

Studying Challenges in Medical Conversation with Structured Annotation

Nan Wang

Hunan University

nwang.christine@gmail.com

Yan Song

Sinovation Ventures

clksong@gmail.com

Fei Xia

University of Washington

fxia@uw.edu

Abstract

Medical conversation is a central part of medical care. Yet, the current state and quality of medical conversation is far from perfect. Therefore, a substantial amount of research has been done to obtain a better understanding of medical conversation and to address its practical challenges and dilemmas. In line with this stream of research, we have developed a multi-layer structure annotation scheme to analyze medical conversation, and are using the scheme to construct a corpus of naturally occurring medical conversation in Chinese pediatric primary care setting. Some of the preliminary findings are reported regarding 1) *how a medical conversation starts*, 2) *where communication problems tend to occur*, and 3) *how physicians close a conversation*. Challenges and opportunities for research on medical conversation with NLP techniques will be discussed.

1 Introduction

Medical conversation is at the core of medical care. Through conversation, doctors collect the information needed to form a diagnosis and provide a treatment recommendation for the patient's condition. Effective communication is essential for achieving optimal medical outcomes. Yet breakdowns in doctor-patient conversation are common in medical practices. For example, the largest proportion of hospital and community health services complaints in the UK were about communications with medical professional in 2017-2018 (NHS, 2018). Thus, a better understanding of medical conversation (e.g., how it is conducted; what practical problems and dilemmas doctors and patients face) could improve not only the quality of care, but also the efficiency of the healthcare system.

In this paper, we first review the major issues that medical conversation research has investigated; we

then introduce the data and methods that we use to analyze medical conversation; lastly, we present some preliminary findings based on our analysis of the corpus and conclude with a discussion on implications of the study and our future work.

2 Research on Medical Conversation

Research on medical conversation has a long tradition. Earlier studies that use naturally occurring medical conversation data can be traced back to 1970s in the United Kingdom and the United States. Having audio-recorded and analyzed 2500 recordings of medical conversation, British researchers Byrne and Long (1976) were regarded as the pioneers in medical conversation research. At about the same time, American physicians Korsch and Negrete (1981) conducted one of the most influential studies on medical communication, based on 800 audio-recordings of conversation collected in the Los Angeles Children's Hospital. These studies showed that medical communication practices significantly affect patient health outcomes.

More recently, a substantial body of research using conversation analysis (Drew et al., 2001; Heritage and Maynard, 2006) emerged and investigated a wide range of topics in medical conversation. These topics generally fall into three categories.

How is medical conversation conducted? This stream of research examines the *process* and *constituent activities* of medical conversation. In other words, how do physician and patient coordinate in this social encounter and how is medical conversation organized?

Unlike many other types of conversation, medical conversation are treated as having a discernable overall structure. For example, acute primary care encounters in the UK were found ordinarily beginning with an opening sequence, progressing through problem presentation, history taking,

physical examination, diagnosis, and treatment recommendations, and then onto a closing sequence (Byrne and Long, 1976).

While this overall structure of conversation can be considered as socialized through physicians' training in medical school and patients' repeated exposure since childhood, it is also considered a product of coordination and negotiation between physicians and patients in local interaction context. Therefore, a lot of research in this stream has been done to investigate questions such as how one activity transits toward another in medical conversation, and what constitutes a complete sequence within some particular activity phase.

For example, studies find that there are varying expectations for what constitutes a complete diagnosis sequence and a treatment recommendation sequence. While patients' no response or weak response (e.g., *mm hmm*) is treated as sufficient for a diagnosis sequence to complete (Heritage and Sefi, 1992), patients' explicit acceptance of treatment recommendation is oriented as necessary for completing a treatment recommendation sequence (Stivers, 2005).

What are the practical problems and dilemmas in medical conversation? The second stream of research on medical conversation concerns with more concrete problems in medical conversation. For instance, how physicians' and patients' *actions* are designed and *sequences* are organized to deal with various kinds of practical challenges and dilemmas in medical conversation.

For example, physicians use various forms of questions to solicit patients' problem presentations, and these different action types can afford different opportunities for patients' contribution to medical conversation. Specifically, the length of patients' problem presentation is significantly longer if physicians use open-ended solicitation questions, as compared with close-ended questions (Heritage and Robinson, 2006).

In addition, when delivering a diagnosis to terminal patients, it is found that physicians can deploy a rather complicated form of sequence (the so-called *News Delivery Sequence (NDS)*), in which patients' perspectives of their health condition are incorporated, in order to prepare them for the bad news (Maynard, 1997).

How does medical conversation affect medical outcomes? The third stream of medical conver-

sation research examines the associations between doctor-patient interaction and outcomes in medical conversation, such as patient satisfaction or medication adherence.

For instance, research shows that if a candidate diagnosis (e.g., *I don't know whether she's got a sinus, but she has a lot of drainage in her nose.*) appears in patients' problem presentation, physicians are more likely to perceive the patient as expecting antibiotics (Stivers et al., 2003), and are more likely to prescribe antibiotics inappropriately (Mangione-Smith et al., 1999).

Similarly, toward the end of medical conversation, it is found that physicians' action design has a significant impact on the effect of medical communication. For instance, by replacing the word *any* in '*Do you have any other concerns you want to address today?*' with the word *some*, it reduces the likelihoods of patients' unaddressed concerns by up to 50% (Heritage and Robinson, 2006).

In sum, a substantial amount of effort has been made to obtain a better understanding of medical conversation. As a result, many important issues have been discovered and effective solutions have been provided to improve the practice of medical communication and the quality of care.

However, the existing resources and tools for medical conversation research suffer from two major problems, which put significant obstacles to the advancement of the field. These problems are: 1) collecting and analyzing natural conversation data in the medical setting requires a tremendous amount of resources (e.g., labor, time); 2) few standard coding frameworks exist, which allow for systematic analysis of medical conversation that takes into account the interactivity of utterances. Although coding schemes such as Roter Interaction Analysis System (RIAS) (Roter and Larson, 2002) attempt to implement an exhaustive classification of the events in medical conversation, these schemes tend to treat utterances in conversation as isolated units (Heritage and Maynard, 2006).

Motivated by these considerations, we construct a corpus that consists of 1,000 medical conversations and develop a coding scheme that captures the deep structure of conversation. Then we conduct a systematic microanalysis of the medical conversation, which takes into account of both the content and the context of utterances. While not being a focus of this paper, we will conclude the paper with

a few potential use cases of our proposed work.

3 Data and Corpus

In this section, we introduce the the corpus that we construct for medical conversation research.

3.1 Video-recorded Data

A total of 1,000 medical conversation were video-recorded in Chinese pediatric primary care settings. Participants involve 14 physicians and 1,000 patients with their caregivers in 9 hospitals in northern, central and eastern China.

For each conversation, a complete course of medical consultation is included, starting from the patients getting seated, progressing through the discussion of patients' health conditions, and toward the patients leaving the office.

Topics in the conversation are mostly concerning children's acute respiratory tract infections (ARTIs) problems, which involve health complaints such as fever, cough, etc.

The conversation is primarily between physicians and caregivers, similar to pediatric primary care conversation in other countries. It should be noted that although most of the conversation is dyadic, multi-party conversation is also common in our corpus, as more than one caregiver can be present and contribute to conversation.

3.2 Transcribed Data

The video-recorded medical conversation data are transcribed manually by trained research assistants. Adopting the Conversation Analysis transcribing conventions (Jefferson, 2004), each conversation is segmented into turns at turn-taking positions. Besides capturing the verbatim of each turn, the transcription also captures a series of para-linguistic features (e.g., dysfluencies, intonations, overlaps of turns, noticeable silence in and between turns, non-verbal actions such as nodding, etc.), which are essential aspects of natural spoken language.

In addition, the transcribed text is automatically segmented into words using an in-house CRF word segmenter trained on the Chinese Treebank (Xia et al., 2000), so as to provide the necessary basis for conducting related NLP tasks.

3.3 Ethical Considerations

All research procedures were reviewed and approved by UCLA IRB and UW IRB. All identifying information (e.g., person, institution, location names) has been removed from the corpus.

3.4 Our Analysis

In this paper, we describe a series of findings based on the corpus regarding the following aspects.

Overall organization and opening a medical conversation Similar to ordinary conversation, medical conversation is a social encounter where physicians and patients build rapport and social relationship (Schegloff, 1968). Thus, quite often at the beginning of medical conversation, physicians and patients engage in social exchange activities such as greetings and identifying. We refer to this kind of exchange as *opening phase* in medical conversation. Ex 1 illustrates an example of *opening phase* in our corpus.¹

Ex 1: Opening in conversation

1 **DOC:** Hi. @NAME@? How are you?

2 **DAD:** Yes, that's us. How are you?

3 **DOC:** What's going on today?

However, medical conversation is also where patients and physicians deal with patients' health concerns. Thus, in many cases, conversation starts with physicians and patients talking about the patients' health problems, without going through *opening phase*. We refer to this kind of activity as *problem presentation phase*.

Although medical conversation can be opened with either the *opening phase* or the *problem presentation phase*, there seems to be a distributional difference in different cultural and medical settings. We will discuss this in more detail in Section 5.

Sequence expansion and making treatment decisions A second aspect of our analysis focuses on *sequences* within some particular phases in medical conversation.

For example, within *treatment recommendation phase*, we examined how treatment recommendations are delivered and received. Specifically, it is found that treatment decisions (e.g., antibiotic prescriptions) can be negotiated between physicians and patients by patients withholding acceptance of physicians' recommendations. Thus, the minimal form of 'recommendation-acceptance' sequence can be expanded quite extensively, in order to secure the patients' explicit acceptance of the physicians' treatment recommendation. Ex 2 and 3 illustrate examples of a non-expanded form and

¹To save space, we omit the Chinese line and show the translation only

an expanded form of treatment recommendation sequence, respectively.

Ex 2: Non-expanded treatment sequence

1 **DOC:** I'll probably put her on some antibiotics.

2 **MOM:** Okay.

Ex 3: Expanded treatment sequence

1 **DOC:** I'll give her some decongestant.

2 **MOM:** Decongestant?

3 **DOC:** Yeah. It's a viral infection. Antibiotics won't kill.

4 **MOM:** Okay, let's use some decongestant.

Although arguably the most ideal form of treatment recommendation sequence consists of two turns (like in Ex 2), patients' acceptance may be delayed and will then be pursued by physicians (like in Ex 3).

Action design and initiating conversation closures Similar to ordinary conversation, the physician and the patient also need to coordinate to close a medical conversation, rather than simply falling silent (Sacks and Schegloff, 1973).

In medical conversation, past research has shown that, upon reaching a point when the treatment decision is made, physicians produce various forms of actions to initiate the closure of the medical visit (West, 2006). Although the patients can always resist such attempts to close and the conversation may go back-and-forth to other phases, the conversation is considered as entering the *closing phase*, when these closure initiation actions are produced.

After physicians secure the warrant from patients to terminate the conversation, the two parties can then proceed to the terminal exchange of the conversation. Ex 4 illustrates an example of the closure initiation action in our corpus.

Ex 4: Closure (a) Making future arrangement

1 **DOC:** Okay. Follow up in two days, ok?

2 **DAD:** Okay. Thank you.

3 **DOC:** You're welcome.

In Ex 4, the physician initiates the closure of the medical conversation by making a future arrangement for the patient's follow-up visit at line 1. The patient's father accepts the proposal and the two parties immediately proceed to the terminal exchange (*thank you-you're welcome*) at lines 2-3.

Besides the closure initiation action (*a*) *making*

future arrangement shown in Ex 4, there are two other forms of action recurrently observed in the medical conversation: (*b*) *summarizing treatment plans* and (*c*) *announcing closures*. Ex 5 and 6 illustrate two examples.

Ex 5: Closure (b) Summarizing treatment plans

1 **DOC:** Just use these three medications, ok?

2 **MOM:** Okay

3 **DOC:** Alright.

4 **MOM:** Ok. Goodbye, doctor.

Ex 6: Closure (c) Announcing closures

1 **DOC:** Okay. That's it.

2 **MOM:** Thank you, Doctor.

3 **DOC:** You're welcome.

The above examples show that, when implementing particular actions in conversation, there can be different turn designs so as to accommodate the particular contingencies arising from the interaction context. These choices of turn design may afford different opportunity for the recipient's participation, and thereby have different impact on the subsequent development of the conversation. A close examination of this phenomenon thus provides a window to uncover the practical constraints that the patients and the physicians face. Actionable solutions can then be developed to deal with these constraints.

4 COSTA Scheme

To enable a systematic analysis of medical conversation in the dimensions that we described above, we developed an annotation scheme that marks up the structure of medical conversation at multiple levels. In addition, application-dependent labels can be created and added on top of the structural annotation, tailored to particular researchers' interests.

Below we briefly introduce how medical conversations are analyzed using the CONversational Structures and Actions (COSTA) scheme, in terms of 1) conversational structures, and 2) application-dependent labels for conversational actions.

4.1 Annotating conversation structure

Figure 1 illustrates how the hierarchical structure of medical conversation is annotated according to the COSTA scheme. Detail of the COSTA scheme can be found in (Wang et al., 2018).

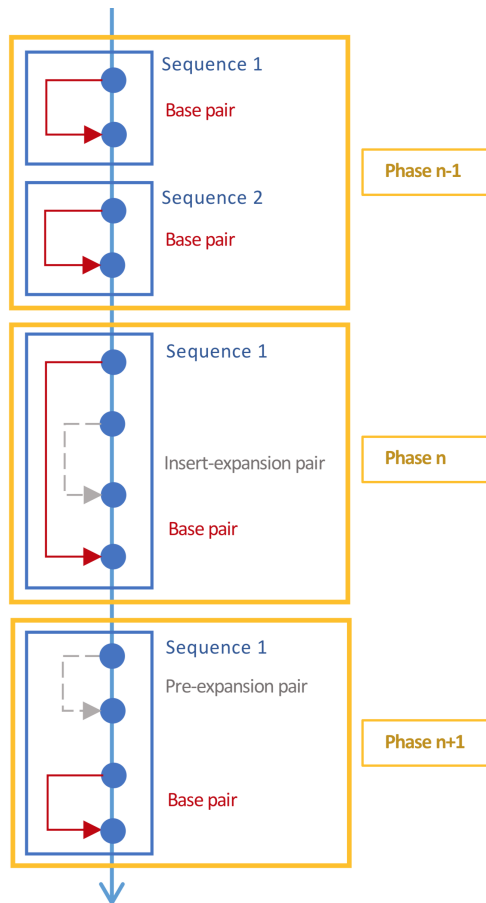


Figure 1: A hierarchical organization of medical conversation. The blue dots represent turns lined up in the temporal order of a conversation. The yellow boxes represent *phases* in the medical conversation (e.g., opening, history-taking, treatment recommendation, etc.), which are consisted of one or more sequences. The blue boxes represent *sequences*, which can be minimally consists of one pair, or multiple pairs with one base pair and its expansion pairs. Within a sequence, the red arrows link the two turns of a *base pair*; whereas the gray arrows suggest that the two connected turns belong to an expansion pair, which is dependent on the base pair.

Overall organization At the highest level, medical conversation is segmented and marked up with a series of component *phases*. Based on related findings on the overall organization of medical conversation from past research (Robinson, 2003; Byrne and Long, 1976) and analysis of our corpus, labels of the phases are created, including, (P0) *opening*, (P1) *problem presentation*, (P2) *history taking*, (P3) *physical examination*, (P4) *diagnosis*, (P5) *treatment recommendation*, (P6) *lifeworld discussion*, (P7) *closing*, (P8) *additional health concerns*. Disruptions (e.g., physicians interrupted by calls) are also common in medical conversation and we mark up them as (P9) *unrelated activities*.²

Despite that past research has shown these constituent phases are ordered in a normative sense,

²Labels of phases can be adjusted for different types of medical conversation.

it is not unusual that the physicians and patients go back and forth between these phases. By annotating which phase a turn belongs to, we are not only able to show where the boundaries are among the phases in medical conversation, but also how transitions are coordinated by the participants.

Pair and turn dependency Unlike many other types of discourse, conversation is interactive in nature. Thus, turns in conversation cannot be understood alone. Instead, each turn should be understood regarding whether they set up an expectation for a next turn or fulfill the expectation set up by a prior turn. Pairs of turns which are linked by conditional relevance is referred to as *adjacency pairs* and considered the building block of conversation (Sacks et al., 1974; Schegloff, 2007).

Based on this idea, turns are annotated with respect to which turn they are connected to, and within a pair, a first pair part is distinguished from its second part (e.g., question vs. answer, request vs. grant, greeting vs. return greeting, etc.).

This type of dependency relationship between two turns in conversation has been attended to in related work such as the SWBD-DAMSL coding scheme (*forward-communicative-function and backward-communication-function*) (Jurafsky et al., 1997; Core and Allen, 1997). It has demonstrated significant value for NLP tasks such as dialog act modeling (Stolcke et al., 2000).

Sequence and pair dependency As mentioned above, *adjacency pair* is considered the most basic unit of conversation (Sacks et al., 1974; Schegloff, 2007). It is also considered as the minimal form of *sequence* in conversation. In the most ideal scenario, a sequence is complete with one base pair (e.g., a question gets its answer, an invitation gets its acceptance), as shown in Ex 2. But more commonly, pairs are expanded to accommodate various types of contingencies in interaction (e.g., repairing a problem of hearing, checking understanding) (Schegloff, 1980, 1992, 1997), as shown in Ex 3.

In these cases, several pairs cluster into a coherent sequence, with one base pair and multiple expansion pairs dependent on it. Based on the sequential position of the base pair and the expansion pairs, there are pre-expansions, insert-expansions, and post-expansions (Schegloff, 2007).

Based on this idea, the COSTA scheme marks up the dependency relationship between pairs and distinguishes the base pairs from the expansion

pairs. In Figure 1, we illustrate a sequence with an insert-expansion pair in *phase n-1*, and a sequence with a pre-expansion pair in *phase n+1*.

In sum, the hierarchical annotation scheme of the COSTA describes the deep structure of medical conversation. It thus allows for systematic study of conversation at multi-level granularity, including *phases, sequences, pairs, and turns*.

4.2 Application-dependent labels

According to researchers’ specific research interests and application scenarios, additional labels can be created and added to any particular level of conversation (e.g., *phase, sequence, pair, turn*). Systematic analyses of the labeled data in the medical conversation corpus can help provide answers to various kind of research questions.

In this study, we ask the following questions: (1) How does a medical conversation start? (2) Where do communication problems tend to occur? (3) How do physicians close a conversation?

To answer question (1), we examine the labels of the initial *phase* of the medical conversation. If a conversation is opened with physicians and patient caregivers identifying and greeting each other, it is annotated as (*P0*) *Opening phase*; if the opening involves physician asking and/or patient presenting health problems, it is marked as (*P1*) *Problem presentation*.

To answer question (2), we examine the organization of *sequence*. Sequences that consist of only one base pair without any expansions are considered as produced with less difficulty. This is compared with sequences that consist of multiple pairs, with the base pair expanded with several dependent pairs. Communication problems tend to occur in phases where there are more expanded sequences.

To answer question (3), we create labels to distinguish different types of physicians’ closure initiation actions. Based on past research on closing in medical conversation (West, 2006) and preliminary analysis of our data, physicians’ closure initiation actions can be classified into three types: (*a*) *making future arrangement*, (*b*) *summarizing treatment plans*, and (*c*) *announcing closures*. Examples of the three types of action are in Ex 4-6.

Item	Number
Number of Visits	187
Number of Hospitals	5
Number of Physicians	9
Number of Patients	187
Average length of a visit	4.9 minutes

Table 1: Meta information of the subset in this study.

Item	Total	Average per visit
Characters	275,303	1472.2
Words	158,798	849.2
Turns	23,060	123.3
Pairs	11,833	63.3
Sequences	5,359	28.7

Table 2: Statistics of the subset in this study. Total number of visits in this subset is 187.

5 Results

5.1 Corpus Statistics

As it is an ongoing project, here we present some statistics based on a subset of acute visits in the corpus. Table 1 shows the meta data of the subset. Table 2 shows the statistics of the transcribed data of this subset. All the experimental results in this section are based on this subset.

5.2 How does medical conversation start?

Since medical conversation is both a social encounter where relationship is built and a task-oriented activity organized with a clear goal, we find that medical conversation in our corpus start with either the (*P0*) *Opening phase* or the (*P1*) *Problem presentation phase*. Table 3 describes the distribution of the two types of conversation opening in our dataset. As shown in the table, a majority of the conversation starts with participants going directly to discuss the health problem of patients.

This is compared with ordinary conversation, in which the initial exchanges almost always involve a *summons-answer* sequence (SA sequence) (Schefflo, 1968). Typical SA sequences include *telephone ring-hello, Johnny?-Yes, Bill-looks up*, etc. After the channel for communication is established through the SA sequence, the conversationalists then proceed to the reason for the talk.

When comparing with the findings on conversation opening in the American primary care, we find that there is a small variance in the distribution of the two types of opening. In the American primary care conversation, it is reported that less than 10% of the cases are opened with the (*P0*) *Opening phase* (Heritage and Robinson, 2006; Robinson

Conversational opening type	Count	%
(P0) Opening phase	62	33
(P1) Problem presentation phase	125	67
Total	187	100

Table 3: Distribution of the two types of conversational opening in the subset.

Phase type	Seq # per phase	Turn # per seq
(P0) Opening phase	1.67	2.15
(P1) Problem presentation phase	3.11	3.95
(P2) History-taking phase	6.50	4.71
(P3) Physical examination phase	3.03	3.21
(P4) Diagnosis phase	2.12	4.49
(P5) Treatment phase	5.32	6.63
(P6) Lifeworld discussion phase	3.18	3.01
(P7) Closing phase	2.14	3.72
(P8) Additional problem phase	0.83	4.13

Table 4: Average number of sequences in each phase of medical conversation and average number of turns in each sequence in those phases. The total number of conversation is 187. (P9) *Unrelated activities phase* is not included in this table, as they do not directly contribute to the understanding or progressivity of the conversation.

and Heritage, 2005). However, in the Chinese medical conversation, 33% of the cases are opened with (P0) (see Table 3). Thus, in a greater proportion of the Chinese medical conversation, the participants do engage in social activities, such as identity confirmation, greetings, or even ‘intimacy ploy’.

It should be noted that although the findings from the Chinese pediatric primary care are not directly comparable to that in the American primary care setting, this distributional variance in conversational opening highlights the difference in the norms and service procedures of medical interaction in two cultures. Specifically, while patients in the American primary care are normally received by nurses or medical assistants first in their medical visits, patients are directly seen by their physicians in the Chinese consultation room. The higher proportion of the *opening phase* in the Chinese corpus thus can be explained by the practical constraints that physicians have to confirm patients’ identity at the beginning of the medical consultation.

5.3 Where do problems tend to occur?

When examining the *process and overall organization* of the medical conversation, we find that there are considerable variances in the shape of various *phases*. Table 4 shows the average number of sequences and turns in each type of phase in the subset of the corpus.

Among all the phases, the *treatment recommen-*

dation phase is where sequences are most likely to be expanded. Specifically, a sequence in the *treatment recommendation phase* takes an average of 6.63 turns to complete. In comparison, the average number of turns for a sequence in the *problem presentation* take 3.95 turns; and that number is the lowest in the *opening phase*, averaging 2.15 turns.

Looking into the *sequences* in the treatment recommendation phase, it is observed that physicians’ treatment recommendations are not always immediately accepted by patient caregivers in the next turn. In face of such patient resistance, physicians must to pursue caregivers’ acceptance, and the sequence continues to expand until the patients’ explicit acceptance is displayed (as shown in Ex 3).

In addition, in our prior work, we labeled and analyzed the caregivers’ *actions* that they use to overtly advocate for antibiotic treatment in the *treatment recommendation phase*. The results showed that, when caregivers use one or more of the following actions *a) explicit requests for antibiotics, b) statements of desire for antibiotics, c) inquiries about antibiotics, and d) evaluations of treatment effectiveness*, the likelihoods of them receiving antibiotic prescriptions from the physicians increased by over 9 times (Odds Ratio = 9.23, 95% Confidence Interval = 3.30-33.08) (Wang et al., 2018). This finding corroborates the fact that antibiotic over-prescription is prevalent in the Chinese pediatric primary care (Li et al., 2012), and parental pressure on physicians in medical conversation plays a significant role in antibiotic over-prescription (Stivers et al., 2003; Mangione-Smith et al., 1999).

5.4 How do physicians close a conversation?

Closing a medical conversation is a delicate matter, as physicians and patients may have conflicting agendas. While patients may still have unmentioned concerns, physicians may have to terminate the conversation so as to move to the next patient.

To deal with such practical challenge, we find that physicians use several types of actions to initiate the closure of medical conversation. These actions include: *a) making arrangement for future activities, b) summarizing the topic-in-progress, and c) announcing closures*. Table 5 illustrates the relative distribution of these three types of action design in our corpus. Examples of the three types of actions are shown in Ex 4-6 in Section 3.

Compared with closing in ordinary conversation,

Closure initiation actions	Count	%
(a) Making future arrangements	78	52
(b) Summarizing treatment plans	57	38
(c) Announcing closures	15	10
Total	150	100

Table 5: Closure initiation actions and their distributions in the Chinese medical conversation. The total number of cases in this table is 150. In the remaining 37 cases, closures are initiated by caregivers and are excluded from the analysis.

the range of actions that physicians use to initiate medical conversation closures are highly similar. After the topical closure attempts are accepted by the caregiver or the patient, participants move on to the pre-closing sequence, in which they pass the floor to one another and confirm there is nothing more to talk about. Once the warrant to terminate is established, they move on to the termination sequence, in which they exchange farewell (bye-bye), display appreciation (thanks-you're welcome), or acknowledge the closure of the conversation (ok-ok) (Sacks and Schegloff, 1973).

When comparing our findings with related findings in the American primary care, we find that there exist some important variances. Besides the three types of actions that the Chinese physicians use, there is another type of action observed in the US data: *checking patients' unmet concerns* (e.g., 'Do you have some other problems that you want to talk about?') (West, 2006).

Again, although the findings from the Chinese pediatric primary care are not directly comparable to that in the American primary care, this difference in the range of action designs that the physicians use highlights the practical problems and constraints that exist in the Chinese pediatric setting. In the Chinese medical setting, and urban tertiary hospitals in particular, physicians are commonly overloaded (Hu and Zhang, 2016). In a day, a physician could see as many as 100 patients, and the length of the medical conversation tend to be very short, averaging 4.9 minutes for each conversation in our corpus. Absence of this action (i.e., checking patients' unmet concerns) in the Chinese corpus can be partially attributable to this.

6 Discussion

In this section, we discuss several potential use cases of this study.

6.1 Facilitating conversational understanding

One reason that conversation understanding is difficult is because the meaning of utterances often

depends on the context. For instance, the word *yeah* as a response to a yes-no question is doing the action of agreeing. In contrast, the word *yeah* uttered by a speaker when another speaker is in the middle of a long stretch of talk may indicate that the former is listening to the latter; it does not mean that the former agrees with the latter.

Moreover, if we treat the conversation simply as a sequence of turns without internal structure, multi-turn understanding may not be easily achieved. The idea of internal structure is that turns in conversation are not like beads on a string; instead, they are organized in coherent clusters. As a result, the two turns within an adjacency pair are not always adjacent. For instance, in Ex 3, Lines 1 and 4 form an adjacency pair, with the word *Okay* in Line 4 responding to the treatment recommendation made in Line 1. Lines 2-3 in between form another adjacency pair, which is an insert-expansion pair of the base pair. These dependent pairs form up one coherent sequence, and sequences of similar kind form up a coherent phase in conversation.

Thus, the whole conversation is represented as a tree structure, similar to dependency structure for a sentence. Compared with treating the conversation as a sequence of turns, having such tree structure information would make it much easier to infer that the word *Okay* in Line 4 indicates the acceptance of the recommendation in Line 1. In this sense, the conversational structural information helps multi-turn conversational understanding.

Manually annotating such a tree structure is labor intensive and time consuming, but once such a corpus is created, automatic tools can be trained on the corpus, the same way that dependency parsers are trained on treebanks. The tools can then be used to process new conversations. Our corpus thus is the first conversation treebank annotated with conversational structures and actions according to the COSTA scheme. In the scheme, the label set (e.g., phase labels and action labels) is application-dependent, whereas the structure levels (e.g., *phase*, *sequence*, *pair*, *turn*) should remain mostly the same for many applications.

6.2 Extracting information from medical conversation

Due to the nature of medical conversation, there are often natural correspondences between phases in medical conversation and sections in Electronic Health Record (EHR). For example, *problem pre-*

sentation phase in the medical conversation corresponds to *symptom section* in the EHR; *treatment phase* in the medical conversation corresponds to *prescription section* in the EHR, etc. Therefore, the medical conversation structures and labels (e.g., phase types) could provide valuable cues when building NLP systems for tasks such as information extraction. For instance, prescribed medication is more likely to appear in the *treatment phase* of medical conversation, rather than *opening* or *closing phase*. While the *history-taking phase* may also contain medication names, such medication concerns primarily with the medication history of the patient, rather than the medication prescribed during the current visit.

Apart from information extraction, the structural representation of medical conversation can help other NLP tasks such as automatic summarization of patient medical visit. Building high-quality BioNLP systems for such tasks has great potential to reduce physicians' workload and increase the time they spend on treating patients.

6.3 Conducting more communication-related research with automatically processed data

Our current study looks at some of the major issues in medical communication. For future work, we plan to apply the same methodology to other issues in medical communication and conversations in other domains.

While the current study relies on manual annotation of the conversation, once NLP tools have been trained on annotated data (as described above), we can use the tools to analyze a large amount of new conversations automatically, and significantly speed up the analytical process of conversation.

7 Conclusion

In this paper, we introduced some of the major issues that existing medical conversation research has focused on; we described the data that we use for conducting medical conversation research. To analyze medical conversation more systematically, we proposed an annotation scheme, which can capture the hierarchical structure of the medical conversation and be extended to include application-specific labels. Based on a subset of the annotated data, we report findings regarding how medical conversation is opened and closed in the Chinese pediatric consultations and how one can identify

places that problems tend to occur.

This study makes several contributions to medical conversation research. First, to our best knowledge, the corpus that we are constructing is the first medical conversation dataset with structural annotation. It is a valuable resource for conducting medical communication research, and can also be used to train NLP systems such as a conversation parser. Second, COSTA is a general scheme for annotating conversational structures and actions. The annotation facilitates systematic analysis of medical conversation and there are other potential use cases as outlined in the previous section. While we use the scheme to build a Chinese corpus consisting of medical conversations, COSTA can be applied to conversation in other domains or in other languages.

For future work, we will finish annotations of the corpus and release it to the public. We will start training NLP tools in order to evaluate the usefulness of the corpus for the use cases mentioned above.

References

- Patrick S. Byrne and Berrie E. Long. 1976. *Doctors Talking to Patients: A Study of the Verbal Behaviour of General Practitioners Consulting in Their Surgeries*. H.M. Stationery Office, London, UK.
- Mark Core and James Allen. 1997. Coding Dialogs with the DAMSL Annotation Scheme. In *Proceedings of AAAI Fall Symposium on Communicative Action in Humans and Machines*.
- Paul Drew, John Chatwin, and Sarah Collins. 2001. [Conversation analysis: A method for research into interactions between patients and health-care professional](#). *Health Expectations*, 4(1):58–70.
- John Heritage and Douglas W. Maynard. 2006. *Communication in Medical Care: Interaction between Primary Care Physicians and Patients*. Cambridge University Press, Cambridge.
- John Heritage and Jeffrey D. Robinson. 2006. [The structure of patients' presenting concerns: Physicians' opening questions](#). *Health Communication*, 19(2):89–102.
- John Heritage and Sue Sefi. 1992. Dilemmas of advice: Aspects of the delivery and reception of advice in interactions between health visitors and first time mothers. In Paul Drew and John Heritage, editors, *Talk at Work*, pages 3–65. Cambridge University Press, Cambridge.
- Yinhuan Hu and Zixia Zhang. 2016. Skilled doctors in tertiary hospitals are already overworked in china. *The Lancet Global Health*, 3(12):e737.

- Gail Jefferson. 2004. Glossary of transcript symbols with an introduction. In Gene H. Lerner, editor, *Conversation Analysis: Studies from the First Generation*, chapter 2, pages 13–31. John Benjamins, Amsterdam / Philadelphia.
- Daniel Jurafsky, Elizabeth Shriberg, and Debra Bisca. 1997. Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual, Draft 13. Technical report, University of Colorado, Boulder.
- Barbara M. Korsch and Vida F. Negrete. 1981. [Doctor-patient communication](#). *Scientific American*, 227(2):66–74.
- Yongbin Li, Jing Xu, Fang Wang, Bin Wang, Liqun Liu, Wanli Hou, Hong Fan, Yeqing Tong, Juan Zhang, and Zuxun Lu. 2012. Overprescribing in china, driven by financial incentives, results in very high use of antibiotics, injections, and corticosteroids. *Health Affairs (Project Hope)*, 31(5):1075–1082.
- Rita Mangione-Smith, Elizabeth McGlynn, Marc N. Elliott, Paul Krogstad, and Robert H. Brook. 1999. The relationship between perceived parental expectations and pediatrician antimicrobial prescribing behavior. *Pediatrics*, 103(4):711–718.
- Douglas W. Maynard. 1997. [The news delivery sequence: Bad news and good news in conversational interaction](#). *Research on Language and Social Interaction*, 30(2):93–130.
- NHS. 2018. Data on Written Complaints in the NHS 2017-18. Technical report, Health and Social Care Information Center.
- Jeffery D. Robinson. 2003. An interactional structure of medical activities during acute visits and its implications for patient’s participation. *Health Communication*, 15(1):27–57.
- Jeffery D. Robinson and John Heritage. 2005. [The structure of patients’ presenting concerns: the completion relevance of current symptoms](#). *Journal of the Association for Computing Machinery*, 61(2):481–493.
- Debra Roter and Susan Larson. 2002. [The roter interaction analysis system \(rias\): Utility and flexibility for analysis of medical interactions](#). *Patient Education and Counseling*, 42:243–251.
- Harvey Sacks and Emanuel A. Schegloff. 1973. Opening up closings. *Semiotica*, 8(4):289–327.
- Harvey Sacks, Emanuel A. Schegloff, and Gail Jefferson. 1974. A Simplest Systematics for the Organization of Turn-taking for Conversation. *Language*, 50(4, Part 1):696–735.
- Emanuel A. Schegloff. 1968. Sequencing in Conversational Openings. *American Anthropologist*, 70(6):1075–1095.
- Emanuel A. Schegloff. 1980. Preliminaries to Preliminaries: “Can I Ask You a Question?”. *Sociological Inquiry*, 50(3-4):104–152.
- Emanuel A. Schegloff. 1992. Repair After Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation. *American Journal of Sociology*, 97(5):1295–1345.
- Emanuel A. Schegloff. 1997. Third Turn Repair. In R. Gregory Guy, Crawford Feagin, Deborah Schiffrin, and John Baugh, editors, *Towards a Social Science of Language: Papers in Honor of William Labov. Volume 2: Social Interaction and Discourse Structures*, page 31–40. John Benjamins.
- Emanuel A. Schegloff. 2007. *Sequence Organization in Interaction: Volume 1: A Primer in Conversation Analysis*. Cambridge University Press.
- Tanya Stivers. 2005. Parent resistance to physicians’ treatment recommendations: One resource for initiating a negotiation of the treatment decision. *Health Communication*, 18(1):41–74.
- Tanya Stivers, Rita Mangione-Smith, Marc N. Elliott, Laurie McDonald, and John Heritage. 2003. Why do physicians think parents expect antibiotics? what parents report vs what physicians believe. *The Journal of Family Practice*, 52(2):140–148.
- Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol Van Ess-Dykema, and Marie Meteer. 2000. Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech. *Computational Linguistics*, 26(3):339–373.
- Nan Wang, Yan Song, and Fei Xia. 2018. [Coding structures and actions with the COSTA scheme in medical conversations](#). In *Proceedings of the BioNLP 2018 workshop*, pages 76–86, Melbourne, Australia. Association for Computational Linguistics.
- Candace West. 2006. Coordinating Closing in Primary Care Visits: Producing Continuity of Care. In John Heritage and Douglas W. Maynard, editors, *Communication in Medical Care: Interaction between Primary Care Physicians and Patients*, pages 379–414. Cambridge University Press.
- Fei Xia, Martha Palmer, Nianwen Xue, Mary Ellen Okurowski, John Kovarik, Fudong Chiou, Shizhe Huang, Tony Kroch, and Mitch Marcus. 2000. Developing Guidelines and Ensuring Consistency for Chinese Text Annotation. In *Proceedings of the Second Language Resources and Evaluation Conference (LREC)*.