within the dyads, we decided that all of the children should express between 40-60% of all utterances that are expressed within the experiment to ensure there was a balanced share in the conversation. Again, this would aid the analysis of interpersonal synchronization in temporal alignment. In addition, the amount of stimulations made by the experimenter should not be higher than 10% of all utterances of the children and the experimenter together. A higher rate would suggest that the experimenter interrupted frequently in the conversation, which in turn suggests that the children did not spontaneously engage on their own. To evaluate the age appropriateness of the task we will look at the average age of each dyad compared to the time they used to complete the statements as well as the number of utterances they used and the number of simulations they needed. Further, the number of stimulations made by the experimenter could also indicate that the children frequently stopped talking. When the children did use more time to

complete the statements, in combination with less utterances, this could be a sign that they did not understand the task and, instead, were silent. In order to analyse this, the dyads were ordered from youngest to oldest.

In order to evaluate use of non-verbal communication the hand gestures will be counted for each statement. In a previous study with 6-10 year old children that were asked to reproduce a cartoon, the average number of gestures that children used per clause was about 0.30 for 6-year-olds and 0.50 for 10-year-olds (Colletta et al., 2010). In a later study with 4-10 year old children, the children used an average of 2.30 gestures per 100 words when telling a story (Nicoladis et al., 2016). In addition to this, it is also known that the amount of hand gestures made by individuals differs due to several factors such as their personality and cognitive and perceptual skills (Özer & Göksun, 2020). In line with this research, one could expect that within the course of the discussion time in each statement, at least one gesture would be produced. Therefore, we decided that when a statement did not elicit any hand gestures in more than one dyad, it would not be considered as a suitable means to stimulate the use of hand gestures. Lastly, when the total number of hand gestures for one statement (used by the entire sample of children) was further than one standard deviation away from the average number of hand gestures of all children in one statement, it was also considered as not suitable.

Coding

In order to study all the aspects of the experiment that we mentioned in the previous section we operationalized the observation points of the pre-pilot in a quantitative way. We did this with the help of the annotation software ELAN 6.4. The only observation points that have not been operationalized in a numerical way are age appropriateness and complexity of the instruction, since these are a matter of observation. In Table 4 the method of operationalization for each observation point for the pilot is given.

In ELAN 6.4 we created tiers to annotate what was said and what happened during each experiment. As visible in Table 5, we coded the following behaviours (tiers): Questions (statements) asked by the experimenter, answers given by the left (L) or the right (R) child, task unrelated utterances made by the left or the right child, instruction or stimulation provided by the experimenter, any needed remarks and finally the duration of each episode (one episode per statement). The frequency and timing of each tier was automatically counted and directly visible within ELAN 6.4.

Part of the coded data from ELAN 6.4 was used to derive at conclusions about the observation points we mentioned earlier (See Table 6). We used the tiers of the duration of each episode to calculate the duration of the conversation for each statement and for the entire conversation. We also used the moments that the experimenter, child L and child R were talking (L on task, R on task, stimulations) to calculate the number of the utterances per child and the share of the stimulations provided by the experimenter relative to this number. In order to analyse age appropriateness we compared the number of utterances of the children, the numbers of stimulations given by the experimenter and the total conversation time to the average age of the participants in the dyads using Excel. The data that we derived from ELAN is visible in Appendix C.

Finally, hand gestures were analysed by manual observations. The experimenter coded every time the child made a hand gesture with the left or the right arm or hand. Movements that were not coded as hand gestures were moments where the child was fidgeting with their hands, a bracelet or other object or touching their face or the microphone. All the coded raw data was transferred to Microsoft Excel in order to calculate mean values, standard deviations and to create scatterplots.

Table 4

Observation points and operationalization

Observation points	Coding method	Operationalization
1 Complexity of the instruction	Observation	Amount of dyads able to perform the task
2 Length of the conversation	ELAN	Total minutes spent discussing all statements, as coded in ELAN
3 Share of each child in the conversation	ELAN	Total numbers of utterances for each child, as coded in ELAN
4 Enthusiasm of the children	ELAN	Total number of stimulating phrases provided by the experimenter, as coded in ELAN
5 Age appropriateness of the task	Manual/Excel	Average age of the dyad compared to: - Total number minutes spent discussing all statements (ELAN) - Total number of stimulating phrases provided by the experimenter (ELAN) - Total numbers of utterances for each child (ELAN)
6 Amount of hand movements the children used	Manual	Counted number of hand gesture of each child

Table 5

Annotation scheme Elan

Name of tier	Subject/Attribute	Definition
Experiment Question	Experimenter	Any time period during which the experimenter reads out one of the twelve statements/the bonus question
L on Task	Left child	Any time period the left child is saying something in relation to the task
L off task	Left child	Any time period the left child is saying something that is not related to the task
R on task	Right child	Any time period the right child is saying something in relation to the task
R off task	Right child	Any time period the left child is saying something that is not related to the task
Instruction	Experimenter	Any time the experimenter provides feedback or instruction to the children
Stimulation	Experimenter	Any time the experimenter says something to encourage one of the children to answer/elaborate
Interruption	Any interfering person	Any time another person than the children or the experimenter said something
Remark	All	Any annotation of something remarkable that happened during the experiment (e.g. child moving a lot)
Episode	Statements	Any time period of the dialogue on one of the twelve statements (from the experimenters question until the end of the last on task annotation of the left or the right child)

Table 6

Descriptions of observation points derived from ELAN

Name	Description	Tiers from ELAN	Subject	How
Length of the statements	Total time (seconds/minutes) spent discussing one statement	Episodes	Experimenter, Child L, Child R	Abstracting the end time of the last utterance from the moment that the experimenter reads out the statement (starting time)
Length of the total conversation	Total minutes spent on all statements	Episodes	Experimenter, Child L, Child R	Adding up the conversation time of all 12 statements
Number of utterances	Total number of times the subject says something without pausing	R on task L on task	Child L, Child R	Counting the total number of times the subject said something without pausing
Number of stimulations	Total number of times the experimenter gave an encouraging comment to one or both children	Stimulations	Experimenter	Counting the total comments that were coded as a stimulation

Results

In this study we investigated if the This-or-That debate phrases protocol could be used to research interpersonal temporal alignment of hand gestures and speech in dyads of school aged children speaking in a first or second language. All 7 dyads, both in the L1-L1 and L1-L2 condition, did understand the task well enough to participate. All dyads together spent an average of 8.44 minutes discussing all 12 statements (SD=1.30, $M_{L1\text{-}L1*}$=8.62, $M_{L1\text{-}L2}$ =8.19). The longest total duration was 11.02 minutes and the shortest duration was 6.78 minutes. For each statement separately, the children spent an average of 42.18 seconds (SD=6.48, $M_{L1\text{-}L1}$=43.08, $M_{L1\text{-}L2}$ =40.97*)*. The dyad with the longest discussion spent an average of 55.08 seconds on each statement and the dyad with the shortest discussion 33.92 seconds. Furthermore, all separate statements had an average duration of 30 seconds discussion time or more, except for statement 1 (Would you rather be a superhero or a king?). Statement 1 had an average duration of 26.29 seconds. Statement 12 (What came first, the chicken or the egg?), had the longest duration of 71.00 seconds (1 minute and 11 seconds). In Table 1 in Appendix C, the specific duration of each statement is represented.

* When we write $M_{L1\text{-}L1}$ or $_{L2\text{-}L2}$ we refer to within group statistics
** When we write M_{L1} or $_{L2}$ we refer to the mean of all L1 or L2 language individuals in the entire population

The average number of utterances expressed by each child was 62.21 (*SD*=13.98, M_{L1**}=60.18, M_{L2} =69.67*)*. The highest number of utterances expressed by one child was 93 and the lowest number 42. The average share in the conversation for each child was 50.00% (*SD*= 4.62%, $M_{L1\text{-}L1}$ =50.00%, $M_{L2\text{-}L2}$ =0.00%). Within the dyad with the least balanced conversation, the division between the two children was 42.15% (Child L) compared to 57.85% (Child R) of all utterances. In Table 2 in Appendix C the total number of utterances of each child and the balance in conversation for all dyads is visible. The average amount of stimulations provided by the experimenter formed 7.86% of the total utterances expressed by the experimenter and both children together (*SD*=4.83%, $M_{L1\text{-}L1}$=6.71%, $M_{L2\text{-}L2}$=5.45%). The highest share of the experimenter's stimulations was 16.36% of all utterances. In all other dyads their share was below 10.00%, with 4.26% of all utterances as the lowest percentage. In Table 3 in Appendix C the number of stimulations and the total share of the experimenter within the conversation is visible.

The average number of hand gestures of one child within the total conversation was 28.36 hand gestures (*SD*=13.18, M_{L1}=24.73, M_{L2}=41.67*)*. The child with the lowest frequency made a total of 6 hand gestures and the child with the highest frequency 59. For the separate statements, the average number of hand gestures made by each child was 2.36 gestures per statement (*SD*=2.42, M_{L1}=2.24, M_{L2}=3.47). Finally, the average amount of hand gestures made per utterance of each child was 0.44 *(SD*=0.15, M_{L1}=0.40, M_{L2}=0.62). A full overview of all the hand gestures made by the children is given in Table 4 in Appendix C.

The average number of hand gestures used for a statement by all dyads together (total hand gestures of all children together) was 33.08 (*SD*=12.16). Statement 4 (Would you rather be able to fly, or be invisible?) elicited the least amount of total hand gestures *(n*=21). Statement 12 (What came first, the chicken or the egg?), elicited the highest total number of hand gestures (*n*=65). Eleven out of twelve statements elicited at least one hand gesture in 6 or 7 out of 7 dyads. Statement 10 (Would you rather be able to travel in time or to predict the future?) did only elicit hand gestures in 5 out of 7 dyads . Those five dyads together produced a total of 29 hand gestures. None of the statements had a total number of used hand gestures that was more than one standard deviation below the mean (12.16 hand gestures below 33.08).

In the scatterplot in Figure 4 one could see an increasing trendline when comparing the total amount of hand gestures used by both children of the dyad to the average age of each dyad (β=16,60). Most of the dyads are fitting this pattern and are spread along the trendline. The youngest dyad used the highest amount of hand gestures and forms an outlier. Figure 5 shows an increasing trendline for the amount of hand gestures of each dyad relative to the amount of utterances made by these children, compared to their average age (β=0.03). The data points are moderately scattered around the trendline. The datapoints of the youngest and oldest dyad are not consistent with the trendline.

The scatter plot in Figure 6 compares the total conversation time of each dyad with the average age of the dyads and shows a slightly decreasing trendline (β= -0.95). The younger dyads generally spent more time finishing all statements than the older dyads. The datapoints are all scattered quite near the trendline. The dyad with the average age of 7 has the highest value and is inconsistent with the trendline. The total number of stimulations provided by the experimenter to the children, compared to their average age, also showed a decreasing trendline (Figure 7)(β=-2.70). The datapoints are spread relatively far from the trendline. The youngest dyad is inconsistent with the trendline.

The amount of utterances expressed by both children together, compared to their average age, showed an increasing trendline (Figure 8)(β=11,10). The datapoints are all close to the trendline but are showing some inconsistencies with the trend. Finally, the amount of utterances relative to the total conversation time for each dyad (number of utterances per minute) compared to their average age showed an increasing trendline (Figure 9) (β=2.91). The datapoints are scattered around the trendline in a relatively consistent manner. The datapoints of the youngest dyad and one of the dyads around the age of 8 are a bit higher than expected.

Figure 4

Total number of hand gestures compared to age

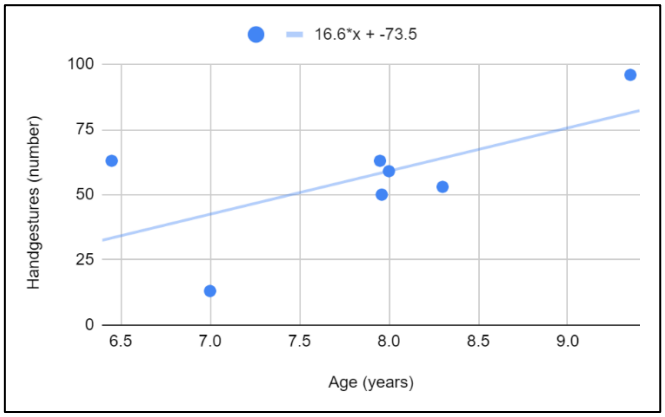


Figure 5

Hand gestures per utterance compared to age

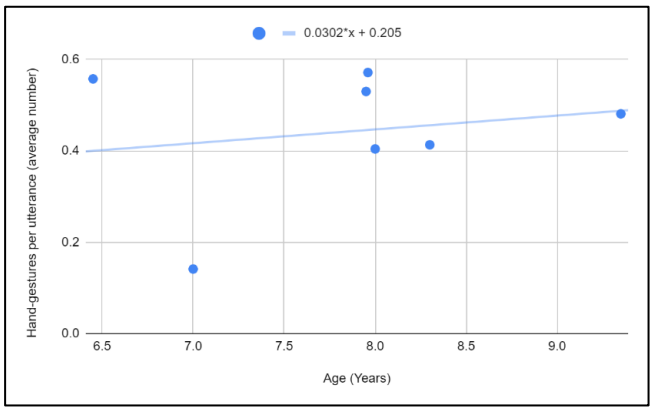


Figure 6

Total discussion time of all statements compared to age

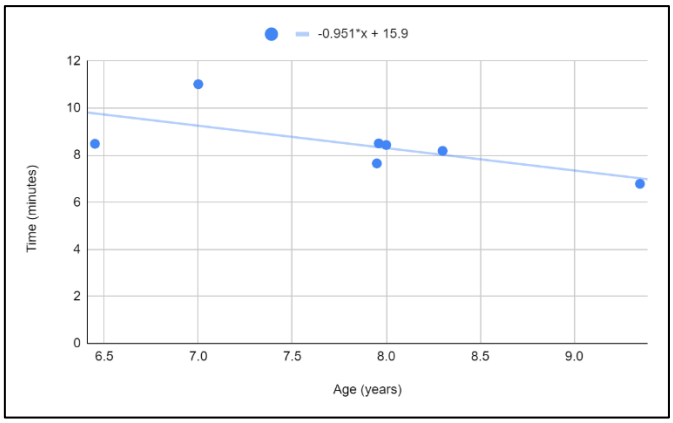


Figure 7

Stimulations given by experimenter compared to age

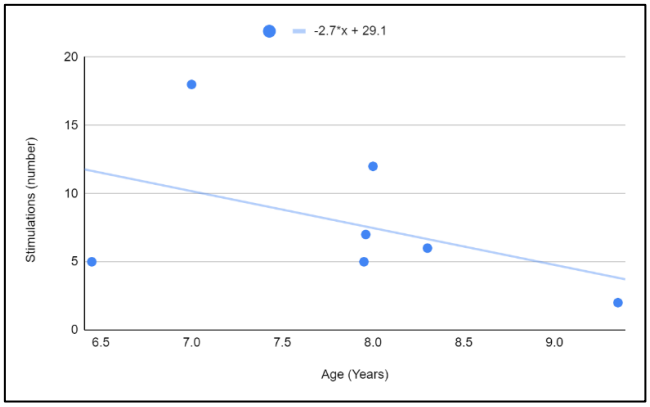


Figure 8

Number of utterances compared to age

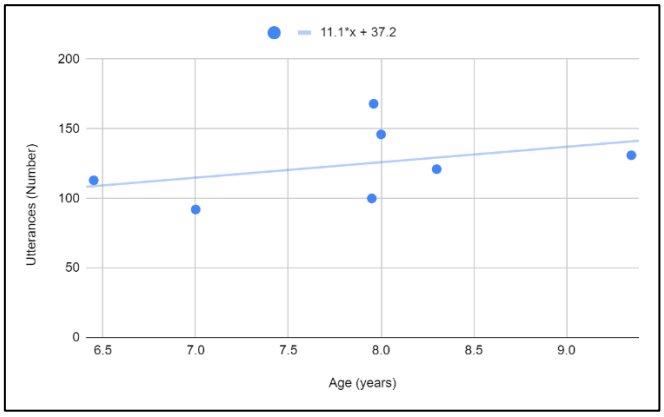
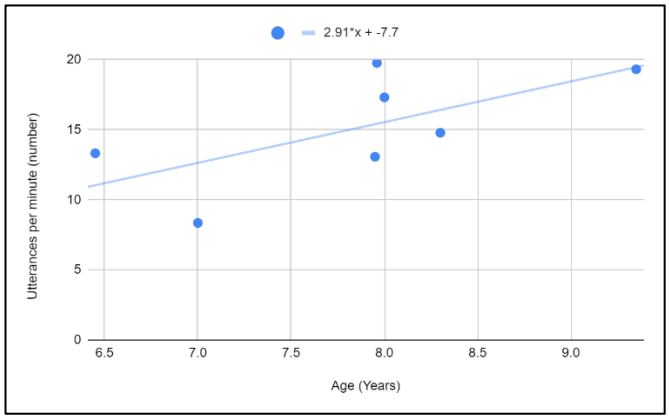


Figure 9

Utterances per minute compared to age



Discussion

The first goal of this study was to explore whether the This-or-That debate phrases task is a suitable task to elicit spontaneous communication and hand gestures in children. Four dyads of Dutch children between 6 and 9 years old as well as three Dutch-Frisian dyads successfully completed the experiment protocol. They all expressed both verbal and non-verbal communication within a dialogue. The length of the total conversation, the share of each child within the conversation, the enthusiasm of the children to engage and finally the children's use of hand gestures all met the requirements set by the experimenters. This means that the results of the previous study, that suggest that the This-or-That debate phrases protocol is a suitable task to elicit spontaneous and non-verbal communication in primary school aged children, are supported by this study. Concluding that the task can be used in both L1 and L2 children with an age between 6 and 9.

Yet, while pertaining to the second goal of this study we discovered some specific differences between the statements. Statement 1 (Would you rather be a superhero or a king?) had an average duration that was shorter than the prerequisite 30 seconds. Therefore one could suggest replacing this statement with a statement that is more similar to a statement that elicited a longer response. However, since the order of the statements was the same for each dyad, the fact that this statement occurred at the beginning of the experiment task could be an explanation for its shorter duration. Statement 10 (Would you rather be able to travel in time or to predict the future?) did not elicit any hand gestures in 2 of the 7 dyads. Yet, the total number of hand gestures that was elicited within the other 5 dyads was relatively high. Therefore, the absence of hand gestures in dyad L2.1 and dyad L2.2 might be coincidental. In the video material of video L2.1 we can see that during this statement Child L was fidgeting with her shirt and in video L2.2 we can hear the children from other groups returning from their break time. Both factors might have distracted the children from thorough engagement with this statement. Statement 12 (What came first, the chicken or the egg?) both produced the longest duration and the highest total amount of hand gestures. An explanation for this could be the fact that the content of this statement can be perceived as the most difficult. The difficulty of a task has been related to different patterns in speech and gesture use in earlier studies (see Goldin-Meadow, 2003; Church & Goldin-Meadow, 1986: De

Jonge-Hoekstra et al., 2016). If this protocol would be used in the future, it could be recommended to use statements with a similar level of difficulty and abstractness as statement 12.

With regards to individual differences and the (within) group differences we discovered a couple of differences. The amount of hand gestures made within the dyads was highly variable. First of all, there was a pattern between an increasing average age of the dyad and the amount of hand gestures used. This pattern was still recognizable when the amount of hand gestures was compared to the amount of utterances that the children expressed, although this time it was less apparent. This finding is coherent with earlier data that shows that older children use more hand gestures than younger children (Colletta et al., 2010). We also discovered that within the L2 group the average amount of hand gestures within the course of the experiment was greater than that of the L1 group. Again, this difference was still present when comparing it to the amount of utterances the children made. Although these findings could be a result of the large difference in sample size (n_{L1} = 11, n_{L2} =3), they might relate to earlier studies in which non-native participants produced a higher amount of hand gestures. Within these studies L2 speakers did more frequently substitute words they did not know for hand gestures, compared to native language speaking participants (Gullberg, 2011 ; Kosmala et al., 2019). On the other hand, Bosma and Blom (2019) found that Frisian children have less difficulty to switch from Frisian to Dutch than the other way around. This finding is not in line with the theory that the higher amount of gestures of the L2 children is related to translations problems. Lastly, the higher frequency of hand gestures used in the Frisian children could possibly be a result of a general higher use of hand gestures in Frisian language. Yet except for the studies by Bosma and Blom (2020) and Nota et al (2016) on inflection rules and intonation, no studies on the use of hand gestures in Frisian have been found in the literature search. With regards to future research, a recommendation would therefore be to further investigate the frequency of hand gesture use in Frisian.

There were a few shortcomings within the design and execution of this study. In both the locations for the data collection of the L1-L1 and L2-L2 condition, there were several distracting factors present in the environment. Within the L1-L1 condition, one of the mothers was present during the experiment. She also interfered within the experiment, by emphasizing parts of the explanation or answering to remarks of the children during the experiment. Within the school setting, the children in

dyad L2.2 and L2.3 were participating in the experiment during the playtime break. It could therefore be that they were less concentrated than they normally would be. Though, these circumstances are naturally present in children's environment so they might add ecological validity to the results as well (Roberts, 2014). Another problem we encountered is that the children were moving more than we foresaw. The protocol was designed in a way that the children were not holding anything in their hands, to enable them to use their hands freely. However, several children were fidgeting with their bracelets or clothing or taking up their legs on their chairs. In addition they were touching the microphone on their cheek quite often. This could have kept them from using their hands more often while talking. When this protocol is used in the future, any loose objects that the children bring into the experiment could be removed beforehand to limit fidgeting behaviour. However, this would in turn lower the ecological validity of the setup. An additional solution could be letting the children play an introduction game before engaging in the actual task, whereas this might reduce their nervousness and therefore decrease their fidgeting behaviour.

Lastly, we encountered some methodological issues during the analysis. The extensive movements of the children made it harder to decide whether to code their behaviour as a hand gesture or not. Furthermore, we did not define a standard measurement to code the utterances of the children. In practice it could therefore be that two phrases that were coded as one utterance for one child would be coded as two utterances in another child. Therefore, some of the results of this pilot study should be interpreted with caution.

Study 3: Pilot using OpenPose

In the third study we investigated if the software Openpose could be used to study video data of the This-or-That debate phrases task to analyse interpersonal temporal alignment of hand gestures and speech in school-aged children speaking in a first or second language. In order to do this we used the video data of one L1-L1 dyad and one L1-L2 dyad from the pilot of the second study. We processed the video data by using Openpose. The processed data from Openpose was used to investigate differences in temporal alignment of gestures and speech within and between the two dyads.

Participants

We used data of the first dyad from the control and experimental group. The control dyad includes two girls with Dutch as their L1. Their ages are 6.3 and 6.6 years (M_{age}=6.45). The first dyad from the experiment group includes one girl with Frisian as L1 (8.3 years) and Dutch as L2 and another girl with Dutch as L1 (7.6 years) (M_{age}= 7.95).

Material

To analyse the video data (recorded with a GoPro Hero 9) we made use of the software ffmpeg, ImageJ, OpenPose and R (Matahelumual, 2022; Cao et al., 2019). We used ffmpeg to extract frames from the video with a frequency of 1/5 Hz (so 1 image per 5 seconds). ImageJ was used to indicate the location of each child in these videoframes. OpenPose is an open source software that enables users to track a wide range of body key points by either analysing 2D video recordings or live footage (Cao et al., 2019). OpenPose can detect 25 different points within the body. For this study we used the datapoints detecting the child's left and right wrists because we are interested in the peaks in movements of the arms only. We used a custom R-script to combine the ImageJ data and OpenPose data and thereby extract the time series of the participants' hand movements from the OpenPose data. To analyse the audio data (recorded with TasCam DR-10L microphones) we used Adobe Audition. Furthermore, we used an R-script by Pouw et al (2020) to extract the amplitude envelope of the speech signal. Finally, the speech and hand movement data was combined using a custom R-script after which the programme was also used to visualize it into graphs and tables.

Procedure

Data preparation

In order to analyse the applicability of the material for OpenPose, the data first had to be prepared. This procedure was done in a similar way as the procedure in the thesis of Matahelumual (2022). Firstly, using ffmpeg, the videos of the experiment were saved with a frequency of 1/5 Hz (so 1 image per 5 seconds). Secondly, the video data was analysed in OpenPose by tracking the 25 body keypoints for both children. For this study, we then selected the data of the wrist keypoints. Thirdly, ImageJ was used to register the location of each child in each frame in order for the custom R-script to correctly identify the children in the OpenPose data. Specially we clicked on the nose of each child in each videoframe. Fourthly, the audio data was edited in Adobe Audition to remove the background noise in the recordings. Fifthly, Using a R-script by Pouw et al (2020), we extracted the amplitude envelope of the speech signal for both participants. Subsequently, with a custom R-script we combined the ImageJ data and OpenPose data to correctly identify the children for each OpenPose data point and calculated wrist movements for both children. The final dataset of speech and hand movements had a sample rate of 100 Hz (1 datapoint per 10 ms). Finally, with the use of a custom R-script, the amplitude envelopes of speech and the movement time series of the wrists were combined per dyad, and per child. We used findpeaks-function from the pracma-package to automatically find peaks in the amplitude of speech and the amplitude of wrist movements. For speech, the minimal peak height was the mean of the whole timeseries of speech, and the minimal distance between peaks was 100 ms. For gestures, the minimal peak height was the mean of the whole timeseries of the wrist + 0.5 x the standard deviation of the all the wrist movements, and the minimal distance between peaks was 200 ms. Speech and gesture peaks that did not have any overlap were excluded from the analysis, because we considered these independent peaks as nonrelevant for temporal alignment concerning speech and gesture.

This complete procedure resulted in a graph with three different lines for each statement and corresponding dialogue of the This-or-That phrases protocol per child. These graphs represent amplitudes of the speech (red lines), velocity of the left hand (green lines) and velocity of the right hand (blue lines). The dots in the graphs resemble the peaks in speech (red) and the peaks in gestures of the left hand (green dots) and the right hand (blue dots). In figure 10 and 11 an example of the amplitudes

of the first episode of the children in the experiment group is visible. The other graphs are included in Appendix D.

Figure 10

Amplitudes of the speech, left hand & right hand of the left child in Task episode 1

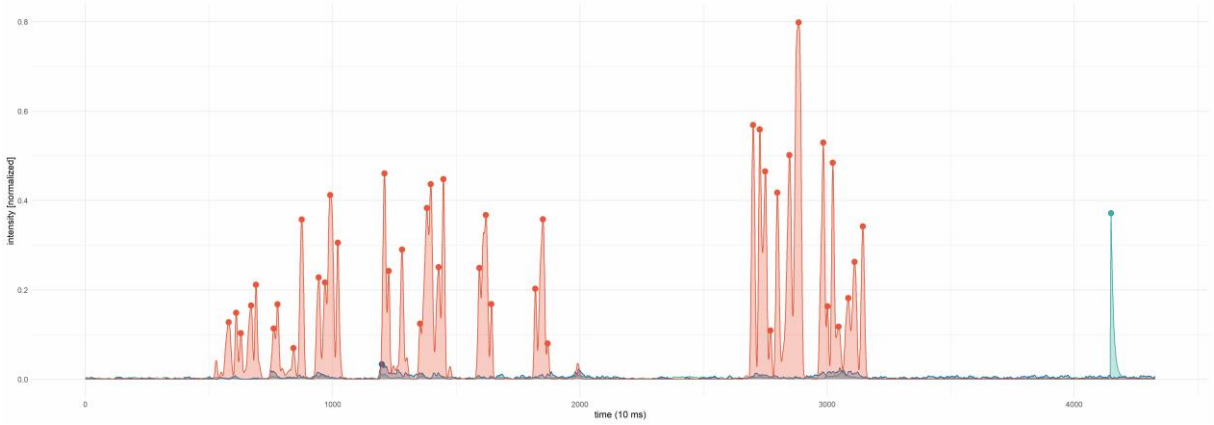
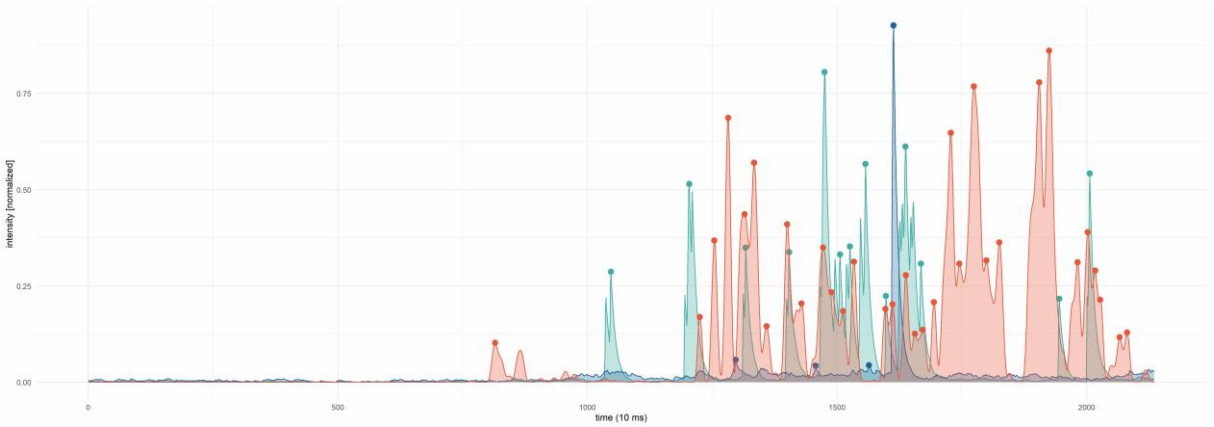


Figure 11

Amplitudes of the speech, left hand & right hand of the left child in Task episode 1



Quantitative Analysis

Differences in temporal alignment of speech and gestures were analysed using peaks within the different timeseries. When a peak in speech was close to a peak in hand gestures, we could analyse it as alignment by counting the amount of milliseconds between the peaks of speech (red dots) and gestures (green and blue dots). When there is a smaller difference between the peaks of speech and gestures, there is more temporal alignment within one child. When the peaks of the left (green dots) and right wrist (blue dots) were 100 ms (milliseconds) or even closer to each other, they were counted as one

combined gesture peak. Differences between peaks that were greater than 1500 ms were considered too large and instead they were viewed as separate peaks in speech or gesture. By comparing the average absolute time difference (ms) between the aligning peaks for both children in a dyad, we were able to quantify differences in the temporal alignment within and between the participants. Finally, the within dyad differences of the experiment conditions (L1-L2) were compared to the control condition dyads (L1-L1) to detect interpersonal differences in alignment between the two conditions.

Results

In this study we investigated if differences in interpersonal temporal alignment between dyads with the same or a different L1 could be detected using the software OpenPose. The difference in timing of the children's peaks in hand gestures compared to the peaks in their speech varied from 0 ms to 1100 ms in all children (*M* = 122.75, *SD* = 154.80). In Table 7 and 8 one can observe the average peak differences per child per episode a well as the average of these average differences (the total average) within each dyad.

For the control dyad (L1-L1), OpenPose detected 71 gestures in the left child and 108 gestures in the right child (see Table 9). Within this dyad the differences in speech-gesture peaks ranged from 0 ms to 1010 ms (*M* = 125.88, *SD* = 151.28) within the right child and from 0 ms to 1100 ms (*M* = 144.79, *SD* = 206.21) within the left child. The absolute time difference between the averages of the two children in the control group is 18.91 ms. In Table 7 the speech-gesture peak differences between the children are represented for each Task episode. The average absolute time difference between the children within an episode is 132.73 ms *(Min*= 7.32, *Max*= 316.67, *SD*=151.60).

For the experiment dyad (L1-L2) there were 66 gestures registered for the left child (Frisian native) and 17 for the right child (see Table 9). No gestures were registered for the right child in episode 1,2,3 and 6. Within episode 6 and 10 no gestures were registered for both the left and the right child. Within this dyad the peaks differences ranged from 0 ms to 230 ms (*M* = 104.13, *SD* = 75.92) in the right child and from 0 to 570 ms (*M* = 95.61, *SD* = 99.59) in the left child. The absolute time difference between the averages of the two children in the experiment group is 8.51 ms. In Table 8 the speech-gesture peak differences between the children are represented per Task episode. The average absolute time difference between the children within an episode is 28.93 ms (*Min*=0.00, *Max*= 75, *SD*=29.64*)*.

When comparing the average absolute time difference of the gesture-speech peak differences between the children in the experiment dyad and the control dyad (within-group differences), the between group difference is 5.38 ms. The between group time difference is 103.80 ms when comparing the average of the within-group differences per episode (see Table 7 & 8)

Table 7

Average peak gesture difference per episode per child for the control dyad (L1-L1)

Task episode	Left child (ms)	Right child (ms)	Absolute time difference (ms)
1	210	450	240
2	86.25	93.57	7.32
3	147.5	120	27.50
4	223.75	150	73.75
5	127.5	116.88	10.63
6	110	82	28
7	230	168.26	61.74
8	530	40	490
9	310	90	220
10	58.33	375	316.67
11	118.33	62.5	55.83
12	73.33	12	61.33
Average	**185.42**	**146.68**	**132.73**
Standard deviation	**132.06**	**132.47**	**151.60**

Table 8

Average peak gesture difference per episode per child for the experiment dyad (L1-L2)

Episode	Left child (L2) (Ms)	Right child (L1) (Ms)	Difference (ms)
1	110	-	-
2	58.33	-	-
3	480	-	-
4	140	100	40
5	116	110	6
6	-	-	-
7	80.91	35	45.91
8	85	-	-
9	120	113.33	6.67
10	-	-	-
11	0	0	0
12	60	135	75
Average	**125.02**	**82.22**	**28.93**
Standard deviation	**130.99**	**52.60**	**29.64**

Table 9

Total number of gestures registered in OpenPose

	Left Child	Right Child
Control group (L1-L1)	71	108
Experiment group (L2-L1)	66	17

Discussion

Within this pilot study we attempted to use OpenPose to analyse intertemporal alignment of gesture and speech within and between native and non-native children. The results suggest that the children in the experiment group (L1-L2) showed a higher degree of temporal alignment within themselves, as well as a higher degree of interpersonal temporal alignment than the children in the control group (L1-L1). This difference is present in both the average peak-gesture difference of the children in the entire experiment task as well as in the averages per task episode. In addition, the standard

deviations within the experiment dyad are also lower, indicating that the peak differences in this group are grouped closer together.

Yet, within the right child (Dutch) of the experiment group, hand gestures were only observed in OpenPose starting from the fourth episode. Meaning that no peak differences were registered in the first three episodes of the experiment for this child. The left child was using hand gestures within these episodes, so one could argue there was no interpersonal synchronisation between the children during these episodes. The same event happened in episode 6. An explanation for these differences could be the fact that OpenPose might register or not register (hand) movements of the children that occur naturally, but would normally not be recorded as gestures (e.g fidgeting with bracelets). In the general discussion we will elaborate on this issue.

Although it seems that OpenPose is appropriate software to detect differences between peaks in gesture and speech and consequently differences in temporal alignment between individuals, these results should be interpreted with care. In this pilot we only compared two dyads. The circumstances in which these dyads were tested differed, as well as their mean age. Therefore, this study should be replicated on a larger scale to find out whether the difference in synchronisation between the L1-L1 and L1-L2 dyad could be generalised to a broader population.

General discussion

The main goal of this thesis was to design and pilot an experimental task that can be used to investigate differences in temporal alignment of hand gestures and speech within and between dyads of children speaking in a first or second language. In order to do this we conducted three consecutive studies. In the first study, we explored different tasks to elicit hand gestures and spontaneous conversation in children. In the second study we used an improved version of one of the tasks from the first study to run a pilot using a sample of dyads of children speaking in their first or second language. Finally, in the third study we used the video and audio data of two dyads from the second study to analyse the intensity of hand movements with the open software OpenPose and the intensity of speech via the amplitude envelope in R. Within the following section, we will combine the results of the three studies and elaborate on limitations and recommendations for future research.

Altogether, the results of the three studies illustrate that the This-or-That debate phrases task could be well used to elicit spontaneous speech and gesture use in dyads of young children speaking in a first and second language (i.e L1-L1 or L1-L2 combinations). Moreover, when recorded, this task can provide video and audio data eligible for OpenPose. Where Matahelumual (2021) successfully used OpenPose to analyse differences in temporal alignment within individual adults, we now confirmed that it can also be used to study differences in interpersonal temporal alignment between dyads of children.

We discovered several individual differences in gesture use throughout the entire thesis, both in L1 and L2 speaking children. In summary we found that both the older children as well as the L2 children were using more gestures in addition to their speech compared to younger or L1 speakers. Furthermore, we observed that most children use more gestures when the context of their conversation is more abstract. Finally, within an analysis using OpenPose we discovered that the L1-L2 dyad showed a higher degree of interpersonal temporal alignment compared to the L1-L1 dyad. Although part of these findings are in line with earlier research (Özer & Göksun 2020 ; Robberts 2014 ; Kosmala et al., 2019) we encountered some internal inconsistencies and flaws within the results and design of our study that must be taken into account for the interpretation of our results with regards to future studies.

First of all, we observed an inconsistency between the manual observation we conducted in the second study and the results from OpenPose from the third study. The amount of hand gestures we

counted manually were inconsistent with the amount of hand gestures that were registered in OpenPose (see Table 9 in study 3 and Table 4 in Appendix C). Within the manual count we excluded movements that we interpreted as either fidgeting or touching/scratching oneself. OpenPose registered all movements above a threshold (the mean of the whole timeseries of the wrist + 0.5 x the standard deviation of all the wrist movements), which could explain why the amount of gestures it registered is higher than the manual observations for three out of four children. Yet, it must be noted that the fidgeting and scratching movements could still be synchronised within the dyads, as people tend to align in all types of behaviour (see Richardson et al., 2007). A solution to avoid the difference between manually and automatically registered hand gestures is to lower the threshold set in the R-script used in Openpose to ensure OpenPose registers any movement that the children make. With a detailed coding table, that depicts what would be seen as a hand gesture and what not, all potential gestures registered in OpenPose could be manually controlled and deleted afterwards if they are not a gesture. Very recently, Rohrer and colleagues (2023) presented a new labelling system named M3D to code gestures in three dimensions (form, prosodic and meaning dimension) instead of set categories. This system could potentially be used for more detailed coding in future studies in addition to our method of analysation.

Secondly, during the analysis with OpenPose we found that the R-script we used did not enable OpenPose to recognize the position of left and the right child in every videoframe. In order to solve this we combined the OpenPose data with ImageJ data to manually locate both children's noses in every frame. To avoid these kinds of time consuming manual corrections, more enhanced software could be used. Yet we believe that the fact that OpenPose is freely accessible makes it a very recommendable research tool, despite the minor practical issues we encountered in this study.

Besides the technical problems we encountered within the analysis there also were some limitations within the designs itself. In all three studies we used a small convenience sample opposed to a random sample. In addition, the subjects we used in the different studies also had different characteristics. Within the first study we only used L1 children (or did not have any knowledge about bilingualism). Within the second study we had a higher number of L1-L1 dyads and L1 children compared to L1-L2 dyads and L2 children, which could have influenced the comparison of means we performed. Finally, within the third study we only compared the first two dyads of both experiment

groups instead of all dyads. Lastly, the circumstances of the studies were not completely standardised, because the context (i.e. at school/home), age (i.e. six/nine years old) and procedure (i.e. interference by third persons/playtime break) differed per dyad.

Despite the limitations within this thesis we recommend future studies to build on our findings. The set-up of the This-or-That experiment protocol that we designed proved to be a promising tool to study spontaneous communication between primary school children and could potentially be used to answer future research questions. In addition we presented a new method to analyse interpersonal temporal alignment within (bilingual) conversation partners with the use of OpenPose which is easily accessible for all researchers. Lastly, in this study we found a higher amount of temporal alignment in gesture and speech in L1-L2 dyads compared to L1-L1. As this is a new finding that potentially holds new insight in the synchronisation of L1 and L2 (young) speech partners, we suggest that our set-up using the This-or-That debate phrases protocol and the analysis in OpenPose must be replicated in a larger sample. This will aid a more in depth exploration of the novel relation we found in this thesis. All in all we think that our research has been a valuable contribution to the field of synchronisation in verbal and non-verbal communication.

References

Bosma, E., & Blom, E. (2019). A code-switching asymmetry in bilingual children: Code-switching from Dutch to Frisian requires more cognitive control than code-switching from Frisian to Dutch. *International Journal of Bilingualism*, *23*(6), 1431-1447.

Bosma, E., & Blom, E. (2020). Language activities in a minority–majority language context: Book-reading at home is more important for Frisian than for Dutch. *Journal of Child Language*, 47(2), 289-308. doi:10.1017/S0305000919000023

Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., & Sheikh, Y. (2019). OpenPose: realtime multiperson 2D pose estimation using Part Affinity Fields. *IEEE transactions on pattern analysis and machine intelligence*, 43(1), 172-186. https://arxiv.org/abs/1812.08008

Cienki, A., & Müller, C. (2008). Metaphor, gesture, and thought. *The Cambridge handbook of metaphor and thought*, *483*, 501.

Colletta, J-M. (2004). *Le développement de la parole chez l'enfant agé de 6 à 11 ans. Corps, language et cognition.* Sprimont: Mardaga.

Colletta, J.-M., Pellenq, C., & Guidetti, M. (2010). Age-related changes in co-speech gesture and narrative: evidence from french children and adults. *Speech Communication*, *52*(6), 565–576. https://doi.org/10.1016/j.specom.2010.02.009

De Jonge-Hoekstra, L., Cox, R. F., Van der Steen, S., & Dixon, J. A. (2021). Easier Said Than Done? Task Difficulty's Influence on Temporal Alignment, Semantic Similarity, and Complexity Matching Between Gestures and Speech. *Cognitive science*, *45*(6), e12989.

De Jonge-Hoekstra, L. (2021). *How hand movements and speech tip the balance in cognitive development: a story about children, complexity, coordination, and affordances* (dissertation). University of Groningen.

De Jonge-Hoekstra, L., Van der Steen, S., Van Geert, P., & Cox, R. F. (2016). Asymmetric dynamic attunement of speech and gestures in the construction of children's understanding. *Frontiers in psychology*, *7*, 473.

Nicoladis, E., Marentette, P., & Navarro, S. (2016). Gesture frequency linked primarily to story length in 4–10-year old children's stories. *Journal of psycholinguistic research*, *45*, 189-204.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Harvard University Press.

Gooskens, C., & Heeringa, W. (2004). The position of Frisian in the Germanic language area. *On the boundaries of phonology and phonetics*, 61-87.

Gullberg, M. (2021). 13 bimodal convergence: how languages interact in multicompetent language users' speech and gestures. In *Gesture in language : development across the lifespan* (pp. 317–334). essay, Berlin, Boston : De Gruyter. https://doi.org/10.1515/9783110567526-013

Gullberg, M., De Bot, K., & Volterra, V. (2010). Gestures and some key issues in the study of language development. *Gestures in language development*, 3-33.

Gullberg, M. (2011). Multilingual multimodality: Communicative difficulties and their solutions in second-language use. Embodied Interaction: *Language and Body in the Material World*, 137–151.

Heaton, J. B. (1966). *Composition through pictures*. Longman Group United Kingdom.

Hodges, L. E., Özçalışkan, Ş., & Williamson, R. (2018). Type of iconicity influences children's comprehension of gesture. *Journal of Experimental Child Psychology*, *166*, 327–339. https://doi.org/10.1016/j.jecp.2017.08.009

Hoetjes, M. W. (2018). Transfer in Gesture: L2 Placement Event Descriptions. *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, 1791-1796.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, 16(5), 367–371. doi:10.1111/j.0956-7976.2005.01542.x

Iverson, Jana M., Olga Capirci, Virginia Volterra, & Susan Goldin-Meadow (2008). Learning to talk in a gesture-rich world: Early communication of Italian vs. American children. *First Language*, 164–181.

Informatievoorziening, provincie Fryslân (2020). *De Fryske Taalatlas 2020: Friese taal in beeld.* Provincie Fryslân. Retrieved on September 9, 2023 from https://friesland.databank.nl/report/Taalatlas%202020%20%28Nederlands%29.pdf

Jaffe, J., Beatrice, B., Stanley, F., Crown, C. L., & Jasnow, M.D. (2001). Rhythms of dialogue in

infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development*, *66*(2), vi–131. https://doi.org/10.2307/3181589

Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and cognitive processes*, *24*(2), 145-167.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48, 16-32

Kosmala, L., Candea, M., & Morgenstern, A. (2019). Synchronization of (dis) fluent speech and gesture: A multimodal approach to (dis) fluency. In *Gesture and Speech in Interaction*.

Leikin, M., Schwartz, M., & Tobin, Y. (2011). Current issues in bilingualism: A complex approach to a multidimensional phenomenon. In *Current issues in bilingualism: Cognitive and socio-linguistic perspectives* (pp. 1-18). Dordrecht: Springer Netherlands.

Leonard, Thomas, Cummins, Fred, 2010. The temporal relation between beat gestures and speech. *Language and Cognitive Processes* 26 (10),1457–1471.

Lowie, W., & Verspoor, M. (2022). A Complex Dynamic Systems Theory perspective on speaking in second language development. In T. Derwing, M. Munro, & R. Thomson (Eds.), *The Routledge Handbook of Second Language Acquisition and Speaking* (pp. 39-53). (The Routledge Handbooks in Second Language Acquisition). Routledge.

Matahelumual, M.D. (2022). *The Temporal Alignment of Hand Gestures and Speech in a Second Language* (master thesis). University of Groningen. (Unpublished).

Mayberry, R. I., & Nicoladis, E. (2000). Gesture reflects language development: Evidence from bilingual children. *Current Directions in Psychological Science*, *9*(6), 192-196.

McNeill, David, 1992. *Hand and Mind: What Gestures Reveal about Thought.* University of Chicago Press, Chicago.

Nagpal, J., Nicoladis, E., & Marentette, P. (2011). Does proficiency or task difficulty explain bilinguals' gesture? *International Journal of Bilingualism*, *15*, 205–214.

Nota, A., Hilton, N. H., & Coler, M. (2016). Word and phrasal stress disentangled: Pitch peak alignment in Frisian and Dutch declarative structures. *Speech Prosody 2016*, 464-468.

doi.org/10.21437/SpeechProsody.2016-95

Ordin, M., & Polyanskaya, L. (2014). Development of timing patterns in first and second languages. *System*, *42*, 244-257.

Özçalışkan, S., Gentner, D., & Goldin-Meadow, S. (2013). Do iconic gestures pave the way for children's early verbs? *Applied Psycholinguistics*, 35(6), 1143–1162. doi:10.1017/S0142716412000720

Özçalışkan, S., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition*, 96, B101–B113. doi:10.1016/j.cognition.2005.01.001

Özer, D., & Göksun, T. (2020). Gesture use and processing: A review on individual differences in cognitive resources. *Frontiers in Psychology*, *11*, 573555.

Perrault, A., Chaby, L., Bigouret, F., Oppetit, A., Cohen, D., Plaza, M., & Xavier, J. (2019). Comprehension of conventional gestures in typical children, children with autism spectrum disorders and children with language disorders. *Neuropsychiatrie de l'Enfance et de l'Adolescence*, *67*(1), 1-9.

Pouw, W., Trujillo, J. P., & Dixon, J. A. (2020). The quantification of gesture–speech synchrony: A tutorial and validation of multimodal data acquisition using device-based and video-based motion tracking. *Behavior research methods*, *52*, 723-740.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The Art of Conversation Is Coordination: Common Ground and the Coupling of Eye Movements During Dialogue. *Psychological Science* (0956-7976), 18(5), 407–413. https://doi.org/10.1111/j.1467-9280.2007.01914.x

Roberts, L. J. (2014). Conflict, real life, and videotape: Procedures for eliciting naturalistic couple interactions. In *The Sourcebook of nonverbal measures* (pp. 481-492). Psychology Press.

Rohrer, P. L., Tütüncübasi, U., Vilà-Giménez, I., Florit-Pons, J., Esteve-Gibert, N., Ren-Mitchell, A., Shattuck-Hufnagel, S. & Prieto, P. (2023). The MultiModal MultiDimensional (M3D) labeling system. https://doi.org/10.17605/osf.io/ankdx

Rusiewicz, H. L., & Esteve-Gibert, N. (2018). Set in time: Temporal coordination of prosody and gesture in the development of spoken language production. In *The Development of Prosody in First Language Acquisition* (pp. 103-124). John Benjamins.

Schmidt, R. C., Nie, L., Franco, A., & Richardson, M. J. (2014). Bodily synchronization underlying joke telling. *Frontiers in Human Neuroscience*, 8. https://doi.org/10.3389/fnhum.2014.00633

Schneider, S., Ramirez-Aristizabal, A. G., Gavilan, C., & Kello, C. T. (2020). Complexity matching and lexical matching in monolingual and bilingual conversations. *Bilingualism: Language and Cognition*, *23*(4), 845-857.

Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication, 57*, 209-232.

Appendices

Appendix A: Experimental Designs study 1

Protocol 1: Filling in the story

Language: All

Procedure The two children are welcomed to the experiment and placed at a table facing each other. Before the experiment starts everything is explained to them and there will be some time to get to know each other. For the first step the experimenter calls them away one by one for about 3 minutes. Both children will get about 3 minutes to look at a part of the pictures from figure 1 and ask questions about the experiment. The first child gets to see picture 1 & 2. The second child gets to see picture 3 & 5. When the children are back at their chairs picture 6 is shown to both of them and they are asked to tell each other what they saw in the pictures to find out what happened in the story. When they are correct they will get a reward. When they are not correct they will get a reward because they participated in the experiment. The whole task will take about 20 minutes.

Figure 1

'Dog Story' by Heaton (1966)

Reference

Heaton, J. B. (1966). *Composition through pictures*. Longman Group United Kingdom.

Protocol 2: Snack attack

Language: All

Age: 7+

Procedure: The two children are welcomed to the experiment and placed at a table facing each other. Before the experiment starts everything is explained to them and there will be some time to get to know each other. Then they will both be shown the same first part of the short video 'Snack attack' (See figure 2) until minute 2:23. Afterwards they have to discuss together what might have really happened in the story. When they finish discussing the end of the video will be shown to them. When they are correct they will get a reward. When they are not correct they will get a reward because they participated in the experiment. The whole task will take about 20 minutes.

Figure 2

Still from 'Snack attack'

Reference

E. Verastegui (June 14, 2016). Snack Attack [YOUTUBE]. Retrieved from https://www.youtube.com/watch?v=38y_1EWIE9

Protocol 3: The dog and the bird

Language: All

Age: 7+

Procedure The two children are welcomed to the experiment and placed at a table facing each other. Before the experiment starts everything is explained to them and there will be some time to get to know each other. Then they will both be shown the same video 'Dog and Bird' (See figure 3) until minute 2:36. Afterwards they have to discuss together if the dog did something wrong or not. They will be given a cue if needed: 'He gave away something from his boss, but was that a right thing to do?' The children are asked to think of a reason why the dog did something wrong and why he did something good. They will be told that they get a reward if they think of two correct reasons. If they do not succeed, they will get a reward anyway because they participated in the experiment.

Figure 3

Still from 'The Dog and the Bird'

Reference

A Joy Story [YOUTUBE]. JD.COM (February 7, 2018). Retrieved from A Joy Story: Joy and Heron - YouTube.

Protocol 4: The mole who knew it was none of his business

Language: English/Dutch

Age: 6+

Procedure: The two children are welcomed to the experiment and placed at a table facing each other. Before the experiment starts everything is explained to them and there will be some time to get to know each other. Then they will both be shown the same short video of the storybook of the mole who knew it was none of his business (Figure 4). Before this, they will both be taken apart for a short instruction. One child will get the instruction to look at the type of poop they see in the video, the other child will get the instruction to look at the type of animals in the video. Afterwards they have to discuss together what animals appeared in the story and what type of poop they had. They will be told that the more correct animals and poop they name the more points they will get. In the end they will get a reward because they participated in the experiment.

Figure 4

Still from 'The mole who knew it was none of his business'

Reference

Juf Lisette TV - Over een kleine mol [YOUTUBE]. Lisette van der Beek (2020, april 6). Retrieved from https://www.youtube.com/watch?v=wohOz_V8Z_c

Protocol 5: Debate phrases

Language: All

Age: 7+

Procedure The two children are welcomed to the experiment and placed at a table facing each other. Before the experiment starts everything is explained to them and there will be some time to get to know each other. With the help of a speaker debate phrases will be read out loud to them. The children have to decide together which phrase is true. They have to mention at least two arguments for their choice, before they will receive the new statement. The more statements they pass, the more points they will get. In the end they will get rewarded for taking part in the experiment.

Statements (from concrete to abstract):

1. Would you only eat sweet things or salty things if you had to choose?
2. Which animal is better to have as a pet: A dog or a cat?
3. Which season is better: Summer or Winter?
4. What would be better: Being a king or being a superhero?
5. What would be better: Being able to fly or to be invisible?
6. What is better: Giant feet or giant hands?
7. What is better: Eating your favorite food for the rest of your life or never again?
8. What came first: The chicken or the egg?
9. What would be better: Not being able to taste or to smell?
10. What would be better: Living on a deserted island or at the North Pole?
11. What would be better: Living in the sea or in space?
12. What would be better: Being the funniest person or the smartest person in the world?
13. What would be better: Being scared of nothing or everything?

Appendix B: Protocol This-or-That Debate phrases

Protocol This-or-That Debate phrases

Name This-or-That Debate phrases (Dutch name: Dit-of-dat debatstellingen)

Purpose Eliciting spontaneous conversation and hand gestures in children, during speech.

Origin Own design, based on observations with children aged 7 to 8 years old.

Duration Approx 15 minutes (5 minutes instruction time, 5-10 minutes execution time).

Materials used Paper Sheet with debate phrases, Go Pro camera, tape, head microphones, two laptops with a google meet connection and an audio speaker, Timer on a laptop, timer on a phone.

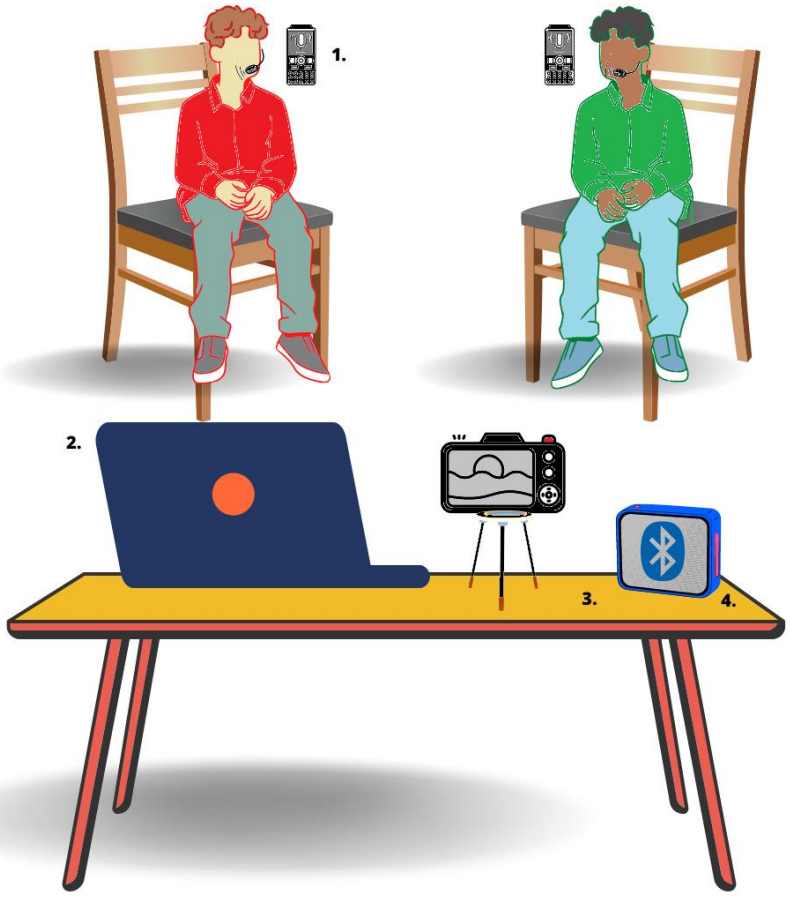
Set up The children will be seated on two chairs that are slightly shifted towards each other (see figure 1). They both have a microphone, attached to their cheek with a piece of tape, that records their voice. On the table in front of them there will be a speaker, a laptop with a timer on and a Go pro. The laptop is connected to google meets with a one-way connection; The children will be visible and audible for the experimenter but the experimenter is only audible for the children. The speaker is connected to the laptop; The experimenter communicates via the speaker. In this way, the experimenter will have the children in vision and can intervene when anything undesirable happens. The experimenter may also help to stimulate the conversation of the children when this is really needed. For both situations, possible responses are preconstructed (see Table 1). Furthermore, they can actively read out the statements without intervening the natural conversation between the two children.

Short description: The aim of this task is to resolve as many This-or-That debate phrases as possible (see figure 2 for the list). To do so, both children need to agree with each other on a topic, by choosing the same side out of two options. In addition, they both need to name at least one argument in favour of their collective choice. The experimenter welcomes the children in a room and explains the experiment to them (see example protocol below). The experimenter sets a timer on both the laptop and their phone, running from 10:00 minutes until 00:00 minutes and leaves the room (taking their phone with the timer). Now, the experiment officially starts. The statements will be provided to the children via the speaker, in the order provided in figure 2. Each consequent statement will be given by the experimenter to the children when they successfully give two arguments for the same topic choice. The game will continue until twelve ten statements are resolved or when the 10 minutes

running time is over. Afterwards, both children will receive a sticker for their efforts.

Figure 1

Experiment set-up This-or-That debate phrases



1: Head microphones
2: Laptop with google meets connection and running timer
3: Go Pro
4: Speaker connected to google meets

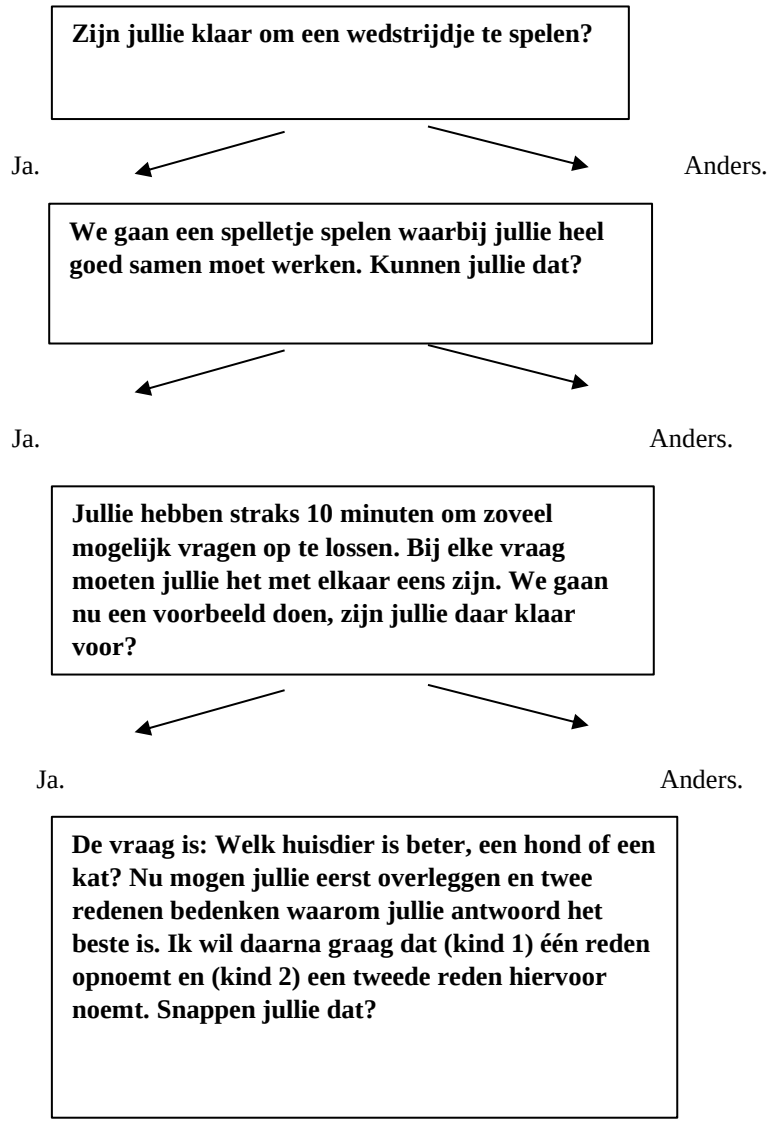
Figure 2

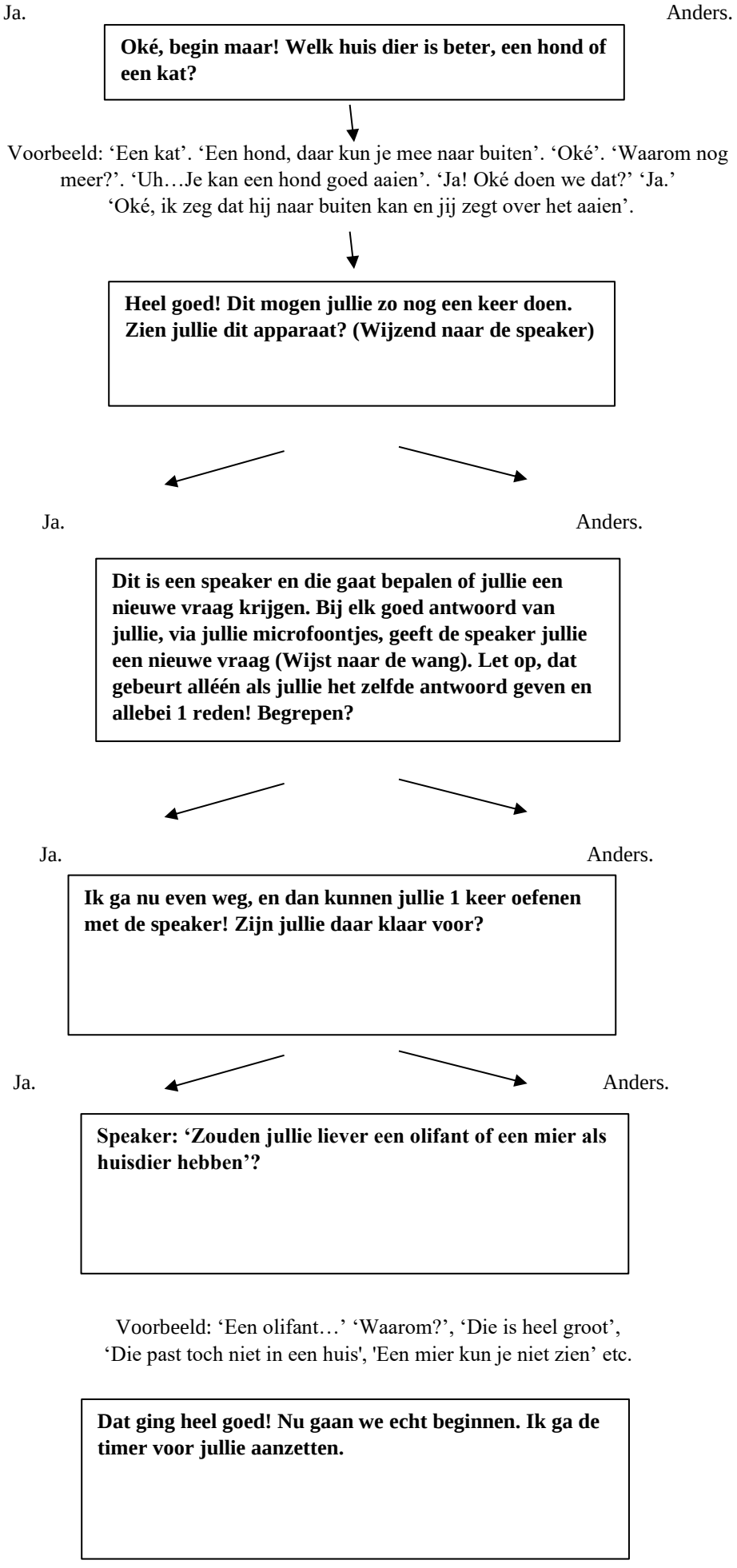
This-or-That debate phrases

This-or-That Debate phrases
1. Zouden jullie liever een koning of een superheld willen zijn?
2. Zouden jullie liever zo klein als een kabouter of zo groot als een reus willen zijn?
3. Zouden jullie liever vier handen of vier benen willen hebben?
4. Zouden jullie liever kunnen vliegen of onzichtbaar kunnen zijn?
5. Zouden jullie liever reusachtige handen of reusachtige voeten willen hebben?
6. Wat zou beter zijn, nooit meer kunnen ruiken of nooit meer kunnen proeven?
7. Wat zou beter zijn, de grappigste of de slimste persoon op de wereld zijn?
8. Wat zou beter zijn, in de zee of in de ruimte wonen?
9. Zouden jullie liever kunnen toveren of de toekomst voorspellen?
10. Zouden jullie liever kunnen tijdreizen of de tijd stil kunnen zetten?
11. Zouden jullie liever met dieren kunnen praten of elkaars gedachten kunnen lezen?
12. Wat was er als eerste; De kip of het ei?

Bonus: Wie denken jullie dat er gewonnen heeft (Jullie of de andere groepjes)

Example of protocol (explanation of the task)




Ja.
Anders.
Oké, begin maar! Welk huis dier is beter, een hond of een kat?
Voorbeeld: 'Een kat'. 'Een hond, daar kun je mee naar buiten'. 'Oké'. 'Waarom nog meer?'. 'Uh…Je kan een hond goed aaien'. 'Ja! Oké doen we dat?' 'Ja.'
'Oké, ik zeg dat hij naar buiten kan en jij zegt over het aaien'.
Heel goed! Dit mogen jullie zo nog een keer doen. Zien jullie dit apparaat? (Wijzend naar de speaker)
Ja.
Anders.
Dit is een speaker en die gaat bepalen of jullie een nieuwe vraag krijgen. Bij elk goed antwoord van jullie, via jullie microfoontjes, geeft de speaker jullie een nieuwe vraag (Wijst naar de wang). Let op, dat gebeurt alléén als jullie het zelfde antwoord geven en allebei 1 reden! Begrepen?
Ja.
Anders.
Ik ga nu even weg, en dan kunnen jullie 1 keer oefenen met de speaker! Zijn jullie daar klaar voor?
Ja.
Anders.
Speaker: 'Zouden jullie liever een olifant of een mier als huisdier hebben'?
Voorbeeld: 'Een olifant…' 'Waarom?', 'Die is heel groot', 'Die past toch niet in een huis', 'Een mier kun je niet zien' etc.
Dat ging heel goed! Nu gaan we echt beginnen. Ik ga de timer voor jullie aanzetten.

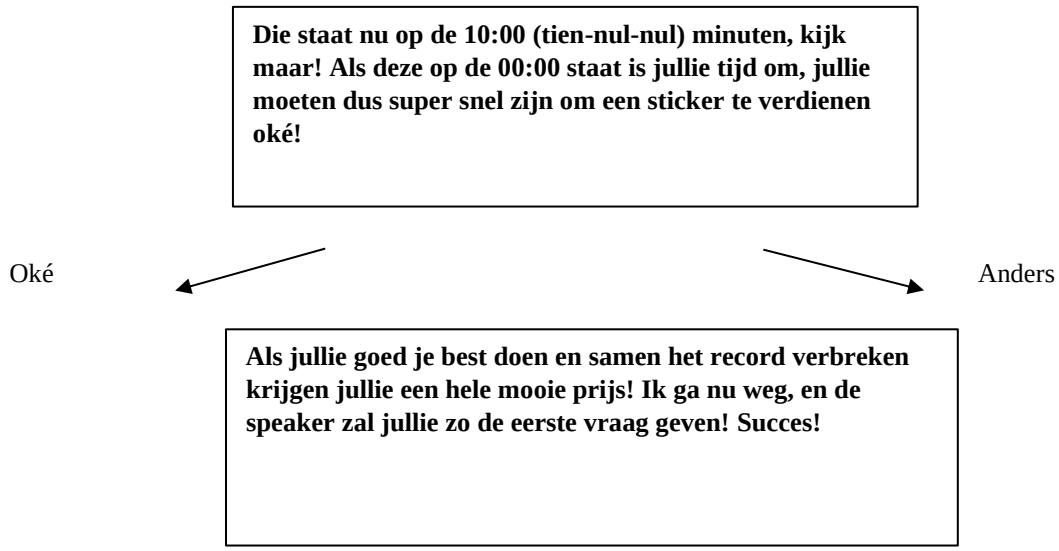


Table 1

Undesirable events and interventions

Situation	**Response**
One of the children gets up from their chair	'[Naam] wil je weer op je stoel gaan zitten?' *[Name], keep sitting down please*
One of the children is not speaking	'[Naam], wat denk jij dat een goed antwoord is?' *[Name], which answer you think is correct?*
The children are forgetting to name two arguments for their statements	'Welke twee redenen hebben jullie bedacht?' *Which two reasons did you think of?*
The children are forgetting to agree on the same choice	'Jullie moeten eerst wel hetzelfde kiezen' *First, you have to chose the same option*
The children start talking about subjects not related to the experiment for longer than 30 seconds	'De tijd loopt verder! Let op, hoe meer redenen jullie bedenken hoe eerder jullie kunnen winnen' *The time is running! Listen, the more reasons you think of, the sooner you will win*
Both children are not speaking	'Denk nog eens goed na, welke reden kunnen jullie nog meer bedenken?' *Try one more time, which other reasons can you think of?*
The children struggle to provide answer for one particular statement, for example because they fail to agree	'Oké, jullie krijgen nu eerst een nieuwe vraag om op te lossen' *Oké, now you will firstly get a new question to solve*

Appendix C: Additional data study 2

Table 1

Duration of each statement for each dyad

Dyad	Condition	Age Left child	Age right child	State-ment 1 (sec)	State-ment 2 (sec)	State-ment (Sec)	State-ment 4 (Sec)	State-ment 5 (Sec)	State-ment 6 (Sec)	State-ment 7 (Sec)	State-ment 8 (Sec)	State-ment 9 (Sec)	State-ment 10 (Sec)	State-ment 11 (Sec)	State-ment 12 (Sec)	Total duration (min)	Average per statement (sec)
L1.1	L1-L1	6.3	6.6	41	44	22	31	70	27	78	21	38	29	51	57	8.48	42
L1.2	L1-L1	6.6	7.4	31	34	33	79	64	58	76	35	45	82	61	63	11.01	55
L1.3	L1-L1	9.3	9.4	21	30	30	20	28	32	20	38	48	32	33	75	6.78	34
L1.4	L1-L1	8.8	7.8	20	30	17	16	21	51	62	43	22	27	43	140	8.18	41
L2.1	L2-L1	8.3	7.6	17	21	34	34	35	42	78	31	60	27	34	46	7.65	38
L2.2	L2-L1	8.1	7.9	25	25	22	30	36	87	21	25	61	37	74	63	8.43	42
L2.3	L2-L1	8.3	7.6	29	40	90	38	31	29	29	40	51	30	49	54	8.5	43
Mean				**26.29**	**31.86**	**35.43**	**35.43**	**40.71**	**46.57**	**52**	**33.29**	**46.43**	**37.71**	**49.29**	**71.14**	**8.44**	**42.18**
Standard deviation				**0.82**	**0.81**	**2.49**	**2.07**	**1.87**	**2.13**	**2.75**	**0.81**	**1.35**	**1.98**	**1.47**	**3.17**	**1.30**	**6.48**

Table 2

Number of utterances within each dyad

Dyad	Condition	Age left child	Age right child	Number of utterances left child	Percentage of total utterances	Number of utterance right child	Percentage of total utterances	Total number of utterances
L1.1	L1-L1	6.3	6.6	58	51.33%	55	48.67%	113
L1.2	L1-L1	6.6	7.4	50	54.35%	42	45.65%	92
L1.3	L1-L1	9.3	9.4	64	48.85%	67	51.15%	131
L1.4	L1-L1	8.8	7.8	70	57.85%	51	42.15%	121
L2.1	L2-L1	8.3	7.6	45	45%	55	55%	100
L2.2	L2-L1	8.1	7.9	71	48.63%	75	51.37%	146
L2.3	L2-L1	8.3	7.6	93	55.36%	75	44.64%	168
Mean				**64.43**	**51.62%**	**60**	**48.38%**	**124.42**
Standard devation				**15.90**	**4.40%**	**12.61**	**4.49%**	**26.45**

Table 3

Number and share of utterances expressed by experimenter

Dyad	Condition	Age left child	Age right child	Number of stimulations (experimenter)	Percentage of all utterances	Utterances of children	All utterances
L1.1	L1_L1	6.3	6.6	5	4.24%	113	118
L1.2	L1_L1	6.6	7.4	18	16.36%	92	110
L1.3	L1_L1	9.3	9.4	2	1.50%	131	133
L1.4	L1_L1	8.8	7.8	6	4.72%	121	127
L2.1	L1_L2	8.3	7.6	5	4.76%	100	105
L2.2	L1_L2	8.1	7.9	12	7.59%	146	158
L2.3	L1_L2	8.3	7.6	7	4%	168	175
Mean				7.86	6.17%	124.43	132.29
SD				5.40	4.83%	26.45	25.69

Table 4

*Number of hand gestures for the left and right child per episode **

Dyad	Condition	Age left child	Age right child	L1	R1	L2	R2	L3	R3	L4	R4	L5	R5	L6	R6	L7	R7	L8	R8	L9	R9	L10	R10	L11	R11	L12	R12	Tot. L	Tot. R	Tot. dyads	Utterances	Utterances/Hand gestures
L1.1	L1_L1	6.3	6.6	3	1	1	0	4	0	2	2	1	3	3	1	3	11	0	1	2	7	3	5	2	3	2	3	26	37	63	113	0.56
L1.2	L1_L1	6.6	7.4	0	0	1	1	0	2	0	1	2	0	0	0	0	0	0	0	1	0	0	1	1	2	1	0	6	7	13	92	0.14
L1.3	L1_L1	9.3	9.4	0	1	0	3	1	2	1	0	5	0	2	0	2	0	6	0	2	9	3	7	1	4	9	5	32	31	63	131	0.48
L1.4	L1_L1	8.8	7.8	2	2	5	4	1	1	3	0	3	1	0	1	6	2	4	3	1	1	1	2	0	0	3	4	29	21	50	121	0.41
L2.1	L2_L1	8.3	7.6	5	0	0	1	1	1	1	2	7	1	5	3	5	3	3	0	0	2	0	0	0	2	7	4	34	19	53	100	0.53
L2.2	L2_L1	8.1	7.9	1	0	4	1	3	1	2	2	1	4	5	3	1	1	3	1	2	1	0	0	1	4	9	9	32	27	59	146	0.40
L2.3	L2_L1	8.3	7.6	5	2	4	1	13	4	2	3	1	1	1	4	5	5	4	6	8	5	5	2	3	3	8	1	59	37	96	168	0.57
Total				**16**	**6**	**15**	**11**	**23**	**11**	**11**	**10**	**20**	**10**	**16**	**12**	**22**	**22**	**20**	**11**	**16**	**25**	**12**	**17**	**8**	**18**	**39**	**26**			**397**	**871**	

* L1 represents the number of hand gestures used during the first statement for the left child, R1 for the right child etc.

Appendix D: Additional data study 3

Peaks of hand gestures and speech per episode for dyad L1.1 (L1-L1)

Figure 1

Peaks of hand gestures and speech for the left child, dyad L1.1 episode 1

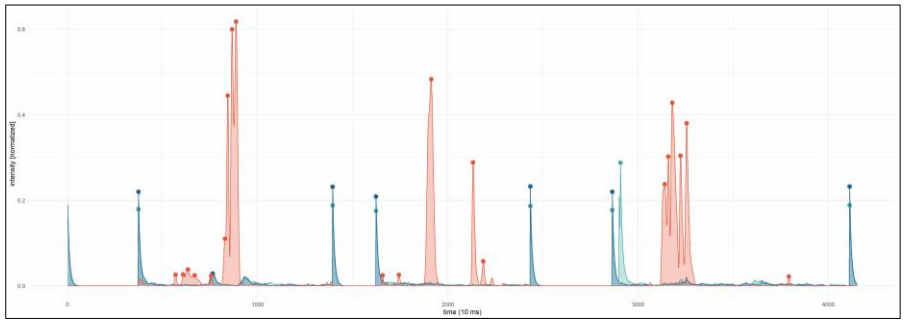


Figure 2

Peaks of hand gestures and speech for the right child, dyad L1.1 episode 1

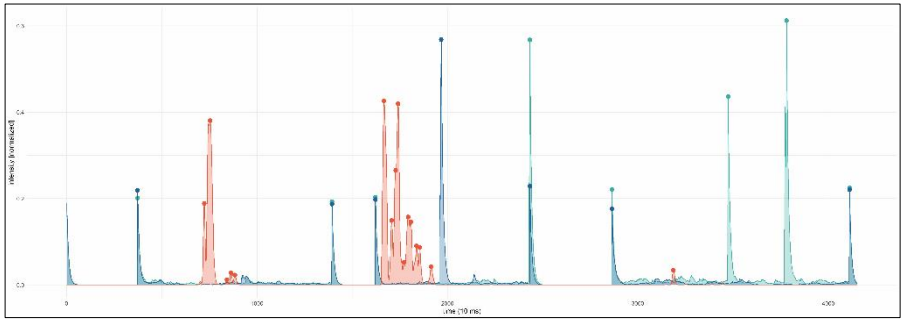


Figure 3

Peaks of hand gestures and speech for the left child, dyad L1.1 episode 2

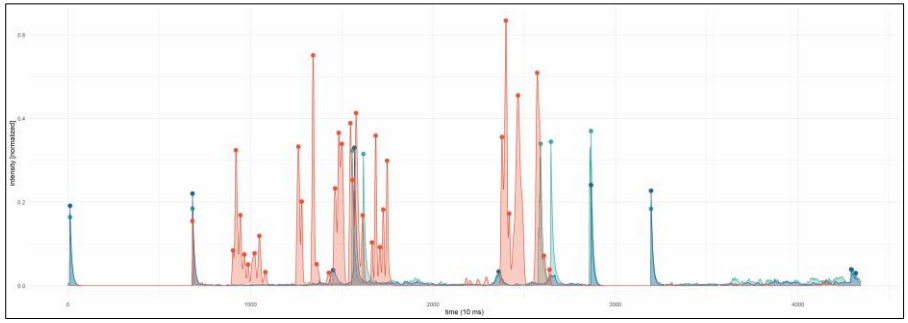


Figure 4

Peaks of hand gestures and speech for the right child, dyad L1.1 episode 2

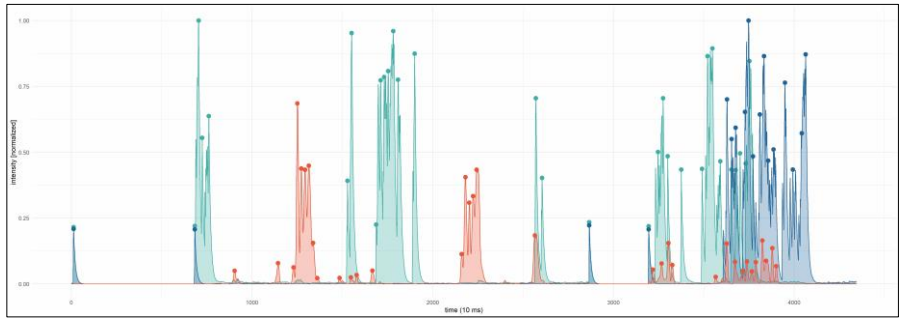


Figure 5

Peaks of hand gestures and speech for the left child, dyad L1.1 episode 3

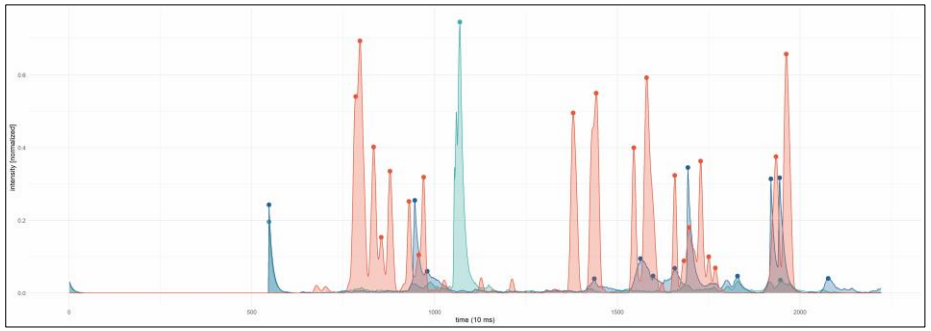


Figure 6

Peaks of hand gestures and speech for the right child, dyad L1.1 episode 3

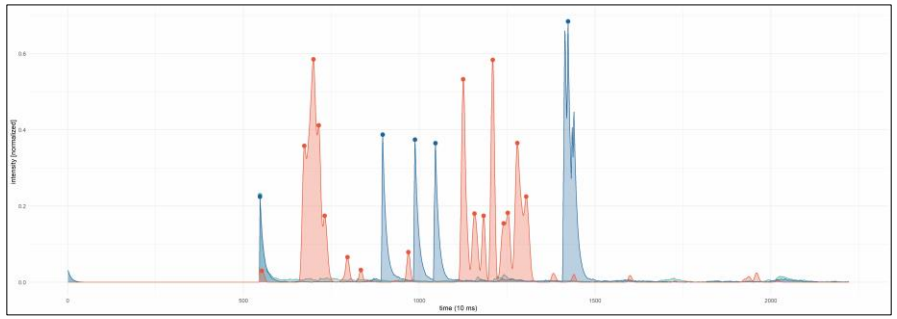


Figure 7

Peaks of hand gestures and speech for the left child, dyad L1.1 episode 4

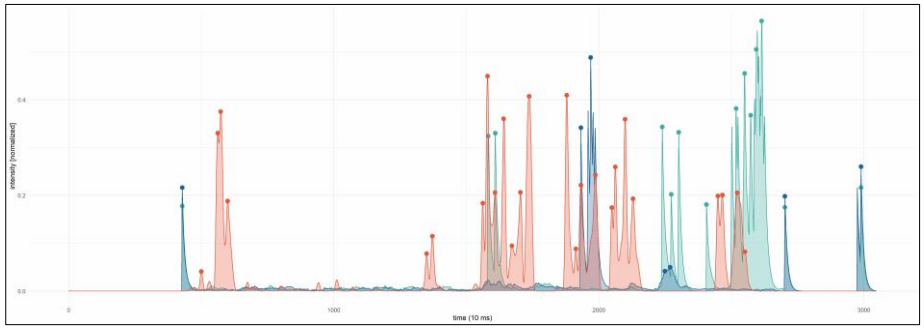


Figure 8

Peaks of hand gestures and speech for the right child, dyad L1.1 episode 4

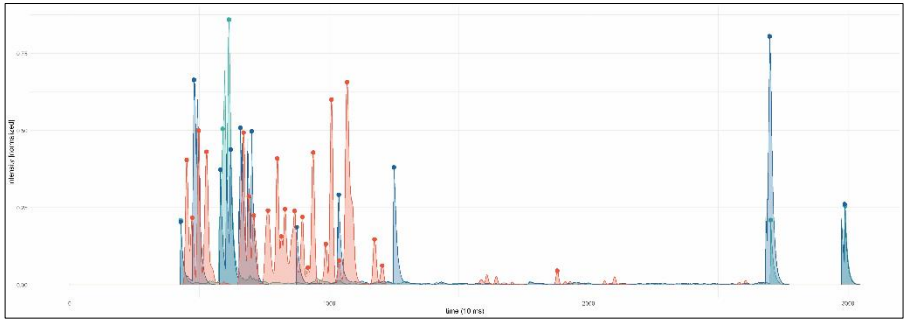


Figure 9

Peaks of hand gestures and speech for the left child, dyad L1.1 episode 5

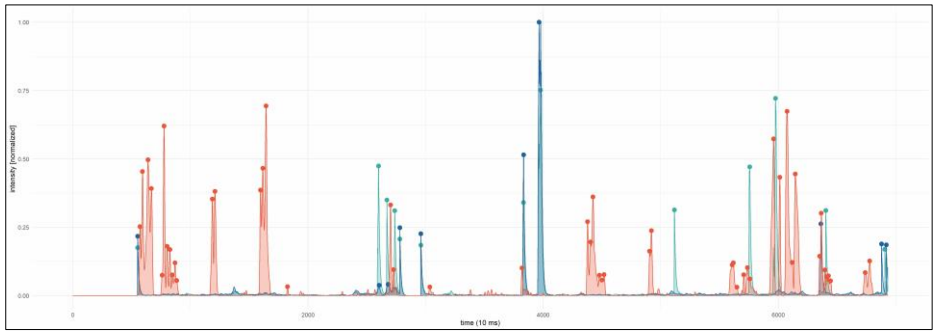


Figure 10

Peaks of hand gestures and speech for the right child, dyad L1.1 episode 5

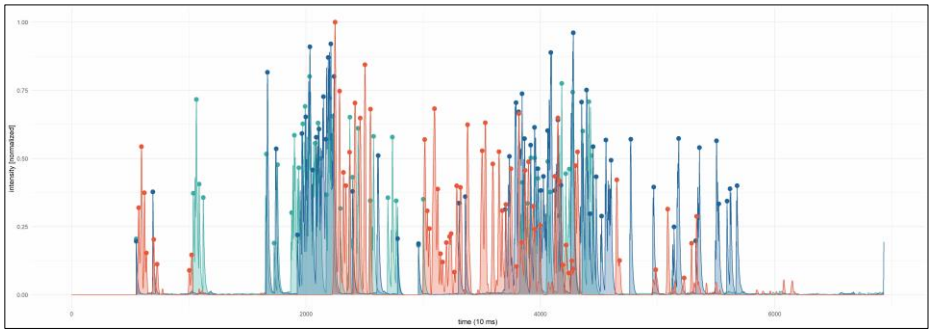


Figure 11

Peaks of hand gestures and speech for the left child, dyad L1.1 episode 6

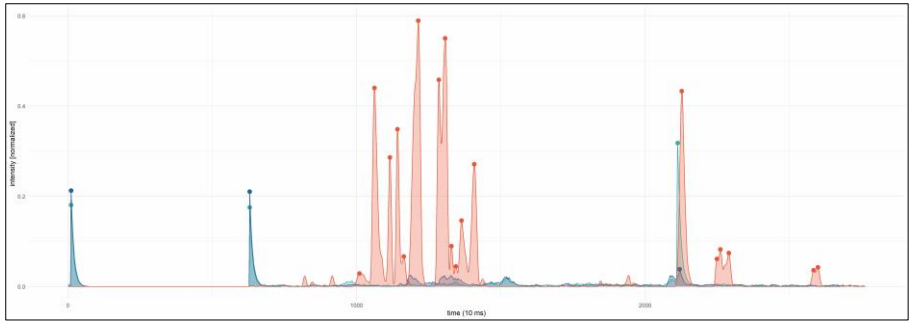


Figure 12

Peaks of hand gestures and speech for the right child, dyad L1.1 episode 6

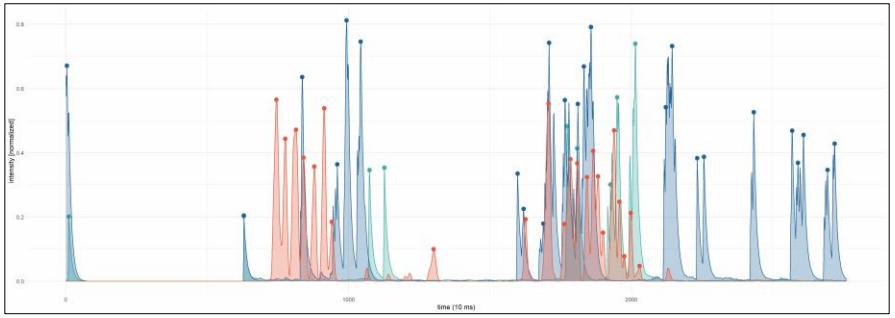


Figure 13

Peaks of hand gestures and speech: left child, dyad L1.1 episode 7

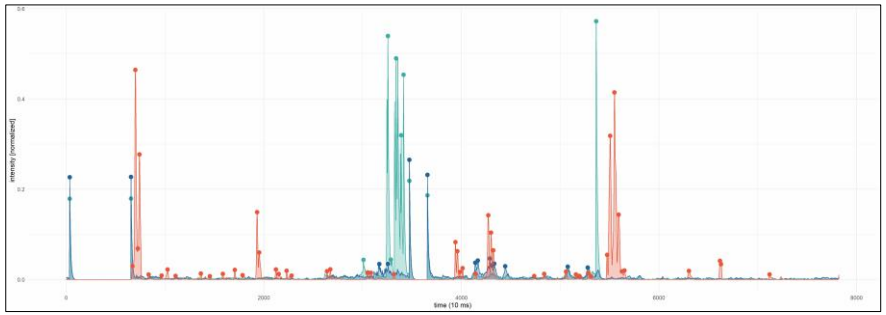


Figure 14

Peaks of hand gestures and speech: right child, dyad L1.1 episode 7

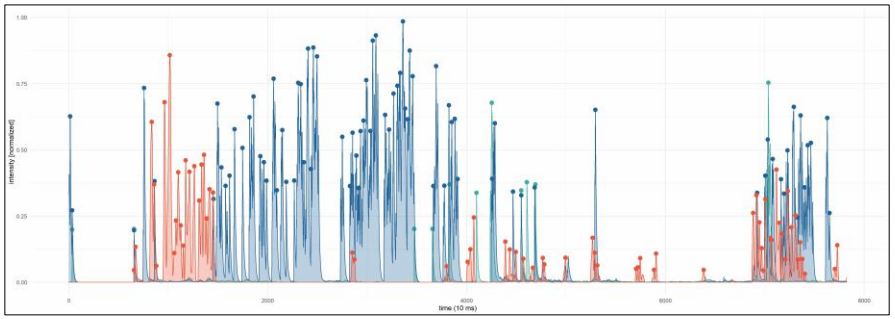


Figure 15

Peaks of hand gestures and speech: left child, dyad L1.1 episode 8

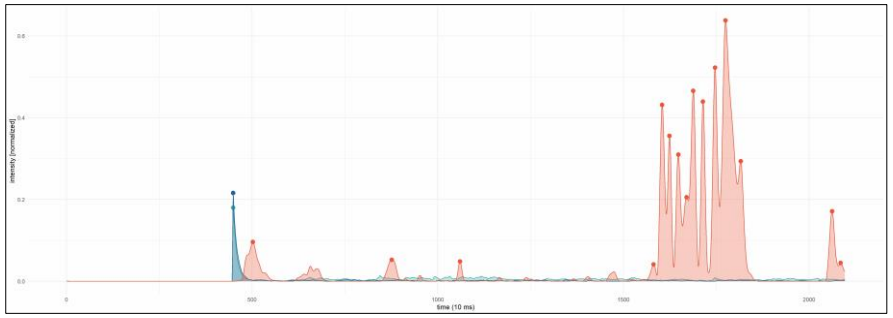


Figure 16

Peaks of hand gestures and speech: right child, dyad L1.1 episode 8

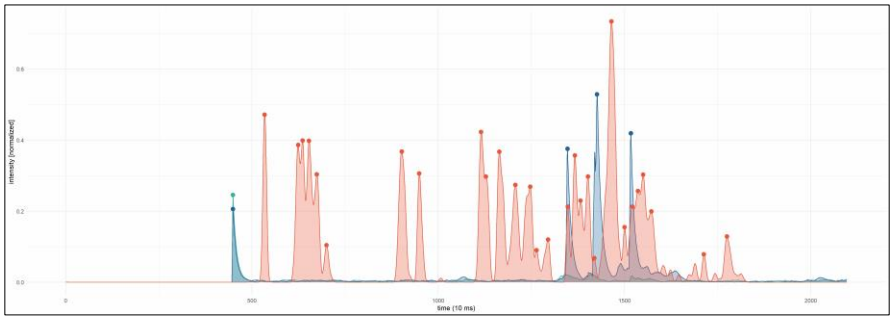


Figure 17

Peaks of hand gestures and speech: left child, dyad L1.1 episode 9

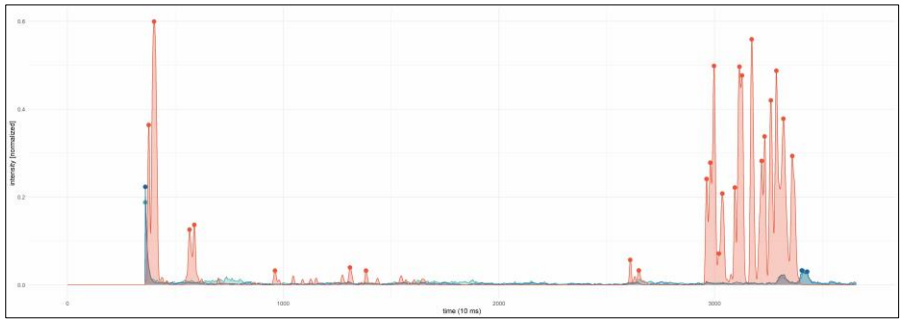


Figure 18

Peaks of hand gestures and speech: right child, dyad L1.1 episode 9

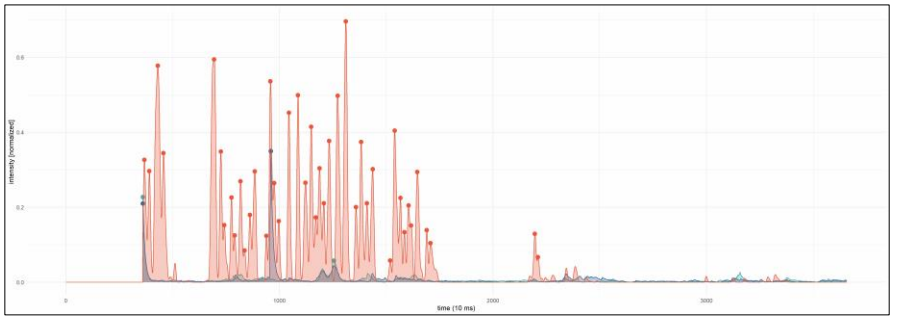


Figure 19

Peaks of hand gestures and speech: left child, dyad L1.1 episode 10

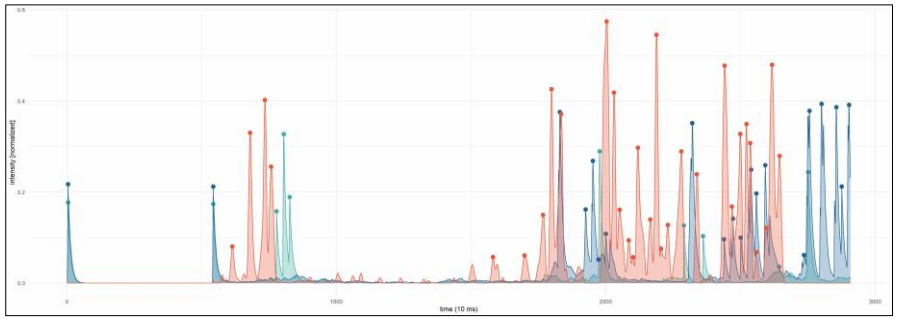


Figure 20

Peaks of hand gestures and speech: right child, dyad L1.1 episode 10

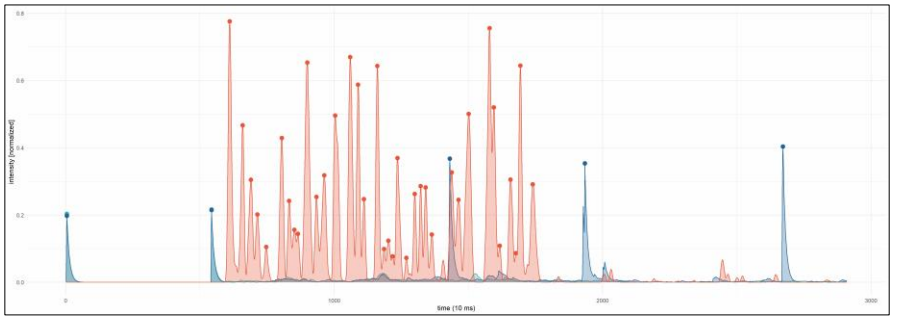


Figure 21

Peaks of hand gestures and speech: left child, dyad L1.1 episode 11

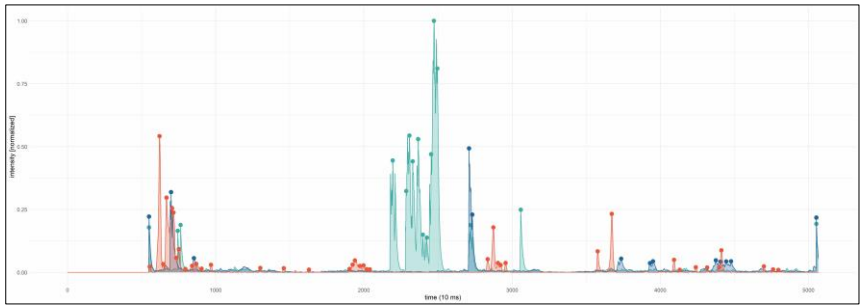


Figure 22

Peaks of hand gestures and speech: right child, dyad L1.1 episode 11

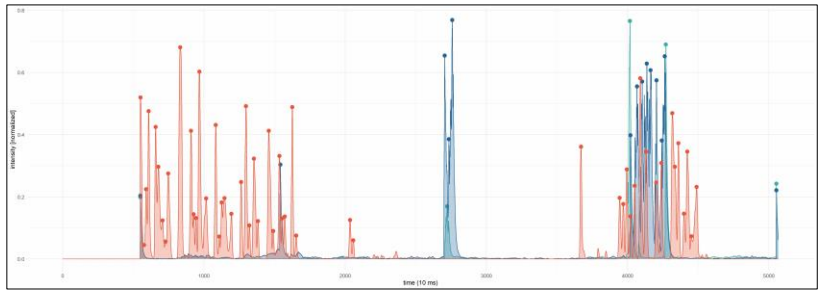


Figure 23

Peaks of hand gestures and speech: left child of dyad L1.1 episode 12

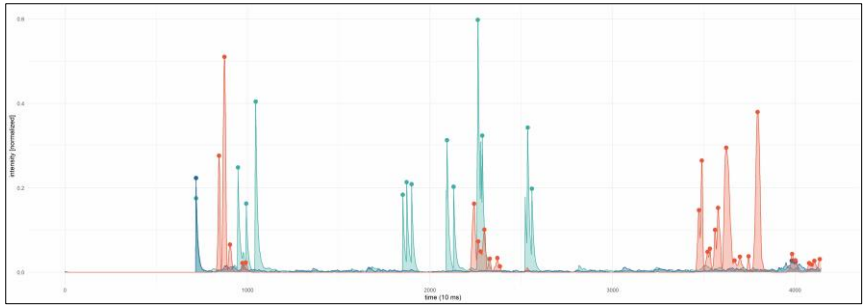
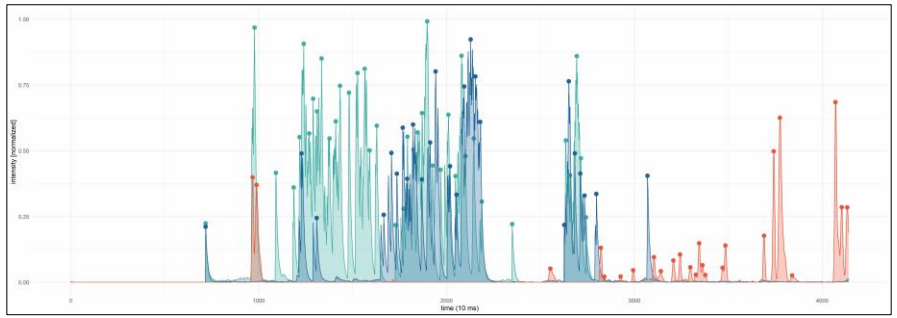


Figure 24

Peaks of hand gestures and speech: right child of dyad L1.1 episode 12



Peaks of hand gestures and speech per episode for dyad L2.1 (L2-L1)

Figure 25

Peaks of hand gestures and speech: left child of dyad L2.1 episode 1

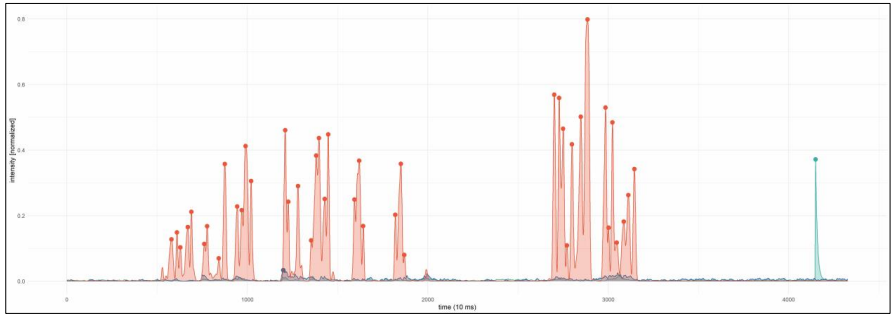


Figure 26

Peaks of hand gestures and speech: right child of dyad L2.1 episode 1

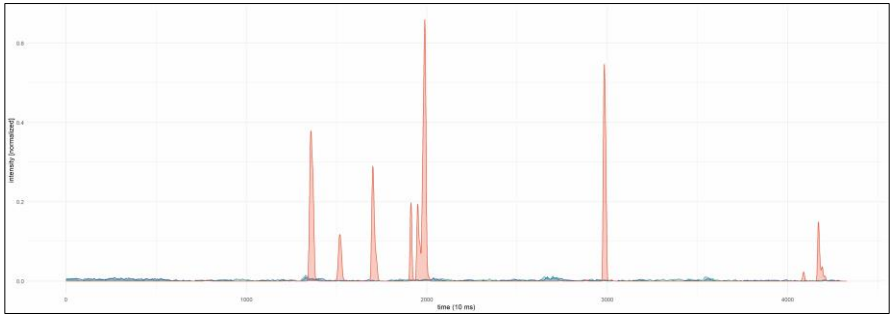


Figure 27

Peaks of hand gestures and speech: left child of dyad L2.1 episode 2

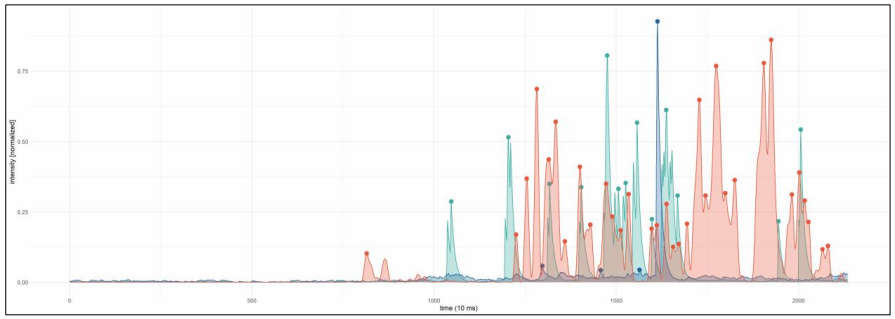


Figure 28

Peaks of hand gestures and speech: right child of dyad L2.1 episode 2

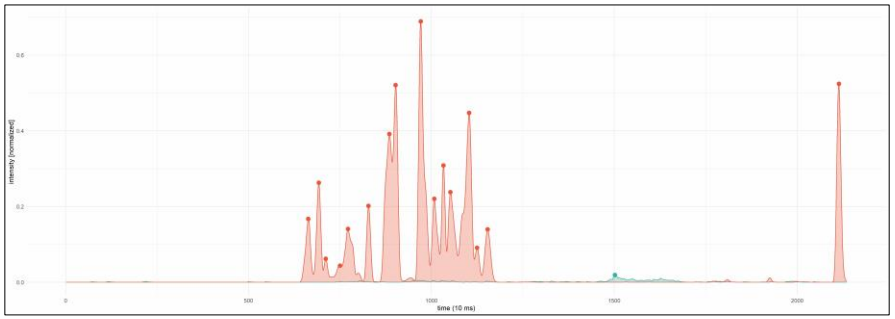


Figure 29

Peaks of hand gestures and speech: left child of dyad L2.1 episode 3

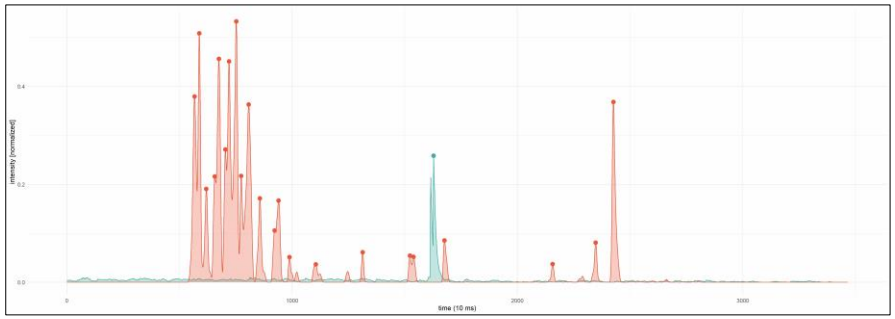


Figure 30

Peaks of hand gestures and speech: right child of dyad L2.1 episode 3

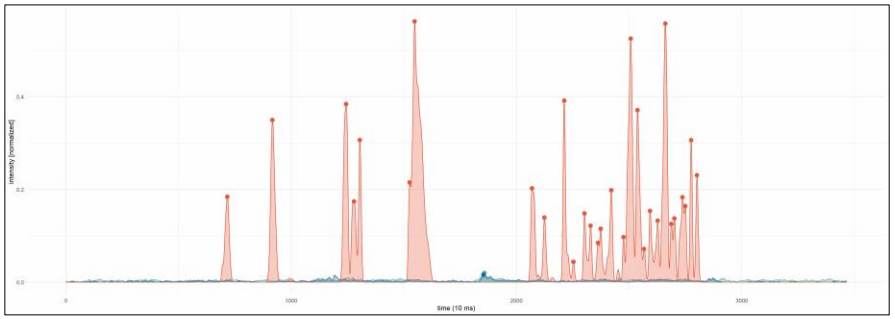


Figure 31

Peaks of hand gestures and speech: left child of dyad L2.1 episode 4

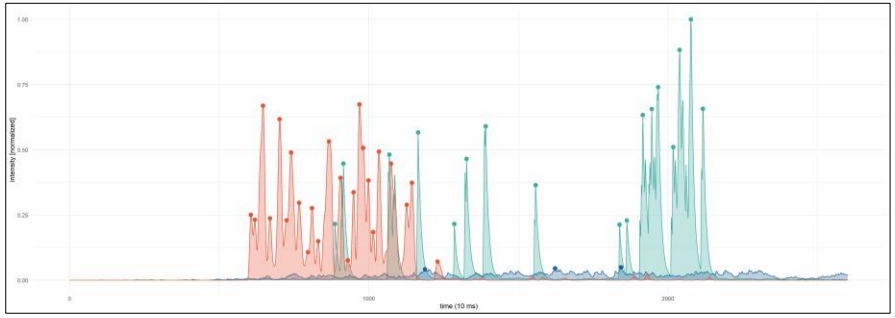


Figure 32

Peaks of hand gestures and speech: right child of dyad L2.1 episode 4

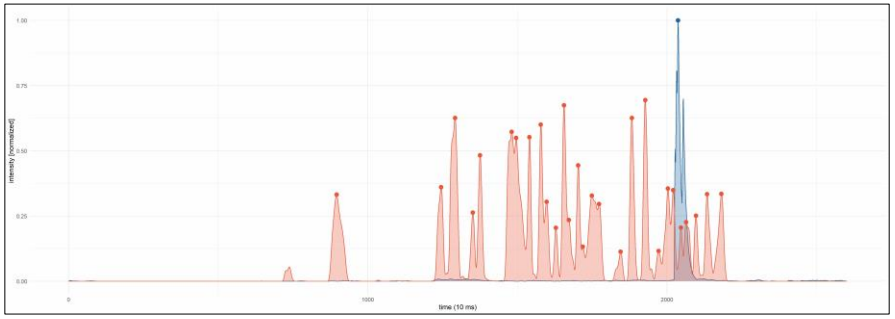


Figure 33

Peaks of hand gestures and speech: left child of dyad L2.1 episode 5

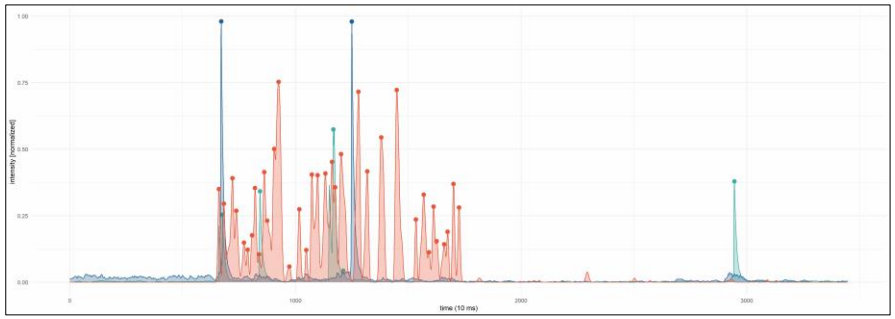


Figure 34

Peaks of hand gestures and speech: right child of dyad L2.1 episode 5

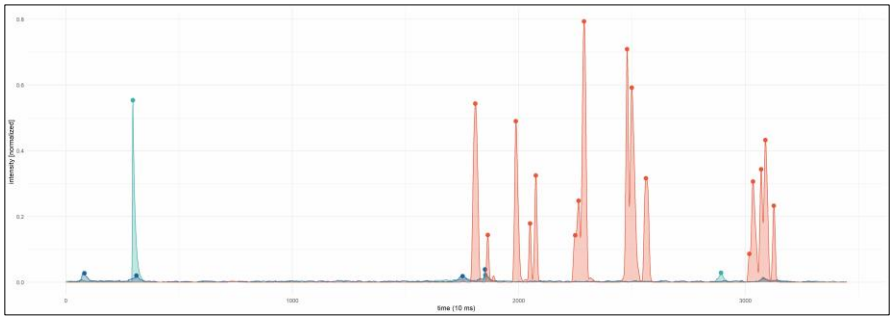


Figure 35

Peaks of hand gestures and speech: left child of dyad L2.1 episode 6

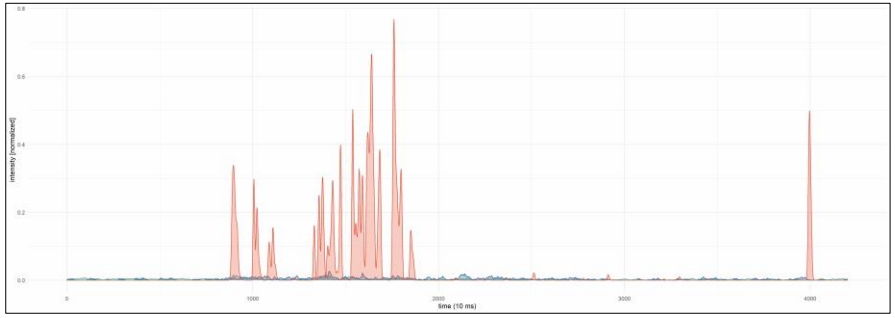


Figure 36

Peaks of hand gestures and speech: right child of dyad L2.1 episode 6

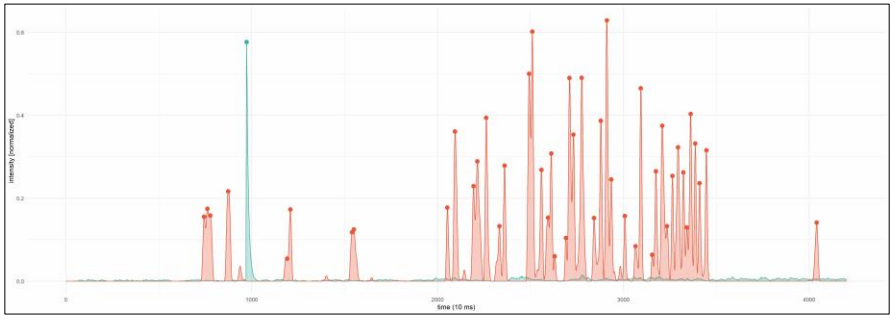


Figure 37

Peaks of hand gestures and speech: left child of dyad L2.1 episode 7

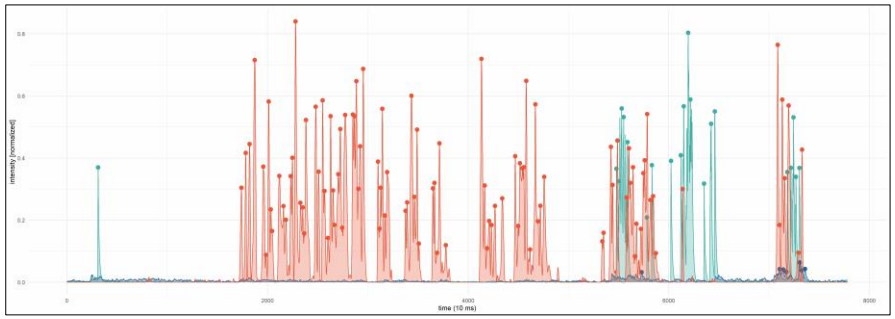


Figure 38

Peaks of hand gestures and speech: right child of dyad L2.1 episode 7

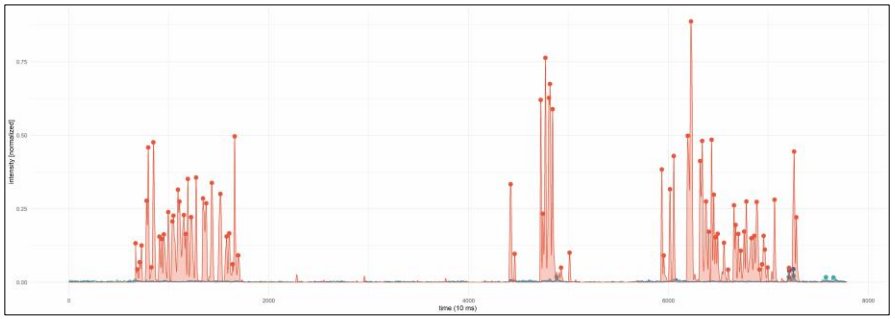


Figure 39

Peaks of hand gestures and speech: left child of dyad L2.1 episode 8

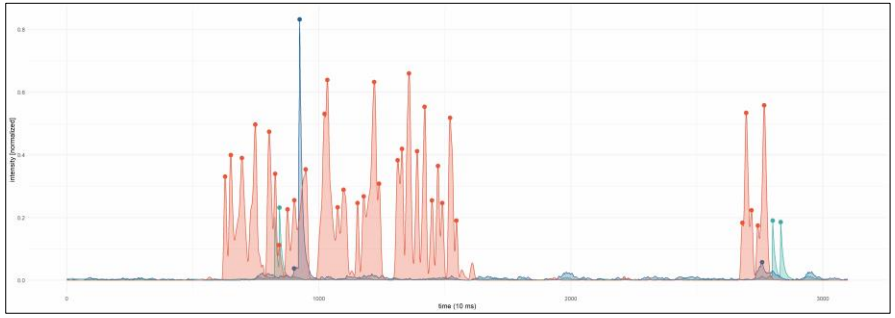


Figure 40

Peaks of hand gestures and speech: right child of dyad L2.1 episode 8

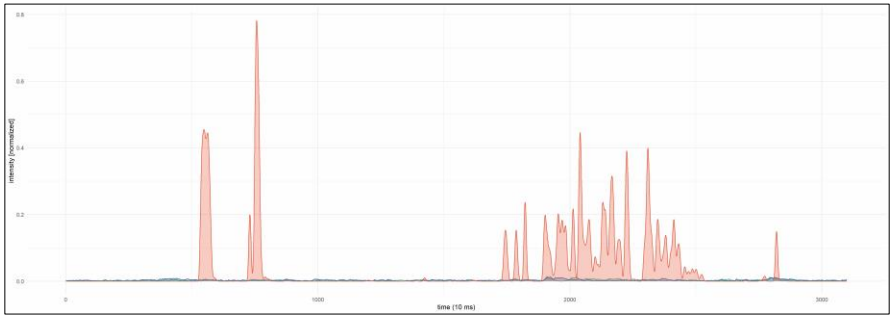


Figure 41

Peaks of hand gestures and speech: left child of dyad L2.1 episode 9

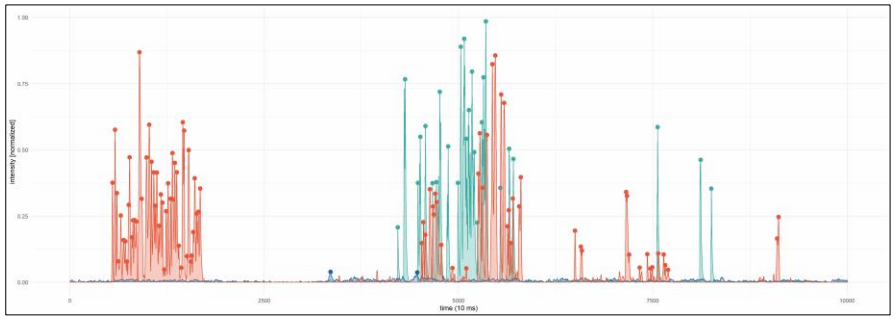


Figure 42

Peaks of hand gestures and speech: right child of dyad L2.1 episode 9

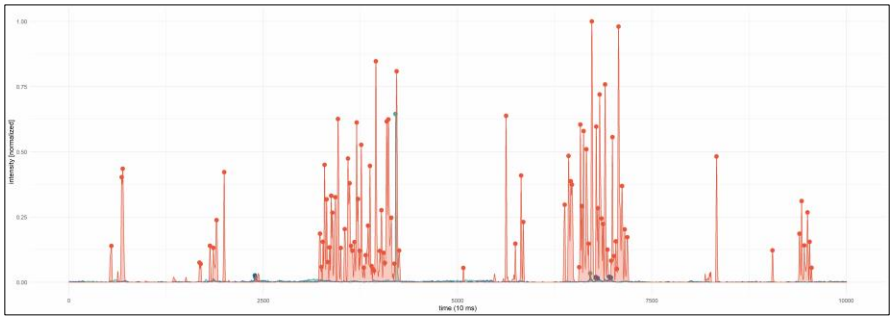


Figure 43

Peaks of hand gestures and speech: left child of dyad L2.1 episode 10

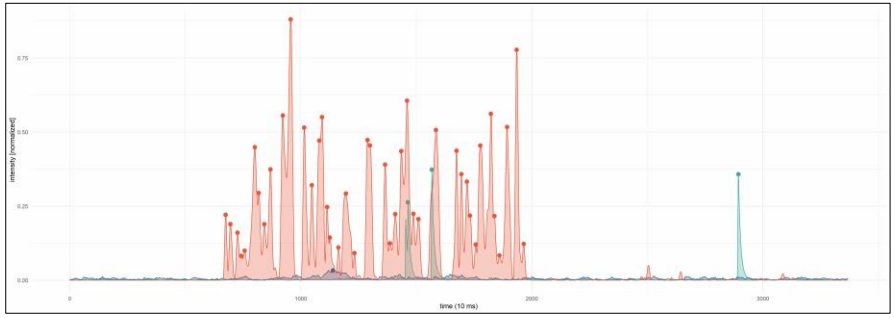


Figure 44

Peaks of hand gestures and speech: right child of dyad L2.1 episode 10

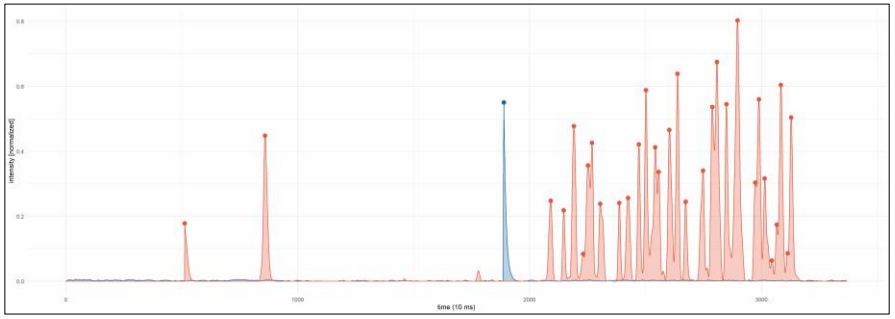


Figure 45

Peaks of hand gestures and speech: left child of dyad L2.1 episode 11

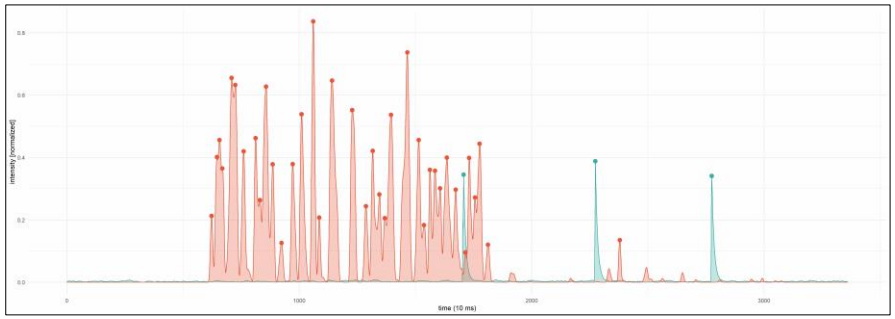


Figure 46

Peaks of hand gestures and speech: right child of dyad L2.1 episode 11

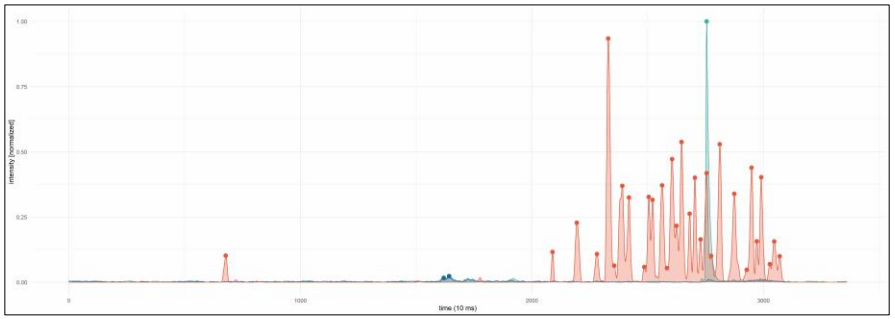


Figure 47

Peaks of hand gestures and speech: left child of dyad L2.1 episode 12

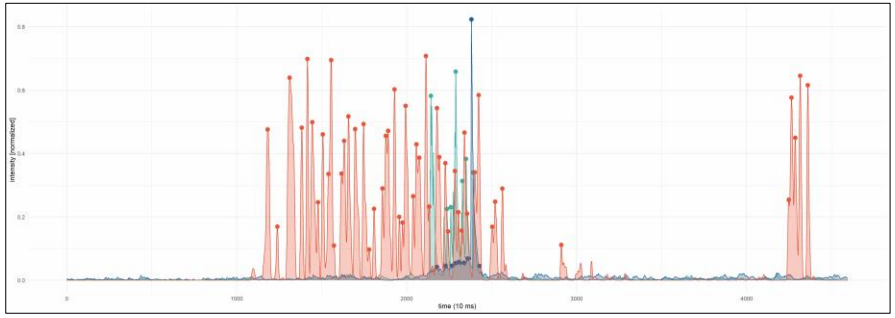


Figure 48

Peaks of hand gestures and speech: right child of dyad L2.1 episode 12