PROJECT REPORTS

Exploring Monological and Dialogical Phases in Naturally Occurring Explanations

Josephine B. Fisher[1] · Vivien Lohmer[2] · Friederike Kern[2] · Winfried Barthlen[3] · Sebastian Gaus[4] · Katharina J. Rohlfing[1]

Abstract

Recent approaches to Explainable AI (XAI) promise to satisfy diverse user expectations by allowing them to steer the interaction in order to elicit content relevant to them. However, little is known about how and to what extent the explainee takes part actively in the process of explaining. To tackle this empirical gap, we exploratively examined naturally occurring everyday explanations in doctor–patient interactions ($N = 11$). Following the social design of XAI, we view explanations as emerging in interactions: first, we identified the verbal behavior of both the explainer and the explainee in the sequential context, which we could assign to phases that were either monological or dialogical; second, we investigated in particular who was responsible for the initiation of the different phases. Finally, we took a closer look at the global conversational structure of explanations by applying a context-sensitive model of organizational jobs, thus adding a third layer of analysis. Results show that in our small sample of conversational explanations, both monological and dialogical phases varied in their length, timing of occurrence (at the early or later stages of the interaction) and their initiation (by the explainer or the explainee). They alternated several times in the course of the interaction. However, we also found some patterns suggesting that all interactions started with a monological phase initiated by the explainer. Both conversational partners contributed to the core organizational job that constitutes an explanation. We interpret the results as an indication for naturally occurring everyday explanations in doctor–patient interactions to be co-constructed on three levels of linguistic description: (1) by switching back and forth between monological to dialogical phases that (2) can be initiated by both partners and (3) by the mutual accomplishment and thus responsibility for an explanation's core job that is crucial for the success of the explanation. Because of the explorative nature of our study, these results need to be investigated (a) with a larger sample and (b) in other contexts. However, our results suggest that future designs of artificial explainable systems should design the explanatory dialogue in such a way that it includes monological and dialogical phases that can be initiated not only by the explainer but also by the explainee, as both contribute to the core job of explicating procedural, clausal, or conceptual relations in explanations.

Keywords Explanation · Co-construction · Conversational jobs · Doctor-patient-communication

✉ Josephine B. Fisher
jbfisher@mail.uni-paderborn.de

Vivien Lohmer
vivien.ebben@uni-bielefeld.de

Friederike Kern
friederike.kern@uni-bielefeld.de

Winfried Barthlen
winfried.barthlen@evkb.de

Sebastian Gaus
sebastian.gaus@evkb.de

Katharina J. Rohlfing
katharina.rohlfing@uni-paderborn.de

1   Faculty of Arts and Humanities, Paderborn University, Paderborn, Germany

2   Faculty of Linguistics and Literature, Bielefeld University, Bielefeld, Germany

3   Clinic of Pediatric Surgery, Evangelisches Klinikum Bethel, Bielefeld, Germany

4   Clinic of Pediatrics, Evangelisches Klinikum Bethel, Bielefeld, Germany

1 Introduction

Algorithms that guide and determine our everyday lives are taking over information processing and becoming more and more complex. They learn and represent data based on hierarchical architectures that are difficult to assess, thereby turning the models into *black boxes*. To make these boxes assessable is the goal of recent explainable artificial intelligence (XAI) approaches. One such recent approach proposed by Sokol and Flach [18] identifies the challenge of satisfying diverse expectations and competing objectives. Derived from the theoretical work of Miller [13], new interactive approaches to this challenge allow the involvement of both partners—the explainer (EX) as the more knowledgeable person and the explainee (EE) as the less knowledgeable addressee of an explanation. However, basic empirical research on how to involve the EE is scarce. In fact, most approaches proceed from the assumption that an explanation is initiated by an EX providing the explanans (the verbal way that an explanation can be expressed), with recent approaches then allowing the EE to ask questions after this "job" [18]. Following this assumption, the current design may involve two phases: a monological phase in which the explanation is delivered, and a dialogical phase in which the addressee is given the possibility to clarify by raising specific questions. This design is justified through research suggesting that when asked spontaneously, EXs provide long-winded and monological statements without involving the EE visibly [2, 9, 11].

However, there are two reasons to question the assumed fixed sequence of a dialogical phase following the monological one. First, linguistic research suggests that conversational partners do not only exchange content in a conversational interaction, but also gradually co-construct the conversational tasks necessary for the accomplishment of a common goal [5, 6]. Second, the sequence found in the above-mentioned research might be a methodic by-product of explanations elicited in semi-experimental settings.

Pursuing the question of how naturally occurring everyday explanations are sequentially structured with regard to monological and dialogical phases, we aimed to study the occurrences of these two types of phases in natural explanations in the specific context of medical consultations. We investigated when and to what extent explanations display the phases, when and which partner initiates them, and to what extent the emerging structure is reflected in what kind of phase. For the occurrence of the phases and based on existing research, we followed the prevalent assumption that the EE will become actively involved later in the interaction, resulting in a first dialogical phase occurring after the first monological phase by the EX. However, we also formulated an alternative assumption against the social framework of

XAI, expecting both types of phases (dialogical vs. monological) alternating several times during the explanation, thus, allowing the EE not only to bring in own interest and knowledge but also to shape the explanandum [15] and, consequently, the explanation itself.

The ability to initiate a monological or dialogical phase requires knowledge about the global structure of an explanatory sequence [14]. It should be noted that everyday explanations do not occur isolated but are usually embedded in a larger conversational context [14]. Consequently, we prefer to speak of explanatory sequences rather than simply explanations. Empirical research in the domain of Conversation Analysis (CA) has pointed out that such sequences can be described as consisting of regularly occurring, genre-specific organizational jobs that are mutually accomplished and thus co-constructed by the conversational partners (that is EX and EE in our case) [10, 14], Kern 2020; [15].

The first two jobs, ((1) *establishing topical relevance* and *(2)*, *constituting an explanandum*,) pave the way for a subsequent conversational explanation, and the last two jobs ((4) *closing,* (5) *transition)* lead the way back to the turn-by-turn talk. The core job in explanatory sequence (3) *explicating procedural, conceptual and/or causal relations*), is sequentially positioned between the second and the fourth job. Following the CA-informed empirical research, it is especially the last two jobs through by which the explanation is implicitly or explicitly ratified [14]. Even when the EX takes the role of principal speaker [19], we aimed to investigate whether jobs are co-constructed in our specific context of medical consultations as an example for naturally occurring interactions, with both partners taking responsibility for them. As previous research has shown, the main effort of the explainers resides in the core job (3) that is about *explicating procedural, conceptual and/or causal relations* [14]. An operationalization of all these jobs enabled us to reveal the interplay between jobs and types of phases to assign the shifting responsibilities for the ongoing explanatory sequence, in terms of who is initiating what, and, finally, to quantify both jobs and types of phases.

2 Methods

Our mixed method pertains to the examination of the three levels of linguistic description on which a co-construction of an explanation can take place (1) the two types of phases, (2) their phase initiation, and (3) the involvement in the conversational jobs by the EE and EX.

2.1 Participants

We collected naturalistic data without any elicitation method in order to obtain naturally occurring explanations. These

data were acquired in the context of patient–physician interactions taking place at the Clinic for Pediatric Surgery, Bethel hospital in Bielefeld, Germany. Overall, eleven consultations were audio- and videotaped. The majority (9/11) of these consultations were held in the children's outpatient department (NoKi); the other two, in the pediatric surgery department. In total, eight doctors, ten patients, and 13 legal guardians took part in the study.[1] Because the consultations concerned different types of surgeries, the length of the explanation differed. This was because certain surgeries took more time to explain than others. The duration of conversations varied between 5 and 26 minutes ($M = 9.5$ minutes, $SD = 6.5$). The data collection took three months. The consultations at the NoKi could only be planned shortly before the appointments, and this was the only time when we could ask participants to participate. This difficult workflow resulted in us spending days at the hospital often without data being acquired. All participants who gave their consent were included in our explorative study. In this type of explanation, the doctors were more knowledgeable and thus the EX, whereas the legal guardians were the less knowledgeable EE. Data collection was terminated because of this workflow placing additional demands on hospital personnel.

2.2 Coding

To explore the engagement of the EE, we coded the content, the different types of phases, the speech of the EE and EX, phase initiations by the EE and EX, and the jobs performed by each participant. The coding of the jobs and phases was done simultaneously by two researchers. After completing this coding process, the two coders switched and recoded two explanatory sequences to check the reliability of their coding schemes. We then calculated an unweighted Cohen's kappa [3, 4]. Explanations corresponded in their temporal length to 20.0 % of the entire data set. The number of turns taken by the EX formed the sum of the sample (*N*). Thereby, both coding schemes were tested for reliability, the coding of the types of phases as well as the coding of the jobs. The coding of the two types of phases—monological and dialogical—resulted in a Cohen's kappa of $k = 0.7$. Cohen's kappa for the coding of the jobs was $k = 0.86$. Both coders thus revealed moderate to near-perfect coding agreement [12]. Following the reliability test, deviations between the two coders were smoothed.

---

[1] The children did not engage in the consultations, except for two out of 12 cases. Mostly due to their young age.

2.2.1 Content-Related Segmentation

Additionally, the explanatory sequences were coded according to the content they provided. The consultations in our dataset included the following three content-related elements: (1) organizational issues, (2) health check of the patient, and (3) information on the surgery. Only the latter element was considered in the analysis, because this constituted the core explanandum. Consequently, the coding started when (3) the surgery was being discussed.

2.2.2 Speech Segmentation

The speech of the participants was segmented into turns and backchannels. A turn can have the length of at least one word, or it can take up to several sentences [16]. Backchannels are defined by Dideriksen et al. (7, p. 262) as "head nods or short utterances consisting of a word (e.g., 'uh-huh,' 'yes,' 'okay'), or short sentences, often repeating the previous turn (e.g., A: 'let's meet Monday at 10,' B: 'Monday at 10')."

2.2.3 Phase Type Segmentation

The different types of phases were segmented into more monological and more dialogical phases. For the analysis, we considered the speech of the participants only, including the backchannel signals. The monological phases could include backchannels but no turns of other speakers besides the current speaker [10, 17]. Therefore, the monological phases started when the principal speakers were the only ones talking, thereby holding the floor, and ended when they addressed the other participants or were being addressed by them. In the dialogical phases, multiple speakers produced turns and also used backchannel signals alternately. Pauses between phases were excluded because it was difficult to assign them to a speaker.

2.2.4 Phase Initiation

In order to describe the explanatory sequences in more detail, we also considered it important to find out who initiated the different types of phases. All doctor phase initiations were coded with EX and all initiations by legal guardians or children with EE.

2.2.5 Job Segmentation

For coding, the content of the turns of EX and EE was considered and processed according to Quasthoff et al. [14].

The job segmentation was based on the model presented above [14] which describes the global conversational

Table 1 Coding scheme for the conversational jobs

Job	Definition	Examples
Establishing topical relevance (1)	Framing the upcoming sequence as a longer mono-logical phase.	EX: What we also need to talk about… (04)
		EX: Planned is quasi for the fracture that we already talked about… (03)
Constituting an explanandum (2)	Specifying the explanandum by localizing the crucial body areas or parts.	EX: As my colleague already said there are two types of circumcision. (11)
		EX: So, he has up here a groove at his penis. (09)
Explication procedural, conceptual, and/or causal relation (3)	Providing information on the processes, concepts, and/or reasons underlying the operation as a whole or for individual steps.	EE: I don't know if it is relevant but he is a premature baby, he had, he came five weeks too early, and that is the risk he had in the past but he outgrew them. (04)
		EX: The procedure is nevertheless the same. So, we will make for one, a cut at the groin. (06)
Closing (4)	Framing the upcoming sequence as a completion/ending of the explanation.	EX: But one cannot exclude it for 100 percent, that is basically it. (04)
		EX: Further questions? EE: No. (04)
Transition (5)	Transfer back to a turn-by-turn structure in a small talk character or by signing the document.	EE: Thank you very much. (04)
		EX: Then we will do it that way, yes? (04)
		EX: Good I have quickly written everything down. (05)
		EX: So, then I get your signature. (05)

The consultations were conducted in German in 10/11 cases (02–11). Sentence structure and semantics were adapted to the English language and thus correspond broadly to the German transcripts

structure of explanatory sequences in terms of organizational jobs co-constructively performed by both EE and EX.

The analysis and coding followed the methods of conversational analysis (see Table 1). Only the verbal level was considered. Other modalities such as gaze, gesture, or positioning of the participants towards each other were excluded from this analysis.

3 Results

We will first present results on the occurrences of the two types of phases, i.e., monological vs. dialogical. Secondly, the initiations of the different phases will be examined. We will then report the findings on the organizational jobs that participants co-constructively performed to accomplish the explanatory sequence.

Our data demonstrated that the explanatory sequences consisted predominantly of the first type of phase, the monological phases in which the EX acted mostly as the principal speaker, and in dialogical phases to a considerably lesser extent in which the EE was more actively involved (cf. Fig. 1).

In more detail, the monological phases made up 73.5 % with 4.9 seconds in total (minimum 57.7 % and maximum 90.8 %, $SD = 11.0$), whereas the dialogical phases made up 26.5 % with a total of 1,396 seconds (minimum 9.2 % and maximum 42.3 %, $SD = 11.1$).

With reference to previous research suggesting that the first monological phase is the longest, we took a closer look at the beginnings of the explanatory sequences. We found that the length of the first monological phases varied immensely in the data set (see Fig. 2): In consultations 2, 7 and 10 (3/11), they constituted the longest phase in the entire sequence. Whereas in the other cases (8/11), the subsequent dialogical phase was initiated earlier which resulted in a shorter initial monological phase.

In the next step of our analysis, we considered whether the EX or EE initiated what kind of a phase in order to investigate the distribution of the responsibilities. The Wilcoxon Signed Rank test on the occurrences presented in Table 2 revealed that the EX tended to initiate the monological phases more frequently than the dialogical phases ($Z = 2.8$, $p < 0.001$, $r = 0.8$). In contrast, it was the EE who tended to initiate the dialogical phases more frequently than the monological phases ($Z = 1.8$, $p < 0.035$, $r = 0.5$). Further, as can be seen from Table 2, there was a clear division of responsibilities, with the EX predominantly initiating the monological, and EE mostly initiating the dialogical phases. However, even though the division of responsibilities seems clear, in 3 consultations, an EE initiated a monological phase whereas in 10 consultations, EX initiated a dialogical phase. This variability supports the fact that there is no rule of responsibility in our sample.

We now turn to the analysis of the organizational jobs in terms of their possible relation with the two types of phases.

Fig. 1 Proportion of monological and dialogical phases in the consultations
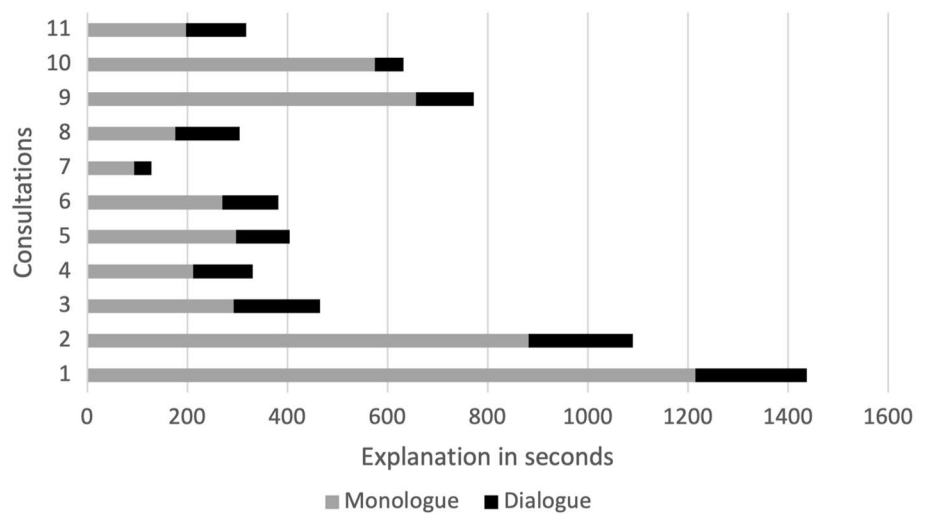
We start by taking a closer look at the distribution of the two types of phases in the different jobs (see Fig. 3). Building on the results, we will then present a case analysis for the core job (3) to highlight the co-constructive character of jobs in the dialogical phases.

Our analysis showed that the jobs coincide with both monological and dialogical phases (cf. Fig. 4).

We can thus assume a connection between our two analytical categories of organizational jobs and phase types. In more detail, the core job was executed in 83.0% of the monological phases (minimum 63.0% and maximum 94.0%, $SD = 9.0$). In addition, in ten out of eleven cases (90.9%), the core job occurred in the first monological phase.

In the following, we shall focus more specifically on the main job *explicating procedural, conceptual, and/or causal relations* (3) to point out the co-constructive characteristics in the monological as well as in the dialogical phases. We present the analysis of the core job (3) in a monological phase being further taken up in the subsequent dialogical phase following the method of conversation analysis (CA) [1].

The example is from consultation 04 and is in the first monological phase. In this case, the first monological phase includes the first two jobs (*establishing topical relevance* (1), *constituting an explanandum* (2)) as well as the core job *explicating procedural, conceptual, and/or causal relations* (3). The EX is the principal speaker. After the second job *constituting an explanandum,* the EX initiates the core job (*explicating procedural, conceptual, and/or causal relations* (3)) by: "äh what we do äh. This is followed by the EX's explication about the planned surgery. First, the surgery is named (EX: "what we do äh is basically uh a closed reposition"). Second, the procedure of the surgery is further explicated (EX: "that mean when he sleeps, you pull it once"; EE: "mhm" EX: "sets it straight again, that you put both äh break ends in front of each other again and then it is fixed with two

wires, maybe [three] (EE: "[mhm]") depending"). The EE regularly uses backchannel signals to display participation in the explanatory sequence and that the EX can continue with his explanation (Schegloff 1981; [10]). Thereby, the EE does not attempt to take the floor. This example shows that even if the EX as principal speaker takes the main effort during the monological phase in the core job, the EE is verbally involved as well, albeit to a reduced extent–the EX remains the principal speaker. However, as the backchanneling signals are important for the continuation of EX's explication, we see this as an indication of co-construction.

In what follows now, the EE initiates a dialogical phase by asking a question, and thus addresses the EX directly (EE: "may I ask you a question" [in between]). The EX immediately agrees to the EE's request (EX: "[yes of course]". Again, both participants are verbally involved, but the EX is no longer the principal speaker; instead, both participants share responsibility for the continuation of the conversation.

After obtaining the right to speak, the EE asks for confirmation on some information they had apparently received from the previous surgery (EE "your colleague has already tried that. Right? That uh was then but mechanically not manageable yet"). The EX supports this information with an acknowledgement token (EX "exactly"). Then the EX continues by explicating why the previous surgery by a colleague had not been successful (EX: "so it is in the end, it has become just a little bit better, but it will not get any better or the danger is simply very very great that this can shift even further"). In contrast to the first example, both EE and EX contribute to the accomplishment of the job in the dialogical phase.

In sum, from our analyses, we derive the following results for the interplay of the different phases and the involvement of the EE and EX in the specific context of medical consultations:
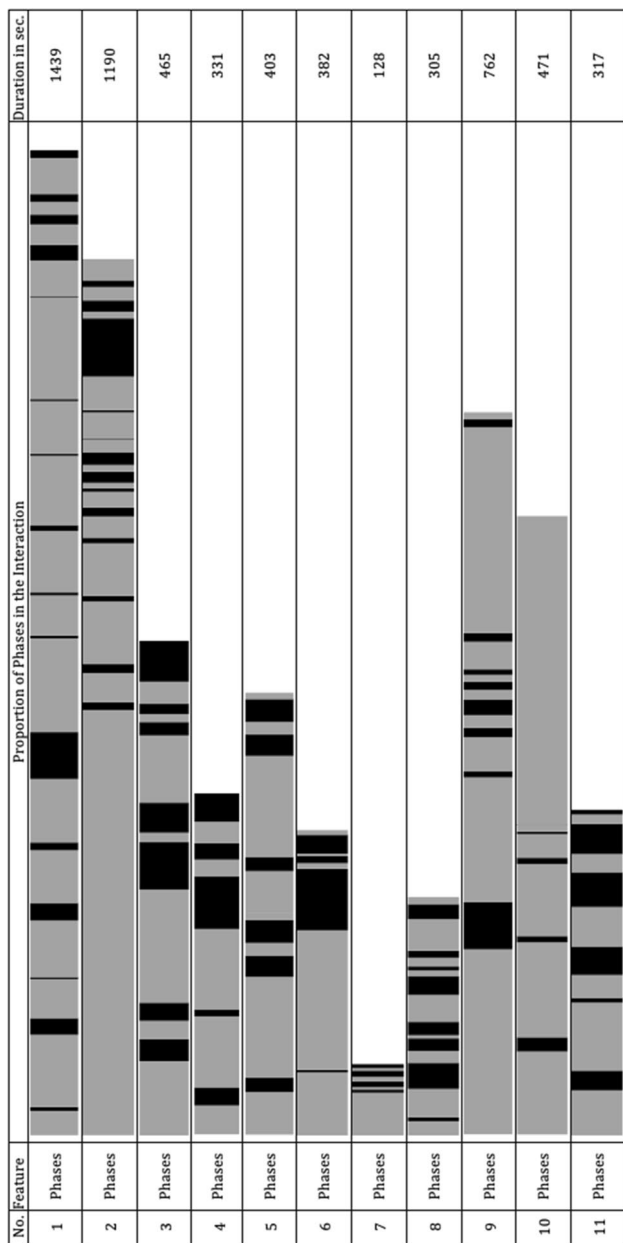
Fig. 2 Monological and dialogical phases in the explanatory sequences ($N=11$) are indicated in gray (monological) and black (dialogical). Phases are displayed in relation to the duration time (t) of the entire sequence

(a) Naturally occurring explanations consist of two types of phases: monological and dialogical, indicating that both types of phases are necessary in explanatory sequences. They are monological to a great extent; hence, the dialogical phases are rather short. Further, an explanatory sequence starts with a monological phase, although the first monological phase is not always the

Table 2 Total occurrences of monological and dialogical phase initiations by explainers and explainees

Consultation	Monological phases initiated by		Dialogical phases initiated by	
	Explainer	Explainee	Explainer	Explainee
1	8	8	10	6
2	14	0	1	12
3	6	1	2	5
4	5	0	1	4
5	4	0	1	2
6	5	0	2	2
7	4	0	1	3
8	9	0	4	4
9	8	0	2	5
10	6	1	2	3
11	6	0	0	6

longest in comparison to the other monological phases produced in the course of explanatory sequences.

(b) When looking at the initiation of the two types of phases in naturally occurring explanations, we found that the EX tended to initiate the monological phases more frequently. Vice versa the EE tended to initiate the dialogical phases more frequently suggesting a typical pattern known from scientific explanations [2].

(c) Adding another analytical layer, when averaged across explanatory sequences, the core job *explicating procedural, conceptual, and/or causal relations* (3) related to monological as well as dialogical phases, thus suggesting that it is omnipresent in both types of phases. Importantly, both EX and EE contribute to the accomplishment of the core job.

4 Discussion

In this study, we investigated naturally occurring explanations with the aim of exploring the involvement of the explainee (EE) as a first step toward an empirical basis to advance the social design of XAI. For this purpose, we sharpened the structure proposed by Quasthoff et al. [14] and further distinguished between monological and dialogical phases within an explanatory sequence. Building on the proposed structure (ibid), we applied a mixed method and were able to discover some patterns. In contrast to current XAI research where the two types of phases follow upon each other, we found monological phases to alternate with dialogical phases throughout the explanatory sequence, unless it is the last phase within a consultation (see Fig. 2). This result supports a study by Kobayashi [11] who found

Fig. 3 Distribution of jobs within the phases and across the 11 consultations: for example, in consultation "No. 04" starts with "M1" referring to the first monological phase and indicating that jobs 2 and 3 was performed. This is followed by "D1" indicating that job 3 was performed in the first dialogical phase

Legend: 1 Establishing topical relevance 2 Constituting an explanandum 3 Explication procedural, conceptual and/or causal relation 4 Closing 5 Transition

No. 01

Phase	Job
M1	1,2
D1	2
M2	2
D2	2,3
M3	3
D3	3
M4	3
D4	3
M5	3
D5	3
M6	3
D6	3
M7	3,4
D7	4
M8	4
D8	4
M9	4
D9	4
M10	4
D10	4
M11	4
D11	4
M12	4,5,2 3,4,5
D12	5
M13	2,3
D13	4,2,3
M14	3
D14	3
M15	3
D15	2,3
M16	3,4
D16	5

No. 02

Phase	Job
M1	1,2, 3,4
D1	3
M2	3
D2	3
M3	3
D3	3
M4	3
D4	3
M5	3
D5	4
M6	4
D6	3
M7	3
D7	3
M8	4,2, 3
D8	4
M9	4
D9	4
M10	4
D10	4
M11	3,4
D11	4,3
M12	4,5
D12	5
M13	5
D13	5
M14	5

No. 03

Phase	Job
M1	1,2, 3
D1	3
M2	3
D2	3
M3	3
D3	3,4
M4	3
D4	3
M5	5
D5	5
M6	4
D6	4
M7	4
D7	4,5

No. 04

Phase	Job
M1	2,3
D1	3
M2	3,4
D2	4,5
M3	1,2,3,4
D3	4,3
M4	3
D4	3
M5	4
D5	4

No. 05

Phase	Job
M1	1,2, 3
D1	4,3
M2	3
D2	3,4
M3	3
D3	4,3
M4	4,5

No. 06

Phase	Job
M1	1,2, 3
D1	3
M2	3
D2	3
M3	3
D3	4
M4	4
D4	4
M5	4

No. 07

Phase	Job
M1	1,2, 3
D1	4
M2	4
D2	3
M3	3
D3	3
M4	4
D4	5

No. 08

Phase	Job
M1	2,3
D1	3
M2	3
D2	3
M3	3
D3	3
M4	3
D4	3
M5	3
D5	3
M6	3
M7	3
D7	4,3
M8	3
D8	4,3
M9	3,4

No. 09

Phase	Job
M1	1,2,3
D1	3
M2	3
D2	3
M3	3
D3	3
M4	3
D4	3
M5	3
D5	3
M6	4,5
M7	3
D7	3
M8	2,3,4
D8	4
M9	4

No. 10

Phase	Job
M1	2,3
D1	3
M2	3,4
D2	3
M3	3
D3	3
M4	3
D4	3
M5	3,4

No. 11

Phase	Job
M1	3
D1	1,2, 3
M2	3,4
D2	3
M3	3
D3	3
M4	3
D4	2
M5	2
D5	2
M6	2

Fig. 4 Different jobs of an explanatory sequence distributed among monological vs. dialogical phases
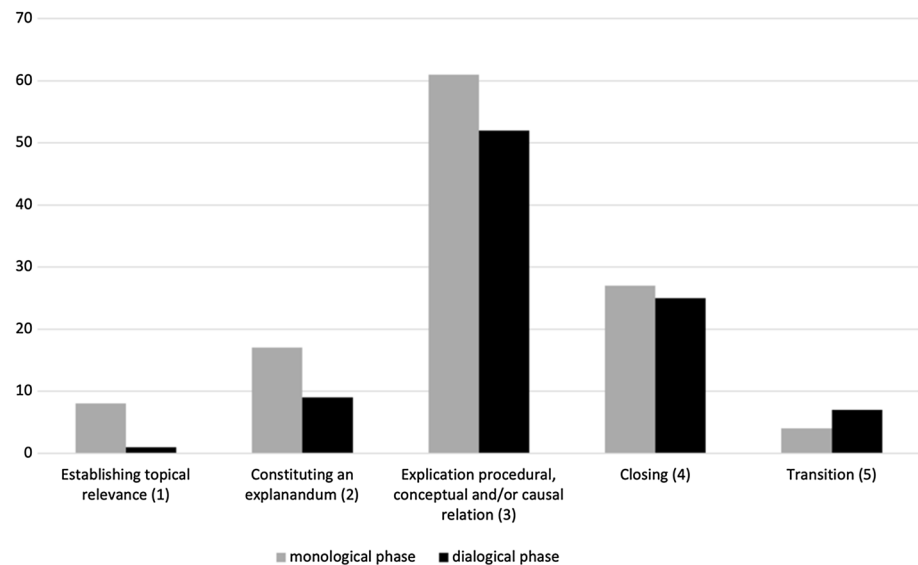
that learning from an explanation is best when partners are involved in both initial (similar to the monological phase in our study) as well as interactive phases (similar to the dialogical phase). Our results are in line with his findings by indicating that in naturally occurring explanations, both kinds of phases can be initiated by both partners. Further, the suggestion that the first monological phase is the longest [2, 11] did not apply generally to the consultations investigated here. Instead, more variation was visible indicating the co-constructive view.

With respect to the initiation of the phases and in line with previous research on scientific explanations [8], we found that most of the monological phases were initiated and performed predominantly by the explainer (EX). Vice versa, the explainee (EE) was found to initiate the majority of the dialogical phases in our sample.

Adding an analytical layer by identifying the conversational jobs, we further found that the core job of *explicating procedural, conceptual, and/or causal relations* (3) was present in both types of phases. That this job emerged during ongoing interaction suggests that the explanatory information might be reformulated, negotiated between the partners, or spread across the partners' contributions, and thus needs to be synthesized or collected.

Taken the three levels of linguistic description together, our exploratory study of naturally occurring explanations yielded frequently alternating phases, limited flexibility in the initiation of the phases, and the (re)occurrence of the core job in both types of phases. We interpret our findings as three possible aspects of co-construction.

These aspects can inform the XAI development about some required features but call for further empirical investigation: Co-constructive XAI could consider to make it possible for digital or embodied agents to switch between the monological and dialogical phases in an explanatory sequence. Furthermore, it should be possible for both the EX and the EE to initiate either phase. Finally, the occurrence of the core job in both phases indicates that the essence of an explanation emerges during the interaction. Future empirical investigation has to test that–based on our result (a) presented above–the monological and dialogical phases reoccur in explanatory sequences; further, with regard to the result (b), we can hypothesize that both partners are involved in the initiation of both types of phases but some responsibilities are assigned to EX (initiating the monological phases) and EE (initiating the dialogical phases); finally, with regard to the result presented in (c), the hypothesis that the core job is omnipresent in both types of phases need to be confirmed in a larger sample. How exactly both partners contribute to it is also a topic of further research analyzing the content transported in the main job.

In addition to the exploratory nature of our study, there are some more limitations to our findings. Even though our statistical analysis yielded some patterns, we must critically remark that our data stem from one hospital. It may well be that in our setting, an established conversational procedure was applied in which explainees are not only informed about the surgery but also need to give their consent at the end of the consultation. To what extent the results obtained here can be generalized to other explanatory settings is clearly a topic for further research that better isolates the contextual factors from properties of the explanatory sequences. Another limitation is that whereas we focused on the different kinds of phases, how they occur and are initiated naturally, we did not further analyze the content of the core job or the properties of the explanatory sequences. This indicates the need for further, more qualitative research. It is of vital interest to gain a deeper insight into the involvement of the EE in

an explanation. Therefore, following Kobayashi [11], one could elicit the two phases systematically and then run a more detailed analysis, focusing on the specific verbal moves that the EE and EX make use of in the respective phases to determine what kind of moves occur at what phases. The selection of specific speaker moves could serve as a conversational technique to guide an explanation and to adapt to the EE. For XAI, adaptation to the EE is of high interest to ensure the relevance of the explanation.

5 Conclusion

Whereas first approaches to interactive explainable systems are currently being developed [18], little is known about the empirical basis for how to involve the explainee in an explanatory sequence. To design XAI systems more adaptable to the demands of human agents, studying the co-construction of naturally occurring explanations may provide helpful insight. Our investigation of naturally occurring everyday explanations indicates that co-construction implies an involvement of explainees at any time during the unfolding explanation. Even though our investigation is of explorative nature, the development of XAI systems might benefit from our results already. They could take the co-constructive nature of naturally occurring everyday explanations more into account: A first step could involve the implementation of alternating monological and dialogical phases during the explanatory sequence. We suggest that XAI might go even further by not only allowing humans to ask questions in the dialogical phases but also by providing them with the agency to intervene at any phase, or at least to initiate the dialogical phases.

A second step would build on existing research into the global structure of explanations described in terms of five subsequent and mutually accomplished organizational jobs. Using this structure as a modelling base for explanations, XAI could offer more fine-tuned adaptation opportunities for human agents' demands. Accordingly, they could intervene and initiate dialogical phases when they consider them appropriate or necessary. In addition, monitoring the core job and a form of synthesis of both partners' contributions could allow for a meaningful and co-constructed extension or reorganization of the explanandum. This way, an XAI could achieve a progressive adaptation of the explanation to the explainee, increasing understanding of the explanandum.

To conclude, based on first exploratory empirical results, modeling could thus take both the alternation of monological and dialogical phases during the explanatory sequence and their relation to its global structure into consideration. As a result, an XAI system could be designed to be more adaptive as it would enable the human to take part in the explanation more actively. Further research is needed to test our assumptions and to provide XAI with detailed information on the composition of jobs regarding the systematics of alternating monological and dialogical phases.

Appendix

See Fig. 3

Author Contributions We specify the following contribution: conceptualization, all authors; methodology, JF and KR; data acquisition, JF; data analysis, JF, VL, FK; original draft preparation, JF, VL; supervision, FK, KR; funding acquisition, FK, KR. All authors have read and agreed to the published version of the manuscript. WB and SG contributed to the patient and data acquisition and discussion.

Data availability The original data presented in the study are summarized in the article (Table 2), further inquiries can be directed to the corresponding author. The data are not publicly available due to informed consent form not including this possibility.

Declarations

Ethical Approval The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Paderborn University on January 11, 2022.

Informed Consent Informed consent was obtained from all subjects involved in the study.

References

1. Bergmann J (2019) Ethnomethodologische Konversationsanalyse. In: Hoffmann L (Ed), Sprachwissenschaft: Ein Reader (4., aktualisierte und erweiterte Auflage). De Gruyter
2. Chi MTH (1996) Constructing self-explanations and scaffolded explanations in tutoring. Appl Cogn Psychol 10(7):33–49. https://doi.org/10.1002/(SICI)1099-0720(199611)10:7%3c33::AID-ACP436%3e3.0.CO;2-E
3. Cohen J (1960) A coefficient of agreement for nominal scales. Educ Psychol Measur 20(1):37–46. https://doi.org/10.1177/001316446002000104
4. Cohen J (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. 8

5. Dausendschön-Gay U, Gülich E, Krafft U (Eds) (2015a) Ko-Konstruktionen in der Interaktion: Die gemeinsame Arbeit an Äußerungen und anderen sozialen Ereignissen. Transcript

6. Dausendschön-Gay U, Gülich E, Krafft U (eds) (2015b) Zu einem Konzept von Ko-Konstruktion. In: Ko-Konstruktionen in der Interaktion. Die gemeinsame Arbeit an Äußerungen und anderen sozialen Ereignissen. Transcript, pp. 21–36

7. Dideriksen C, Fusaroli R, Tylén K, Dingemanse M, Christiansen MH (2019) Contextualizing conversational strategies: backchannel, repair and linguistic alignment in spontaneous and task-oriented conversations [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/fd8y9

8. Garfinkel A (1982) Forms of explanation: rethinking the questions in social theory. Br J Philos Sci 33(4):438–441

9. Graesser AC, Person NK (1994) Question asking during tutoring. Am Educ Res J 31(1):104–137. https://doi.org/10.3102/00028312031001104

10. Kern F (2007) Prosody as a resource in children's game explanations: some aspects of turn construction and recipiency. J Pragmat 39(1):111–133. https://doi.org/10.1016/j.pragma.2005.01.017

11. Kobayashi K (2021) Learning by teaching face-to-face: the contributions of preparing-to-teach, initial-explanation, and interaction phases. Eur J Psychol Educ 37(2):551–566. https://doi.org/10.1007/s10212-021-00547-z

12. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33(1):159. https://doi.org/10.2307/2529310

13. Miller T (2019) Explanation in Artificial Intelligence: Insights from the Social Sciences (arXiv:1706.07269). arXiv. http://arxiv.org/abs/1706.07269

14. Quasthoff U, Heller V, Morek M (2017) On the sequential organization and genre-orientation of discourse units in interaction: an analytic framework. Discourse Stud 19(1):84–110. https://doi.org/10.1177/1461445616683596

15. Rohlfing KJ, Cimiano P, Scharlau I, Matzner T, Buhl HM, Buschmeier H, Esposito E, Grimminger A, Hammer B, Hab-Umbach R, Horwath I, Hullermeier E, Kern F, Kopp S, Thommes K, Ngonga Ngomo A-C, Schulte C, Wachsmuth H, Wagner P, Wrede B (2021) Explanation as a social practice: toward a conceptual framework for the social design of AI systems. IEEE Trans Cognitive Dev Syst 13(3):717–728. https://doi.org/10.1109/TCDS.2020.3044366

16. Sacks H, Schegloff E, Jefferson G (1974) A simplest systematics for the organization of turn-taking for conversation. Language 50(4):40

17. Schegloff EA (1982) Discourse as an interactional achievement: some uses of 'uh huh' and other things that come between sentences. In: Analyzing Discourse: Text Talk, vol 71, pp 71–93

18. Sokol K, Flach P (2020) One explanation does not fit all: the promise of interactive explanations for machine learning transparency. KI—Künstliche Intelligenz 34(2):235–250. https://doi.org/10.1007/s13218-020-00637-y

19. Wald B (1978) Zur Einheitlichkeit und Einleitung von Diskurseinheiten. In: Quasthoff U (ed) Sprachstruktur - Sozialstruktur: zur linguist. Theorienbildung