Research Article

Jarret Geenen*

# Objects and materiality in pragmatic development: Here-and-now to then and-there

https://doi.org/10.1515/mc-2020-0020

**Abstract:** Language acquisition involves more than learning how to produce words in complex strings. It involves a diversity of aptitudes about how, when, with whom and in what way to use language abilities. While it is acknowledged that these skills are learned through social interaction (Blum-Kulka, S. (1997). *Dinner talk: cultural patterns of sociability and socialization in family discourse*. Lawrence Erlbaum Associates, Inc, Mahwah, NJ; Rogoff, B. (2003). *The cultural nature of human development*. Oxford University Press, Oxford), our understanding about precisely how they emerge and how they are taught and learned remains preliminary at best. Additionally, much of our understanding is strictly limited to spoken language. The analysis and arguments herein detail the consequentiality of child directed interaction strategies (CDIS) which facilitate non-verbal actions and motivate episodic retrospection, making a tangible link between the current interaction and past experiences. Through a multimodal interaction analysis (Author and Pirini, J. (2020). Multimodal (Inter)action Analysis. In McKinley, J. and Rose, H. (Eds.) *The Routledge handbook of research methods in applied linguistics*. Rouledge, London, pp. 488–499; Norris, S. (2004). *Analyzing multimodal interaction: a methodological framework*. Routledge, London. https://doi.org/10.4324/9780203379493; Norris, S. (2011). *Identity in (inter)action: introducing Multimodal (Inter)action Analysis*. de Gruyter Mouton, Berlin & New York. https://doi.org/10.1515/9781934078280; Norris, S. (2019). *Systematically working with multimodal data: research methods in multimodal discourse analysis*. Wiley Blackwell, Hoboken, NJ; Pirini, J. (2014). Introduction to Multimodal (Inter)action Analysis. In: Norris, S. and Maier, C. (Eds.). *Interactions, texts and images: a reader in multimodality*. Mouton de Gruyter, New York. https://doi.org/10.1515/9781614511175.77) of the practice of showing material objects during interaction, I show that non-verbal action, material culture and the physical world are crucial to developing a certain socio-cognitive pragmatic aptitude. CDIS motivating 'showing' of tangible objects of personal significance may be the non-verbal antecedent of selecting and introducing new topics during interaction. These CDIS defer interactional agency and motivate non-verbal communicative actions more comfortably within the zone of proximal development. Importantly, the materiality of the objects themselves are of fleeting interactional priority. Instead, the objects provide a bridge between materiality in the here-and-now to past experiences in the there-and-then. Facilitating non-verbal actions of showing help motivate explorations of episodic memory by creating a tangible and immediate link within the unfolding interaction.

**Keywords:** Multimodal (Inter)action Analysis, pragmatic development, multimodal pragmatics, objects-in-interaction, object handling

***Corresponding author: Jarret Geenen**, Radboud Universiteit Nijmegen, Comeniuslaan 4, Nijmegen, 6500 HC, Netherlands, E-mail: j.geenen@let.ru.nl

# 1 Introduction

It is no secret that the true site of communicative development for children is in real-time social interaction (Blum-Kulka 1997; Rogoff 2003). From the minute they are born, they are besieged with various forms of communication. Family members and friends wave, speak to, hug, kiss and perform a whole range of verbal and non-verbal actions showing supreme pleasure when these new humans appear to respond with a coo or a smile. This continues well into the first year of life with parents gleaming at the precocious interactive skills of their preverbal infant. It is a communicative smorgasbord and long before children begin using spoken language, they are active and accomplished social actors able to play simple games and communicate a wide range of things without saying a word.

The communicative abilities which emerge prior to language have been a subject of empirical interest for years (see Stephens and Matthews 2014 for review). The pre-verbal period of development has actually been labelled 'pure pragmatics' as children act and interact, construct and interpret meanings without properly knowing how to speak. While there is some debate regarding intention and parents obviously play a role in evaluating actions as 'communicative', few contest that pre-verbal infants exploit non-verbal resources to 'express things'. This has led some to argue that language acquisition proper is crucially dependant on preverbal pragmatics abilities (Brooks and Meltzoff 2005; Mundy et al. 2007) and that interactive aptitudes in this stage lay the groundwork upon which the ability to communicate through words can emerge (Bates 1976; Carpenter et al. 1983).

There is a growing body of evidence which indeed suggests that non-verbal and verbal actions may be complexly intertwined in both development and communication. There are well documented links between hand gestures and language acquisition (Iverson et al. 1994; Ozcaliskan and Golden-Meadow 2005) and some evidence even suggests that gestural behaviours in preverbal infants are predictive of later language development (Igualada et al. 2015). Admittedly, precise details are still somewhat fuzzy, however, a picture is emerging which seems to suggest that non-verbal communicative abilities and language development are to some extent interrelated.

Although pragmatic development does not end with the onset of language acquisition, little attention has been allocated to the role that non-verbal actions may play in the development of higher-order interactional aptitudes. It seems as though after the shift has been made to verbal language, empirical attention shifts as well. This has even led to a more recent call for a need to analyse "the pragmatics of interaction using language" (Clark 2014: 117) recognizing *interaction* as primary and language as simply one of the modes through which it is accomplished. The championing of this reorientation of empirical attention leads to a number of significant questions regarding the development of pragmatic aptitudes which occur through social interaction. If all social interaction is primarily multimodal by nature (Norris 2004) to what extent do other communicative modes figure as consequential in learning to use language appropriately in interaction? To what extent and how do non-verbal forms of communicative action figure in the development of discrete pragmatic aptitudes? What specific role do interlocutors play in this learning process and are there any implicit or explicit strategies employed to facilitate this development?

The analysis and arguments presented herein seek to provide preliminary though partial answers to the questions above in relation to one particular interactive pragmatic aptitude: the ability to select and introduce new topics during interaction. More specifically, I argue that parentally facilitated non-verbal communicative actions of showing material objects and entities during interaction may help children make connections between the immediate interaction and past (or future) experiences. As a result, younger interlocutors begin to recognize that past experiences of individual personal significance may be interactionally relevant. Parental facilitation and the showing of physical objects appear to be a collaborative form of topic introduction (albeit a multimodal form) and are treated as such by interlocutors. This practice, which is more communicatively manageable for the developing interlocutors, may be the non-verbal antecedent to the ability to select and introduce conversational topics through language.

Through a multimodal interaction analysis (Norris 2004, 2011, 2019, Pirini 2014, 2020) of the practice of showing material objects during interaction, I show that non-verbal action, material culture and the physical world are crucial to developing a certain socio-cognitive pragmatic aptitude. More specifically, Child Directed Interaction Strategies (CDIS) which prioritise 'showing' as a tangible, non-verbal communicative action can bridge a developmental gap to more normative interactive 'sharing' through spoken language. The CDISs accomplish a number of interactive and pedagogical goals which facilitate reflection and explicit consideration about topic introductions and the nuances of conversational relevance. As such, these strategies and the co-produced interactive practices they initiate can lay the socio-cognitive ground work for further pragmatic development of communicative strategies undertaken more holistically through spoken language. These communicative strategies are indeed instrumental in teaching developing interlocutors how to introduce new topics during social interactions, what kind of topics are appropriate for selection and various communicative norms regarding interactive responsibilities.

The analysis and arguments regarding the developmental function of this co-produced interactional practice are framed within a consideration about how these one-time situated mediated actions and their developmental potentials intersect with and connect to larger Discourses (Gee 1990) and scales of action (Norris 2017). In addition, the particular site of engagement (Norris 2011; Scollon 2001) is critically considered to articulate how these actions and practices are mediated by and thus shaped by the intersection of social actors, practices and multiple cultural tools of varying materialities. Thus, the analysis seeks to bridge micro interactive mediated actions with macro discourses, practices and developmental trajectories: a task for which multimodal interaction analysis is perfectly suited (Pirini 2014).

Following the discussion of largely conversation analytic work on topic introductions and topic talk, a brief overview of Multimodal (Inter)action Analysis (Author and Pirini 2020; Norris 2004, 2011, 2019; Pirini 2014) as a methodological framework is detailed. The employment of multimodal interaction analysis leads explicitly to a consideration of the social actor and mediational means intersection. Two representative data samples from the long-distance video-conferencing corpus are analysed to explicate the utility of child directed interaction strategies in the teaching (and learning) of various socio-cognitive pragmatic aptitudes followed by a consideration of the specific site of engagement in which the actions unfold and how these actions link to other actions and practices on different time-scales. The connections across actions existing and persisting on different time-scales connect objects to experiences and object-oriented mediated actions in interaction to object-oriented mediated actions in history. The material and immediate intersects with the historical whereby objects give immediate interactive tangibility to the complexes of experience which become the collective topical foci of the exchanges.

## 2 Conversational transitions and topic introduction

While an understanding of the contribution of non-verbal modes of communication is lacking in the domain of topic selection and transition, conversation analytic work has revealed much about the interactive particulars of these pragmatic aptitudes. We have a refined understanding of the structure and form of topic initiations, how people use them and how they are accomplished naturalistic conversation.

Introductions or initiations are typically recognized as a part of "topic talk" (Schegloff 1990: 52) which is normal conversational activity involving the sharing of experiences, discussing of current events or ongoing situations in each others' lives. Topic talk normally involves the selection of various 'mentionables' (Schegloff and Sacks 1973: 301), which are events and experiences that are conversationally relevant. The notion of "mentionables" connects quite explicitly with topics, themes or experiences which are to some extent conversationally relevant as what constitute something that is mentionable is influenced by the immediate exchange, interpersonal relationships and speech genre being produced at that time-place. Thus, topic talk is *about something* but additionally about something that is interpersonally relevant and/or pertinent given the participatory framework, trajectory of the discourse and interpersonal relationships of interlocutors.

Types of topics and time allocated to discussion varies (Barnes et al. 2013) again contingent on interpersonal relationships and various other dimensions of the interactive situation. Additionally, unlike other practices which may occur only once or twice in any interaction, topic talk continuously recurs with multiple discursive segments making up the 'body' of the interaction itself (e.g. Opening, topicT1, topicT2, topicT3, closing). Given that multiple topic talk segments unfold in any conversation, transitional phases have captured considerable interest as this involves multiple interlocutors having to negotiate turns, input and agency to seamlessly move the conversation forward.

Button and Casey (1985) provide a useful classification of two primary ways topic transitions occur: stepwise and disjunctive. Stepwise transitions are characterised by close continuity between aspects of the bordering topics. As a result, stepwise transitions do not necessitate explicit closings and initiation strategies as there is some locally cohesive thread which ties the upcoming topic to the previous. A stepwise transition may occur as interlocutors move from a discussion about an ice hockey game to a discussion about a police investigation involving an ice hockey player from that particular team. Disjunctive transitions do not exemplify such thematic continuity.

Disjunctive transitions often create distinct boundaries between the topics which unfold during interaction. As Svennevig (1999: 188–199) points out, in disjunctive transitions, topics often close when interlocutor contributions wane or there are simple responses rather than progression forward. These simple responses may come in the form of content summation or laughter but pivotally, do not contribute to extension about the topic at hand. Disjunctive transitions, unlike stepwise, feature explicit forms of topic closure as well as more explicit means to introduce or initiate further topic talk. These transitions also exemplify a thematic discontinuity as there is no thematic thread which can be seen woven between the two topics. An example is when talk about the Ice Hockey game ends in lamentation about the continuous poor performance of the team as a 'closing' strategy then moves to events which may have been discussed a few weeks back: i.e. "did you ever end up looking at that motorcycle that was for sale down the street?". Here, summation and explicit introduction coupled with discontinuity between subjects exemplifies a disjunctive transition.

While both transition types lead to the broaching of different subject matter, the later is perhaps the more treacherous to negotiate (see Barnes et al. 2013 in relation to aphasiac speakers). While not a focus in the work on transitions, both types involve what might be characterised as a negotiation of *conversational relevance* as the new topic should satisfy either local discursive relevance in stepwise transitions (i.e. connect to the previous content in some way) or global interpersonal relevance in disjunctive transitions (i.e. Connect to the interests of or relationship with the other interlocutors). Stepwise transitions present a number of challenges pertaining to topic selection which is thematically cohesive. The latter, however, requires more abstract consideration of a whole host of potential topics and a determination about whether the topic itself is interpersonally relevant given various interactional particulars. The topic also requires explicit introductory methods which to some extent signal its interpersonal relevance so as to elicit expansion strategies from other interlocutors.

While previous work has carefully detailed the collaborative nature of topic transitions (Geluykens 1993), distinguished between types of transitions (Button and Casey 1985), refined our understanding about practices which emerge in transitions (Svennevig 1999) and detailed some difficulties associated with cross turn cohesion and cross turn relevance in younger children (Dorval and Eckerman 1984; Keenan 1974), we have very little understanding about the ontogenesis of these communicative practices. While it appears as though topic transition is complex and a higher order pragmatic aptitude, the acquisition process of this aptitude remains unknown. While it is acknowledged that development occurs through interaction more generally, how adult interlocutors specifically teach and/or facilitate the development of the pragmatic aptitude remains at question. We are additionally without any theorization regarding the socio-cognitive transitions necessary for the development of such an aptitude and how this might occur in child ontogenesis.

The analysis below hones in on one interactional practice which I argue may lay the groundwork for the emergence of the ability to autonomously introduce, through disjunctive transitions, topics of interactional relevance which pertain to significant epistemic experiences during interactions with spatially distant family members mediated by video-conferencing technology. The analytical interest is twofold: on strategies which

recur in the corpus employed by more knowledgeable and pragmatically mature social actors called Child Directed Interaction Strategies (CDIS) and how these strategies contribute to learning and developmental trajectories which form the foundation for the pragmatics aptitudes discussed above. Obviously, an accurate detailing of the developmental process would require extensive longitudinal data collected over many years. The discussion herein makes no such claims or provides evidence in support of such claims. Rather, I argue that the interactive practices which are the phenomenological focus of the analysis herein function to promote certain socio-cognitive transitions upon which these pragmatic aptitudes may emerge.

A central characteristic of the practice under analytical scrutiny but also a fundamental aspect of social interaction in general is that it is by its very nature, multimodal. As such, approaching the phenomenon inductively requires an analytical framework which can be used for the analysis of multiple modes of communication as they are employed in real-time social interaction. Multimodal (Inter)action Analysis (Norris 2004, 2011, 2018; Pirini 2014; Author and Pirini 2020) provides precisely such a framework.

## 3 Multimodal (Inter)action Analysis

Any comprehensive analysis of real-time social interaction necessitates an analytical framework and transcription system which is able to account for the multiple communicative modes through which human beings interact without allocating implicit prioritization on any single mode *a priori*. Multimodal (Inter)action Analysis (Norris 2004, 2011, 2019; Pirini 2014) facilitates the analysis of real-time audio-video data by incorporating empirical and analytical insights generated in various multimodal domains and unifying them through the employment of a single ecological unit of analysis: the mediated action (Norris 2004, 2009, 2011, 2019, Scollon 1998, 2001, Wertsch 1991, 1998). It is widely recognized that human beings utilise a vast diversity of non-verbal communicative modes in coordination with spoken language during social interaction. Multimodal (Inter)action Analysis adopts insights from the study of spoken language (Chafe 1994; Sacks et al. 1974; Tannen 1984; Scollon 1998, 2001) gesture (Kendon 2004; McNiell 1992, 2007), interactive proxemics (Hall 1959, 1966), gaze (Kendon 1990; Kidwell 2005; Rossano 2012), body posture (Scheflen 1964), layout (Hall 1959, 1966) and object handling (Norris 2004; Author 2013, 2017, 2018) and integrates them in an analytical framework with a single unit of analysis.

The primary methodological hurdle when approaching the analysis of real-time social interaction pertains to employing a single unit of analysis. Spoken language, gesture, gaze and proxemics are all modes through which human beings act in real-time interaction, however, they are all structured in very different ways. For instance, as McNeill (1992) articulates, gesture is global and synthetic which means unlike the compositionality of spoken language, a single gesture and the meaning expressed therein is not a result of the individual structural components that make up the gesture itself. One does not combine finger shape and location, stroke trajectory, stroke location in gesture space, speed, rhythm and handshape to get gesture X. Instead, the gesture in its entirety in coordination with other modes (spoken language for instance) may provide insights into the communicative salience of stroke speed but this is in a top down and not bottom up fashion. Alternatively, spoken language is combinatorial and a morphological or lexical alteration in any utterance can indeed affect the meaning of the entire utterance in a bottom up fashion. Multimodal (Inter)action Analysis overcomes this methodological issue by employing the lower-level action as the smallest unit in analysis. A lower-level action is the smallest pragmatic meaning unit of any communicative mode (Norris 2004). In the mode of spoken language this is an utterance, in gesture this is a stroke or a stroke hold, for posture it is a postural alteration etc. This enables the collective analysis of all communicative modes while maintaining a single unit of analysis accounting for the analytical and theoretical insights provided in specific modal domains.

MIA explicitly acknowledges the (Vygotsky 1978) notion of mediation which was later taken up in Wertsch's (1991, 1998) Mediated Action Theory in socio-cultural psychology and subsequently by Scollon (1998, 2001) in Mediated Discourse Theory. From the two later perspectives, the mediated action which constitutes the social actor acting with/through mediational means is the most fruitful ecological unit of analysis as it maintains the inextricable tension between social actor and mediational means. All concrete

social action is mediated whether this be through psychological tools, physical tools or a complex combination of the two. Acknowledging that all action is mediated and prioritising this within the unit of analysis maintains as much of the individual, socio-cultural, institutional and historical complexity of the whole phenomenon within the analytical unit itself. With a specific view toward the multimodality of all social interaction, Norris (2004) operationalised this theoretical unit for the analysis of real-time social interaction which occurs through multiple modes simultaneously. The lower-level action as the smallest pragmatic meaning unit of any mode implicitly prioritises the individual, socio-cultural, historical and institutional permeations in all mediated action while being functionally useful for analysis across communicative modes.

All mediated actions unfold in specific sites of engagement (Norris 2011; Scollon 1998) which are the real-time windows opened up at the intersection of social practices and mediational means. The site of engagement acknowledges the "real-time, irreversible, and unfinalizable nature of social action" (Scollon 1998: 4) insofar as mediated action is not a category or class of action but a one-time unique material moment in history. Indeed, any mediated action may materially resemble others across various time scales and are linked to social practices but the actions themselves uniquely unfold in some specific site of engagement.

Chains of lower-level actions at specific sites of engagement build higher-level actions which are definable as larger co-produced actions with socio-culturally acknowledged and definable beginnings and endings. Various co-produced chains of lower-level actions may contribute to the construction of a dinner party, a casual conversation, a drive to the store etc. Central is that these higher-level actions are built through lower-level actions. In turn, higher-level actions can be multiply embedded. The same complexes of lower-level actions can also contribute to building multiple higher-level actions on differing time-scales simultaneously.

The notion of scales of action (Norris 2017) acknowledges that although actions unfold sequentially in time and space, actions can and often do connect to various other historical or ongoing higher-level actions. For instance, the unfolding of a conversation in an automobile is mutually influenced by the scale of action definable as a drive to the supermarket. The concrete lower-level mediated actions taken at that site of engagement are undeniably affected by its situatedness in a larger scale of action which involves motor vehicle operation. However, the conversation during the drive may connect to and thus build the higher-level action of planning a large dinner party for friends and family. The actions taken at that site of engagement both link and build other higher-level actions occurring on other time-scales. So, while the actions which build the conversation during the drive are indeed embedded in and affected by the higher-level action of driving to the store, they may also be influenced by and additionally build the action of planning a dinner party for friends and family: a higher-level action which extends to and through the conversation during the drive. As described in more detail later, these connections to other scales of action and the influence of those other actions permeates through and affects the actions emerging at that site of engagement. Thus, the lower-level action of showing a toy, can and does connect to a multitude of other actions on different time scales like receiving the toy as a gift, or playing with it during a favorite game.

Multimodal (Inter)action Analysis (Author and Pirini 2020; Norris 2004, 2011, 2018; Pirini 2014) thus provides an analytical framework for the inductive qualitative analysis of real-time social interaction occurring through multiple modes of communication while acknowledging the influence and affect of individual-socio-cultural, historical and institutional trajectories which undeniably affect the materiality of social action in situ.

## 3.1 A note on transcription

As with many ethnomethodologically oriented qualitative analytical methods, MIA acknowledges the dualistic nature of transcription and analysis (Bucholtz 2000) as well as how theory and ideology are exemplified through transcription methods. Given that transcription methods are analytically and theoretically informed, the practice itself is not and can never be simply a means to accurately represent the phenomenon of interest in a static textual manner.

An MIA approach necessitates employing visual transcription method which is realised through the analytical procedure. As such, the transcripts which are frames from the audio-video data with textual

overlays, numbering and time signatures are intended to represent the specific lower-level actions undertaken at a particular site of engagement. Any verbiage through the mode of spoken language is additionally represented as an overlay on top of the frames with speakers indicated by the use of colour. The composition of the verbiage aims to represent the precise materiality of the auditory qualities of speech. Thus, variation in volume is represented through font size alteration and into national patterns indicated by the trajectory of the reading path. Rising intonation through curvature upwards and the reverse for falling. Spelling seeks to maintain intelligibility through conventional orthography while being modified to indicate auditory specificities without employing actual phonetic transcription (see Author, Mateleau and Norris *forthcoming* for detailed discussion of visual transcription).

## 4 Project description and data set

The data set referred to herein was generated for the project Video-conferencing: How is family interaction changing which includes 17 families and over 80 participants. The research questions guiding the ethnomethodologically based audio-video data collection were multiple but the primary interest was on the ways young children (>5) act, interact and are incorporated in, long distance familial video conferencing sessions. The data was generated using multiple high-definition audio-video cameras simultaneously but given that participant groups were separated by great distances, real-time audio-video footage was only collected with one of the interlocutor groups engaged in the video-conferencing interaction. Screen recording software was therefore employed to capture the incoming audio-video stream from that particular location.

One potentially pertinent empirical note is that a single laptop was used for the generation of the audio-video data set comprising the capture of the incoming video-conferencing stream. This methodological decision was practically motivated as alternative methods would involve participant groups having to install, operate and manage this data set individually. Semi-structured interviews were also conducted with participant groups and the research team included questions pertaining to this methodological factor as a means to determine whether using a different technological device had a qualitative effect on the interaction itself. Interview data suggests that this was not an issue insofar as it did not alter the interaction itself in ways that participants could reflexively articulate. In some cases, this was even evaluated as positively affecting their experience insofar as some participant groups claimed that the laptop provided location flexibility that their traditional desktop computers otherwise would not. The lack of influence of this methodological decision was also confirmed phenomenologically as the data set does not include interactional infelicities which appear connected to the device itself.

## 5 Topic transition: Whispered and demonstrative interrogatives

Elsewhere (Author 2017), the frequency at which objects, entities or material components of the interactional surrounding become salient and the collaborative focus of attention has been extensively detailed in relation to the affordances for multimodal identity production. Therein, details regarding the diversity of introduction strategies are explicated in full. For present purposes, introduction and initiation of these embedded higher-level actions are supremely pertinent as the introduction of physical objects or material entities emerge during disjunctive transitions (Button and Casey 1985) which are demonstrably representative of interactive segments with waning conversational contributions including summation, laughter and extended silence (Svennevig 1999).

Figures 1 and 2 below exemplify the points at which interactional trajectory shifts and the collective focus of the social actors involved in the multiparty interaction alters toward a physical object, entity or some subsidiary property of the aforementioned. The representative samples from the video-conferencing corpus exemplify the normative manner in which physical entities ‘enter the fray’ so to speak. In many cases, the momentum of the interaction slows and the previous ‘topic talk’ segment closes giving way to a transition. As

revealed through analysis, the embedded higher-level actions which feature specific conversational topics close according to normative strategies outlined by Svennevig (1999). The transitions exemplify extended silences and social actors appear to be oriented toward determining a new topic – a subject or experience which is 'mentionable' given the conversational trajectory, interlocutors involved, interpersonally history and interpersonal interests.

The transitions and extended silences create conversational exigencies insofar as the social actors involved are required to move the conversation forward with further topic talk or end the interaction through some closing and departing strategies. At these sites of engagement (Norris 2011; Scollon 2001) parents or adult peers take the lead in initiation strategies, but the topics selected and the precise manner in which they are introduced provides explicit scaffolding to the younger social actors involved in the multiparty video-conferencing interaction. It is here that tangible physical objects are identified and introduced within the interaction.

The most salient means through which material objects and artefacts enter the ongoing interaction is through explicit facilitation on behalf of adult interlocutors through CDIS of noticing, explicit/implicit requests or interrogatives. Noticing typically involves explicit mention of some material element visible to interlocutors and is typically precipitated by those not in the same material time-place as the children themselves. Explicit requests can be made by any social actor involved in the multiparty interaction whereas interrogatives as a means for object introduction tends to be propagated on behalf of adult caregivers or guardians. Each form of initiation may arise for any number of interactional exigencies and be precipitated by a diversity of interpersonal motivations, however, they all seem to serve the same function within the interaction: provide explicit direction to the child about something that is potentially interactionally relevant. More importantly, these CDIS which arise during disjunctive transitions, encourage developing interlocutors to explicitly assess the interactional relevance of the proposed subject matter therein facilitating the explicit judgment about conversational relevance (Brown and Yule 1983: 83; Grice 1975).

In Figures 1 and 2, the caregivers (in this case parents) of the younger social actors who are currently also ratified participants (Goffman 1979) within the multiparty interaction, recognize a stagnation in the temporal progression of the interactional exchange unfolding at the site of engagement. The whispered and demonstrative interrogatives appear motivated by marginal contributions from the children themselves coupled with the extended interactive silence. These CDIS simultaneously seek to motivate more autonomous engagement on behalf of the children themselves.

## 5.1 The tooth

In Figure 1, two younger girls and their mother are talking to Matt (their uncle). At 0.03.48 in frame 1, the previous higher-level action is ending. Isabel is finishing providing a humorous explanation regarding the particular reason for a specific Christmas wish of Sara, her younger sister. Matt is well aware of a disposition Sara has acquired in recent months which involves pretending to be a fictional character. As such, he finds Sara's Christmas wish quite humorous. Isabel's explanation of Sara's wish sends Matt into a fit of laugher. In frame 2, Matt's laugher slowly subsides and he remarks through spoken language 'amazing', reaffirming the amusement and his surprise at Sara's dedication to pretend which has led to her to ask for something funny for Christmas. The extended laughter and following silence signal a topic closing (Svennevig 1999).

Between frames 3 and 5 extending from 0.03.50 to 0.03.54, there is a 3.5 s silence. In frame 2, Matt's final remark is made and in frame 3, Sara shifts her gaze and the orientation of head direction up and to the left towards her mother. Simultaneously, she alters her posture slightly and opens her mouth producing a visible smile. These multiple and coordinated lower-level actions through the modes of gaze, posture, head movement and facial expression suggest her attention has shifted momentarily to her mother by virtue of the modal complexity (Norris 2009). In frame 4, she again undertakes multiple lower-level actions altering postural orientation and head direction and displaying a visible smile, seemingly pleased at the laughter she has motivated. In frame 5, after the silence in which there are various lower-level mediated actions involving subtle

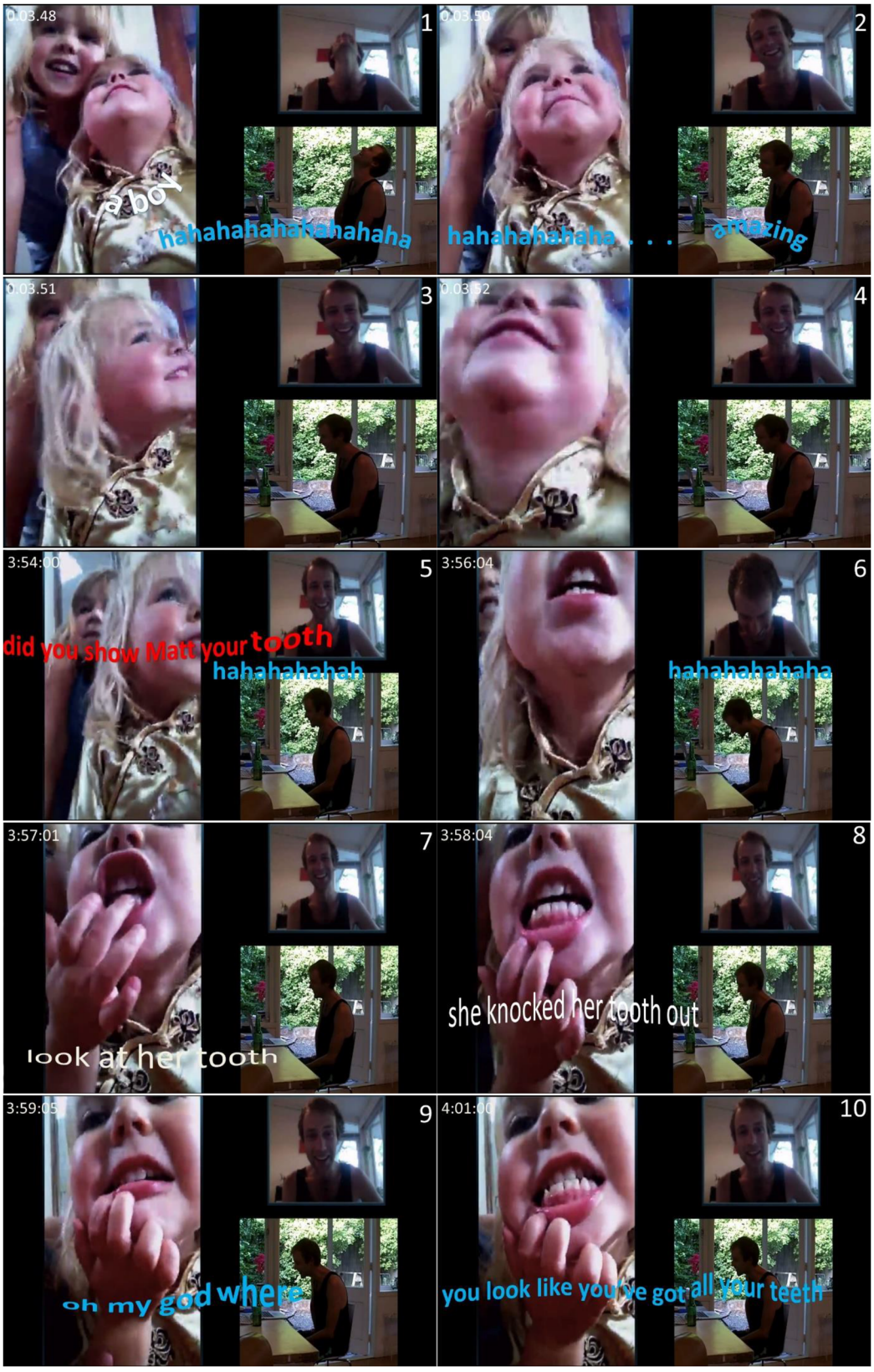


**Figure 1:** Did you show Matt your Tooth?
Figure partially reproduced from Author (2018).

postural shifts, gaze shifts and head movements, Sara's mother produces a whispered interrogative, asking Sara, 'did you show Matt your tooth?'.

The initiation of the lower-level action undertaken through the mode of spoken language attracts Sara's attention immediately. Evidenced in frame 5, Sara undertakes multiple lower-level mediated actions through the modes of posture and gaze indicating that her attention is again with her mother and momentarily directed away from her Skype interlocutor. Approximately 2 s following her mother's whispered interrogative, Sara demonstrably alters her posture, gaze and proxemics and opens her mouth to show her tooth to her Skype interlocutor. Sara thus, considers her mothers interrogative, interprets it as an invitation or suggestion that she should, and obliges in displaying a tooth which is later revealed to be severely damaged due to an accident. Following the multiple lower-level actions which build the showing of the tooth, the tooth and the manner in which it was damaged persists to be the thematic focus of the interaction. This higher-level action and new topic which is initiated with the whispered interrogative, is built further through Sara's interpretation of the interrogative resulting in the 'showing' of her damaged tooth further perpetuating the higher-level action with the tooth as the thematic focus.

The participation framework, proxemics of social actors to each other and to the selected devices used for video-conferencing creates an interactional affordance which makes possible this whispered interrogative. In Figure 1, Sara's mother is completely out of the video-conferencing frame and thus not visible. However, a faintly audible whispered interrogative can be heard asking "did you show Matt your tooth". This interrogative signifies the beginning of a new higher-level action wherein the topical focus is the tooth and an accident which occurred leading to the damage. The precipitating factor for the topic alteration and whispered interrogative is the topic transition point exemplified by the 3.5 s silence which involves various lower-level actions through non-verbal modes, but does not involve any further elaboration on the previous subject matter and no other social actor initiates a new topic opening.

## 5.2 The slug

Figure 2 exemplifies a similar phenomenon differentiated by the specifics of the ongoing thematic foci of the interaction and the precise material nature of the interrogative articulated on behalf of the child's parent. In frame 1 at 1.01.35, Megan is providing affirmative confirmation through the mode of spoken language to an interrogative posed by her mother about a toy which is currently visible in Megan's hands. Moments before, the toy bow and arrow which shoots paper straws and Christopher's marksmanship was the collective focus of the social actors involved in this multiparty video-conferencing interaction. At the closing of this higher-level action and immediately following the affirmation provided by Megan there is a distinct auditory silence between frames 2 and 4 in the transcript running from 1.01.35 to 1.01.40 where the silence is punctuated by an elongated conjunction in the form of 'aaand'. The lower-level action undertaken through the mode of spoken language functions as a bridge between topical foci but the atypical and marked duration of the initial phoneme is simultaneously indicative of an attempt to determine or decide upon a potential new topic to further extend the interaction. It may be that Megan is entertaining various recent events or experiences and trying to determine which of them or if any of them are indeed 'mentionable' as a means to further the interaction.

In frame 5 at 1.01.41, Megan undertakes multiple lower-level actions simultaneously during the initiation of a demonstrative interrogative posed to Christopher. Megan alters her postural orientation, dropping her right shoulder toward Christopher and in doing so, closes the proxemic distance between them. Her right hand, with which she has just placed the toy bow on the table is placed on the table in close proximity to Christopher as she further closes proxemic distance by leaning toward him. Even prior to the verbal address, multiple lower-level actions show that Megan's orientation is no longer with the skype interlocutor solely, but rather, seeks to engage Christopher directly. Immediately after, she addresses Christopher by name, followed by a very brief pause in speech production which is followed by a demonstrative interrogative 'do you want to show grandma your pet'. Approximately 6 s elapse from the point of affirmation in frame 1 to the initiation of the

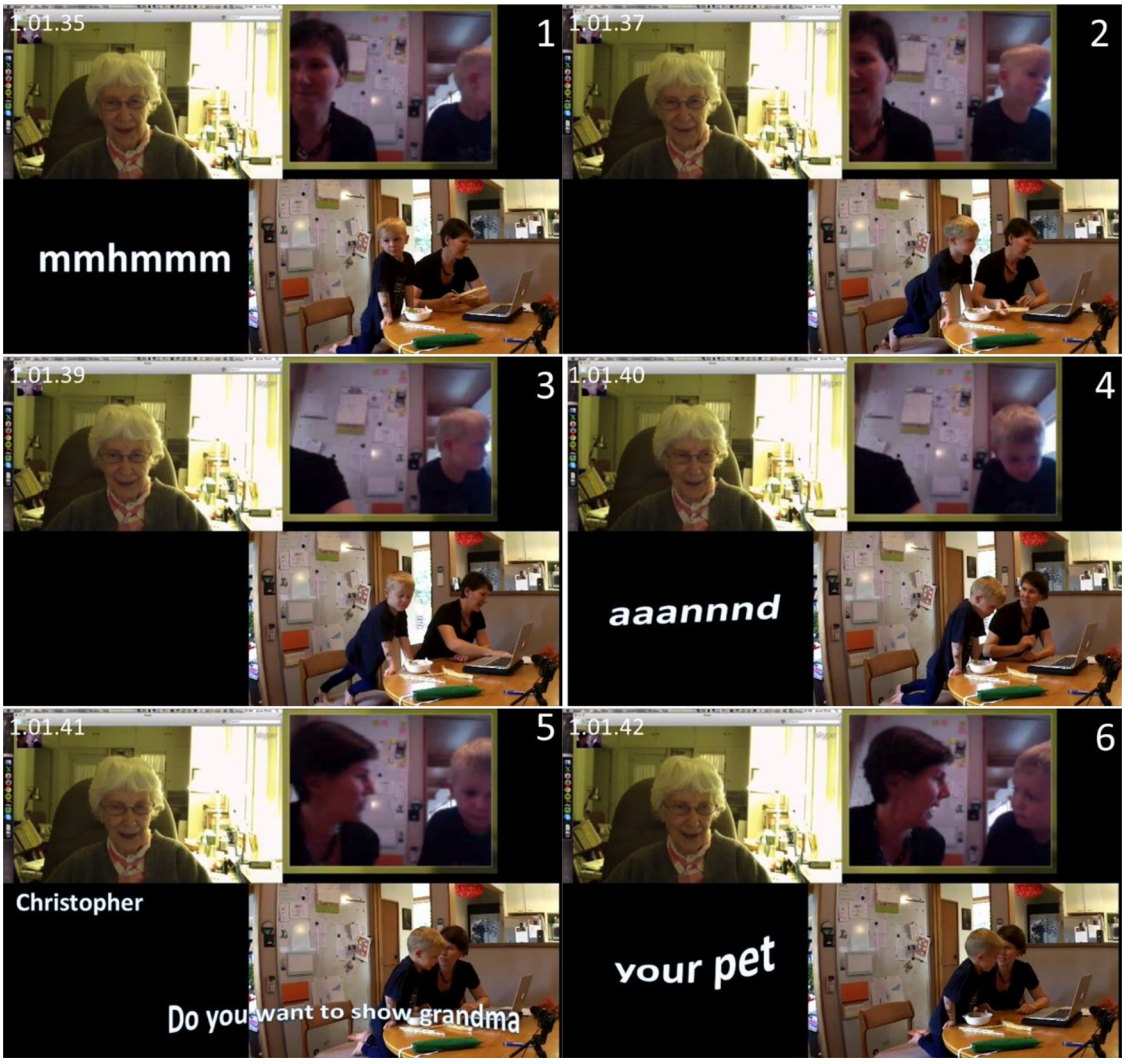


**Figure 2:** Do you want to show grandma your pet?

interrogative in frame 5 punctuated only by the elongated conjunction. Taken together, it seems as though the waning pace of the multiparty interaction has led to Megan's postulating the potential introduction of a new thematic trajectory for the interaction. At this point, she addresses Christopher directly and inquires about whether or not he would like to show his grandmother his pet.

After the marked silence extending approximately 2 s, Christopher responds affirmatively through spoken language to his mother's interrogative. Immediately following, Christopher pushes himself back from the table altering his balance and proxemic distance between himself and the table. The alteration in posture and proxemics appear to be in anticipation to depart and retrieve the 'pet'. Simultaneously, his mother, through a coordination of a gaze shift, deictic gesture and a whisper 'over there' indicates deictically toward the location of the pet in question. Christopher is well aware of the location of his pet and the indicatory deixis is not simply the identification of location but rather could be seen as further specifying the illocutionary nature of the initial interrogative. While Christopher responds affirmatively, his lack of immediate departure and retrieval may have motivated his mother's additional specification that he should indeed retrieve the pet from its current location.

Upon returning with a small canister, in frame 13 at 1.01.53, Christopher extends his left arm up and out with the canister containing his slug in his hand. Seemingly, due to the proxemics constraints or uncertainty

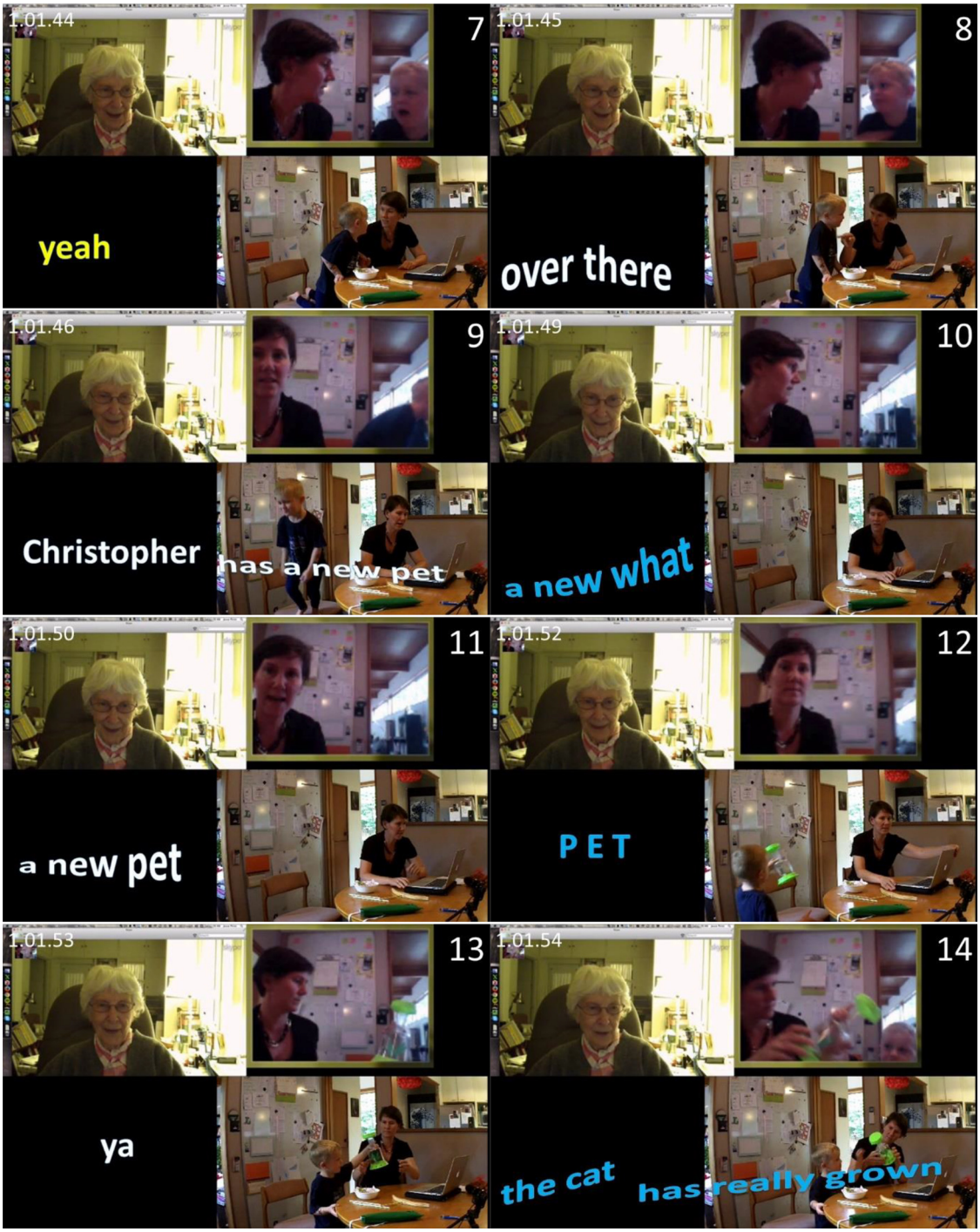


**Figure 2:** (Continued)

regarding grandma's ability to see the small animal, Christopher's mother helps by holding it even closer to the camera in a position wherein grandma can see the slug. Christopher's new pet slug continues to be the thematic focus of the conversation with various questions arising from both mom and grandma about names, particular type of slug and how it was found.

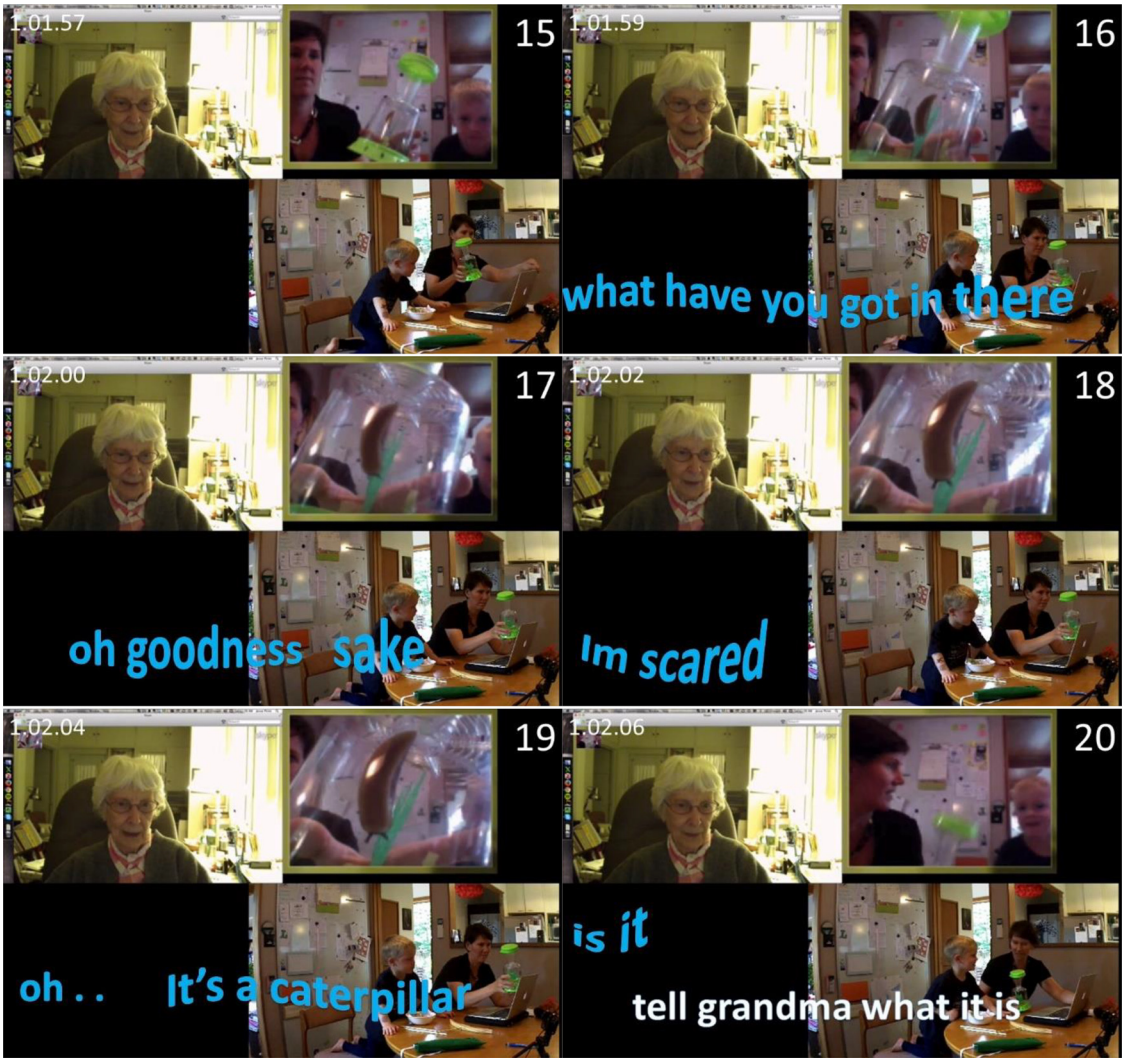


**Figure 2:** (Continued)

In Figures 1 and 2, there are two lower-level actions undertaken through the mode of spoken language in the form of interrogatives directed explicitly at the children though they indeed differ in materiality. In Figure 1, the interrogative is whispered whereas in Figure 2 the audible quality is more representative of the normative speech production though there is a hint of motherese in both. Additionally, both initiation sequences explicitly prioritize the child within the participant framework. Not only are the children the addressed participants, in both cases it is clear that attention is with the child. In Figure 2 this is obvious through the lower-level actions of postural alteration, proxemic alteration, gaze shift and direct address through spoken language. In Figure 1, this is exemplified both through the precise nature of the interrogative itself coupled with how the mother's management of the tablet through object handling maintains visual focus on the children themselves.

Importantly, the interrogatives function as topic proposals. The children are directly addressed and thus, are explicitly positioned as responsible for determining appropriateness of the proposals. Despite the fact that the interrogatives are posed as yes/no questions wherein adjacency pair expectations would predict confirmation or negation, the children do not treat them so simply. The children treat the interrogatives as topic proposals. While in Figure 2, Christopher indeed provides what could be considered a preferred response to the

yes/no interrogative, his resulting departure and return with a small plastic capsule containing his pet suggests that he either accurately interpreted the interrogative as a thematic proposal at its time of production or the deictic specification helped him recover the implication. Either way, it is explicitly clear that he does not treat his mother's interrogative as an inquiry regarding his desire. Similarly, Sara does not treat her mother's interrogative as requesting information. In both instances, the parents' topic proposals are explicitly acknowledged as such by the children.

The interrogative nature of these lower-level actions through spoken language is also central. As interrogatives rather than declaratives, commands or demands like "show Matt your tooth" or "show grandma your pet", these thematic prompts produce a space for explicit consideration and negotiation on behalf of the children therein facilitating the explicit judgment about conversational relevance (Brown and Yule 1983: 83; Grice 1975). This is further evidenced in multiple data segments in the corpus where children refuse to accommodate the proposals with either inaction or demonstrable negation through spoken language and explicit head shaking signifying 'no'. These CDIS additionally seek to allocate interactional agency in a sense nominating the children to explicitly consider the thematic proposal. Among other things, this explicitly though gently motivates a conscious consideration regarding the interactional relevance of the object or entity lexicalised in the interrogative forms.

## 6 Scales of action: Individual-interpersonal interconnectedness

While the analytical focus above is centred on the CDIS which aim to encourage involvement in the interaction, hand over a degree of interactive agency and produce the potential for explicit consideration regarding the topical proposals, equally important are how the lower-level actions of showing unfolding at a specific site of engagement build multiple higher-level actions simultaneously. It is precisely at this intersection where the practice of showing connects to and brings to the fore multiple histories and experiences with and through the objects themselves.

The lower-level actions of showing which emerge at these sites of engagement crucially build the higher-level action that is the immediate interaction with distant family members. Those lower-level actions also connect with larger Discourses (Gee 1990) of family togetherness and history contributing to building and maintaining familial bonds across time and space. However, the actions of showing simultaneously build and connect to other historical and persisting higher-level actions. These interconnections which transcend time-scales and location are pivotal in recognizing how the CDIS and practices of showing connect objects or entities in the here and now to experiences in the recent or distant past and potential future. Through these connections, objects in the here and now and actions of showing which are more interactively manageable for developing conversationalists help bridge the gap to verbal articulations about experiences in the past or future. These practices lay the ground work for and facilitate the development of socio-cognitive aptitudes which will be crucial for extended epistemic recounts or provisional future plans/activities undertaken more holistically through spoken language during interactions in later life.

At these sites of engagement, multiple higher-level actions existing and persisting on different time-scales intersect. The lower-level action of showing the slug consequentially builds the higher-level action of interacting with grandma but simultaneously builds and is a part of the higher-level action of stewardship of the slug for the time it remains a pet. The action of showing builds and is a part of two identifiable higher-level actions which exist on dramatically different time-scales. The production of this action and the interrogative which motivated it bring pet ownership, finding and housing the slug, feeding it well and learning about it into contact with the higher-level action of interacting with grandma. As such, the CDISs and the 'showing' bring those experiences into this immediate site of engagement – teaching young children the interpersonal relevance of those experiences through practices more holistically situated in their zone of proximal development.

The actions of showing Sara's tooth builds the higher-level action of damaging and caring for the damaged tooth and this includes a trip to the dentist which occurred at some recent moment in familial history. The lower-level action of showing also builds the higher-level action of the video-conferencing interaction with her

cousin. These connections across and through time-scales in relation to other scales of action are precisely how aspects of individual experiences at different time-places are introduced as potentially interpersonally relevant within the interaction unfolding at this particular moment.

The lower-level actions of showing which unfold at these sites of engagement contribute to building the higher-level actions of long-distance, extended family interaction. The CDIS and resulting lower-level actions of showing additionally connect to recent experiences or other scales of previous and ongoing actions. It is in this way that the scales of action of more individual/personal significance become situated within the interaction creating intersections across and between individual epistemic significance and interpersonal interactional relevance. The CDIS motivate an agentive and explicit consideration regarding whether the children want to share through the actions of showing, some object or entity. However, it is not the object or entity itself which is significant, nor is it really about 'showing' the object or 'sharing' the object through actions of showing. The practice of showing which is comfortably within the interactional skill set of the children is used as a means to connect the immediate interaction to interpersonally significant past experiences and action. It is precisely in the interconnections across scales of actions which emerge in the CDIS and resultant showing where individual experiences become interactionally relevant. Thus, the CDIS which focus on material objects may function as a supremely useful strategy to contribute to explicit consideration about what individual experiences constitute topics of interpersonal relevance within the ongoing interactions.

## 7 Objects to experience to 'the now'

As exemplified in the representative samples analysed herein, CDIS in the form of interrogatives accomplish multiple functions within these multiparty video-conferencing interactions. Their motivation as is evidenced in the data corpus likely satisfies two interactional exigencies; to explicitly engage children more agentively in the interactional exchange and to propose potential thematic trajectories during disjunctive transitions exemplified by temporal stagnation. The first is demonstrated in both the model density exemplified in the higher-level actions undertaken at those specific sites of engagement as well as the directionality of address. The second is evidenced by the marked and elongated temporal and rhythmic stagnations which tend to precede the complexes of lower-level actions through which the interrogatives are made.

The interrogatives typically take as their thematic focus an object or material entity either within or very near to the immediate physical surround often simultaneously implicating the physical bodily action of showing in some way or form. Often times the actual lexical verb 'show' appears in one of its various forms but alternatives such as requests to 'see' are commonplace within the corpus as well. Given the visibility affordances of video-conferencing technologies it is likely the requests to see or show aim to engage the children in the ongoing interaction through practices more within the zone of proximal development and less contingent on linguistic or conversational maturity.

There is little material continuity across the diversity of objects and entities which come to occupy interactional priority through CDIS to see or show. However, whether the thing itself is a body part, a toy, a pet, a craft, a video-game console or graduation diploma (Author 2017), the object or entity is always individually and/or personally significant to the children themselves. The significance personally can range dramatically. The thing may be a badly damaged tooth which resulted from a physical accident and led to an unexpected visit to the dentist. It may be a slug which was found, captured and is being dutifully cared for as a pet. It may be a graduation certificate which signifies years of hard work and scholastic achievement or a bow and arrow with which the child has developed excellent marksmanship. Most importantly, the object itself is rarely (if ever) significant simply as an object. Instead, the explicit facilitation through CDIS of the practice of showing is an interactional gateway to expansion and expert adult interlocutors consistently prompt the developing conversationalists to provide further commentary about the precise significance of the object itself.

Finally, the object or entity, its materiality, aesthetic qualities, colours and textures do not become the collaborative focus of attention during the unfolding higher-level actions built in the multiparty interactions. Instead, social actors consistently, through various questions and requests, aim to elicit information about the

experiences which connect to the object, knowledge about the object or memory of associated experiences. In this way, the object functions more as a bridge to epistemic recounts, details about likes/dislikes or specialized knowledge gained through learning and experience.

## 8 Conclusion

Children face a diversity of linguistic hurdles during ontogenesis. Beyond the construction of grammatically correct and auditorily coherent sentences children must also learn to employ communicative aptitudes in the appropriate manner, at the appropriate time and with the appropriate people. While considerable amounts of work have charted the ages at which particular pragmatic milestones are achieved, less is known about the real-time acquisition of these aptitudes which occur in situated social interaction.

Based on the analysis above, I argue that CDIS which prioritise material objects and entities are implicit pedagogical strategies which assist in scaffolding the recognition of what constitutes a relevant interactional topic. In line with socio-cultural positions on pragmatic development stipulating that children acquire linguistic skills in real-time social interactions in their culture (Blum-Kulka 1997; Rogoff 2003), I contend that these CDIS and associated practices are pivotal in multiple ways. Despite their interrogative form, the CDIS are actually proposals about potential topics and are indeed treated as such by the children themselves. As interrogatives, the proposals are mitigated, simultaneously creating a space for explicit consideration by the children and this minimizes obligation while also increasing the interactional agency for the younger social actors. As a result, the developing conversationalist is to some extent sharing responsibility for new topic initiation and thus, co-producing this interactive practice with the help of expert adults. This suggests that they may also be learning more generally that they indeed *have* shared responsibility for topic initiation during multiparty interactions.

The CDIS prioritising showing or seeing material objects and entities seek to motivate interactional practices more squarely within the zone of proximal development Vygotsky 1978; Wertsch 1985 for the younger interlocutors. Within the data set, proficiency factors coupled with individual, interpersonal, socio-cultural and historical components of comfortability, relationship specifics, last interaction, interest, attention and multiple others result in great variability in the extent to which younger children autonomously and consistently engage in these video-conferencing interactions. However, regardless of idiosyncratic variation, expert adults appear to recognize the potential of the CDIS as a means to engage even the most docile or passive social actors. Obviously, the intersection of various technological tools and social practices at this site of engagement (Norris 2004, 2011; Scollon 1998) make possible these strategies given the mutual visual access afforded by video-conferencing.

Adult social actors select as topic potentials objects or physical entities which are (a) known to be of individual and personal significance to the children through their relation to other scales of action and (b) have expansion potentials pertaining to those complexes of experiences, skills or personal attributes relevant within the particular configuration of the participant framework. CDIS of topic proposals through interrogatives to see or show material entities can thus contribute to a socio-cognitive transition in the developing interlocutors. They require reflective consideration about a particular object or entity, its relation to other scales of actions and experiences and whether these are worth sharing with their interlocutors. What may at a later stage become explicit reflection/consideration about topic appropriateness, originally manifests intermentally Vygotsky 1978, 1987; Wertsch 1985 between social actors in situ. These strategies then create a unique site for the explicit consideration of various individual, interpersonal and historical components which affect the extent to which any potential set of experiences or conversational topoi meet the criteria for being interpersonally relevant.

The CDIS defer interactional agency, provide a topic proposal, open a space for explicit reflexive consideration and facilitate non-verbal communicative action through tangible material objects of personal affinity as a departure point for interactional extrapolation. In these ways, developing interlocutors are collaboratively responsible for interactional trajectory and thus, are being taught that they have such a

responsibility. They are put in a position to make a decision and must reflectively entertain *at least* two courses of action (oblige or refuse). Obliging minimally involves a course of action by showing or going to get the object specified as the topic of the request. The materiality of the object as mediating the practice of 'showing' is first prioritised (as a method to facilitate topic introduction) and then backgrounded giving way to non-tangible memories and experiences which are introduced through the practice of showing. Herein lies the bridge from tangible non-verbal actions in the here and now, to verbalisation of non-tangible experiences of action in some other time-place.

The analysis and arguments above posit the significance of non-verbal communicative actions for developmental pragmatic approaches and create questions about teaching and learning in-situ, the relationship between tangible objects and experience/memory in early childhood and how they become instantiated in real-time interaction. Making connections across time, space and materiality is central to basic processes of language acquisition and it seems plausible that making these connections across time, space and materiality are equally pivotal in more advanced communicative development.

## References


Author (2013). Actionary pertinence: space to place in kitesurfing. Multimodal Commun. 2: 123–153.

Author (2017). Show (and sometimes) tell: identity production and the affordances of video-conferencing. Multimodal Commun. 6: 1–18.

Author (2018). Multimodal acquisition of interactive aptitudes. A microgenetic case study. Pragmatics and Society 9: 518–544.

Author and Pirini, J. (2020). Multimodal (Inter)action Analysis. In: McKinley, J. and Rose, H. (Eds.), *The Routledge handbook of research methods in applied linguistics*. Rouledge, London, pp. 488–499.

Author, Mateleau and Norris. (2021). Visual transcription: a method to analyse the visual and visualise the audible in interaction. *Images Interactives et Nouvelles Écritures*. (forthcoming).

Barnes, S.E., Candlin, C.N., and Ferguson, A. (2013). Aphasia and topic initiation in conversation: a case study. Int. J. Lang. Commun. Disord. 48: 102–114.

Bates, E. (1976). *Language and context: studies in the acquisition of pragmatics*. Academic Press, New York, NY.

Blum-Kulka, S. (1997). *Dinner talk: cultural patterns of sociability and socialization in family discourse*. Lawrence Erlbaum Associates, Inc, Mahwah, NJ.

Brooks, R. and Meltzoff, A. N. (2005). The development of gaze following and its relation to language. Dev. Sci. 8: 535–543.

Brown, G., and Yule, G. (1983). *Discourse analysis*. Cambridge University Press, Cambridge.

Bucholtz, M. (2000). The politics of transcription. J. Pragmat. 32: 1439–1465.

Button, G. and Casey, N. (1985). Topic nomination and topic pursuit. Hum. Stud. 8: 3–55.

Carpenter, R.L., Mastergeorge, A.M., Hearing, C., and Coggins, T.E. (1983). The acquisition of communicative intentions in infants eight to fifteen months of age. Lang. Speech, 26, 101.

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

Clark, E. (2014). Pragmatics in acquisition. J. Child Lang. 41: 105–116.

Dorval, B. and Eckerman, C.O. (1984). Developmental trends in the quality of conversation achieved by small groups of acquainted peers. Monogr. Soc. Res. Child Dev. 49: 1–72.

Geluykens, R. (1993). Topic introduction in English conversation. Trans. Philol. Soc. 91: 181–214.

Goffman, I. (1979). Footing. Semiotica 25(1/2): 1–29.

Grice, P.H. (1975). Logic and conversation. In: Cole, P., and Morgan, J.L. (Eds.), *Syntax and semantics, speech acts*, *3*. Academic Press, New York.

Hall, E.T. (1959). *The silent language*. Doubleday, New York.

Hall, E.T. (1966). *The hidden dimension*. Doubleday, New York.

Keenan, E.O. (1974). Conversational competence in children. J. Child Lang. 1: 163–84.

Kendon, A. (1990). *Conducting interaction: patterns of behavior in focused encounters*. Cambridge University Press, Cambridge, UK.

Kendon, A. (2004). *Gesture: visible actions as utterances*. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511807572.

Kidwell, M. (2005). Gaze as social control: how very young children differentiate 'the look' from a 'mere look' by their adult caregivers. Res. Lang. Soc. Interact. 38(4): 417–449.

Gee, J.P. (1990). *Social linguistics and literacies: ideology in discourses*. Falmer Press, London.

Igualada, I., Bosch, L., and Prieto, P. (2015). Language development at 18 months is related to multimodal communicative strategies at 12 months. Infant Behav. Dev. 39, 42–52.

Iverson, J.M., Capirci, O., and Caselli, M.C. (1994). From communication to language in two modalities. Cognit. Dev. 9: 23–43.
McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*. University of Chicago Press, Chicago, IL.
McNeill, D. (2007). *Gesture and thought*, 1st ed. Chicago, IL: University of Chicago Press.
Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A.V., and Parlade, M.V. (2007). Individual differences and the development of joint attention in infancy. Child Dev. 78: 938–954.
Norris, S. (2004). *Analyzing multimodal interaction: a methodological framework*. Routledge, London. https://doi.org/10.4324/9780203379493.
Norris, S. (2009). Modal density and modal configurations: multimodal actions. In: Jewitt, C. (Ed.). *The Routledge handbook of multimodal analysis*. Routledge, London, pp. 78–90.
Norris, S. (2011). *Identity in (inter)action: introducing Multimodal (Inter)action Analysis*. de Gruyter Mouton, Berlin & New York. https://doi.org/10.1515/9781934078280.
Norris, S. (2017). Scales of action: an example of driving & car talk in Germany and North America. Text Talk 37: 117–139.
Norris, S. (2019). *Systematically working with multimodal data: research methods in multimodal discourse analysis*. Wiley Blackwell, Hoboken, NJ.
Pirini, J. (2014). Introduction to Multimodal (Inter)action Analysis. In: Norris, S., and Maier, C.D. (Eds.), *Interactions, texts and images: a reader in multimodality*. Mouton de Gruyter, New York, pp. 77–92. https://doi.org/10.1515/9781614511175.77.
Rossano, F. (2012). Gaze in conversation. In: Sidnell, J., and Stivers, T. (Eds.), *The handbook of conversation analysis*. John Wiley & Sons, Oxford, pp. 308–329.
Sacks, H., Schegloff, E., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. Language 50: 696–735.
Ozcaliskan, S. and Golden-Meadow, S. (2005). Gesture is at the cutting edge of early language development. Cognition 93: B101–B113.
Rogoff, B. (2003). *The cultural nature of human development*. Oxford University Press, Oxford.
Scheflen, A.E. (1964). The significance of posture in communication systems. Psychiatry 27: 316–331.
Schegloff, E.A. (1990). On the organization of sequences as a source of coherence in talk-in-interaction. In: Dorval, B. (Ed.). *Conversational organization and its development*. Ablex, Norwood, pp. 51–77.
Schegloff, E.A. and Sacks, H. (1973). Opening up closings. Semiotica 8: 289–327.
Scollon, R. (1998). *Mediated discourse as social interaction: a study of news discourse*. Addison Wesley Longman Limited, London.
Scollon, R. (2001). *Mediated discourse: the nexus of practice*. Routledge, London and New York, NY.
Stephens, G. and Matthews, D. 2014. The communicative infant from 0-18 months: the social-cognitive foundations of pragmatic development. In: Matthews, D. (Ed.). *Pragmatic development in first language acquisition*. John Benjamins, Amsterdam/Philadelphia, pp. 13–36.
Svennevig, J. (1999). *Getting acquainted in conversation: a study of initial interactions*. John Benjamins, Amsterdam.
Tannen, D. (1984). *Coherence in spoken and written discourse*. Ablex, Norwood, N.J.
Vygotsky, L.S. (1978). Interaction between learning and development. In: Cole, M., John-Steiner, V., Scribner, S., and Souberman, E. (Eds.), *Mind in society: the development of higher psychological processes*. Harvard University Press, Cambridge, MA.
Vygotsky, L.S. (1987). *The collected works of L.S. Vygotsky, Volume 1: problems of general psychology, including the volume thinking and speech*. Springer, Berlin.
Wertsch, J.V. (1985). *Vygotsky and the social formation of mind*. Harvard University Press, Cambridge, MA.
Wertsch, J.V. (1991). *Voices of the mind: a sociocultural approach to mediated action*. Harvard University Press, Cambridge, Mass.
Wertsch, J.V. (1998). *Mind as action*. Oxford University Press, New York.