The Communicative Activity of 'Making Suggestions' as an Interactional Process: Towards a Dialog Model for HAI

Christiane Opfermann and Karola Pitsch
University of Duisburg-Essen
Essen, Germany
christiane.opfermann@uni-due.de
karola.pitsch@uni-due.de

Ramin Yaghoubzadeh and Stefan Kopp
CITEC, University of Bielefeld
Bielefeld, Germany
ryaghoubzadeh@uni-bielefeld.de
skopp@techfak.uni-bielefeld.de

ABSTRACT
Dialog modeling of making suggestions in human-agent interaction is a challenge due to the socially delicate nature of a suggestion and ensuing interactional negotiations. A basic first dialog model for making suggestions was tested in the context of schedule management assistance by an embodied conversational agent with elderly and mildly cognitively impaired persons. Analysis showed that users responded according to human social structures with most response types bearing potential challenges concerning the system's language understanding and the users' intention interpretation: next to explicit answers, users produced implicit versions for acceptance or resistance and further requests for information or modifications. Thus, an enhanced dialog model with a newly added clarification sequence and a new multi-conditional entry sequence was tested in a second study with the autonomous system. Initial observations show a promising performance of the dialog model.

Author Keywords
Dialog modeling; Suggestions; Yes-no-question turn design; (Non-)Conforming answers; Conversation Analysis.

ACM Classification Keywords
H.5.2 User Interfaces – Interaction styles.

INTRODUCTION
The activity of 'making a suggestion' is a socially delicate matter that can lead to expansive interactional negotiations because social relationship and personal dispositions come into play [5, 13, 33]: The participant who makes a suggestion promotes a future action to be executed by the recipient only [5]. Thereby, the participant who makes a suggestion deals with matters lying in the personal domain of the recipient's decisions and actions.

HAI '17, October 17–20, 2017, Bielefeld, Germany

ACM ISBN 978-1-4503-5113-3/17/10⋯$15.00
https://doi.org/10.1145/3125739.3125752

Thus, to endow an autonomous robot system or embodied conversational agent with the functionality of 'making suggestions'[1] not only requires careful design of the initial suggestion turn, but also the ability to engage in sequential interaction dealing with the typical human practices for reacting to suggestions. 'Making suggestions' becomes in particular important when providing assistive functions in the planning and management of schedules and spare time activities. The virtual agent system "Billie" attempts to offer such domain specific tasks in order to help elderly and mildly cognitively impaired persons to maintain their daily routines and improving their leisure time activities [40, 41].

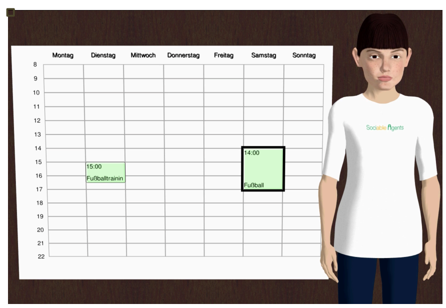

Figure 1. Setting of Embodied Conversational Agent "Billie" and its virtual calendar deployed as Assistive System.

So far, research on human-agent interaction (HAI) has tackled the problem of assistive functions more with focus on socially and/or functionally adequate request behavior [e.g. 10, 16, 43, 44], giving instructions [e.g. 35], entering scheduled tasks [25, 40] and reminding of appointments [e.g. 21, 25]. Yet, how a robotic system can initiate a suggestion and interactively deal with the users' reactions to it has been widely neglected.

This paper presents a reiterative data-based dialog modeling process with two successive human-agent interaction studies. It investigates into adequate dialog structures for a system-initiated suggestion that should enable the system to interpret adequately the user's reactions. Each video-recorded study was carried out with the agent system

[1] By suggestion, we refer to the technical term 'Suggestion*' as described in the taxonomy of directive-commissive social actions by the interactional linguist Couper-Kuhlen [5].

"Billie" and different user groups, i.e. elderly people, cognitively impaired persons and students. Analysis adopts a qualitative approach informed by Conversation Analysis (CA) [32] and addresses the following questions:

(1) How do users react to a system's yes/no-question for making a suggestion and for entering it into a virtual calendar? Are yes/no-questions functional for inviting users to produce explicit and conforming yes/no-answers?

(2) What are system-internal processing requirements and external dialog structures for a system's autonomous dealing with user reactions?

The dialog model for suggestions in the context of schedule management was developed in three steps: First, a Wizard-of-Oz (WOz) study (Study I) provided for a test of the basic design of the suggestion and entry turns in yes/no-question format. Second, based on the qualitative analyses of user reactions, implications were derived for an enhanced dialog model. Third, the enhanced dialog model was implemented in a study with the autonomous "Billie" system (Study II), in which a newly added clarification sequence and a newly designed multi-conditional entry sequence were tested.

Analysis shows that the basic dialog resources for making suggestions deployed in study I are not sufficient for dealing with the variety of human reactions to suggestions. Initial analysis of the enhanced dialog model in study II suggests that it seems to be functional.

SUGGESTIONS IN HAI & HHI

Suggestions in Human-Agent Interaction (HAI)

Many speech based applications and robotic systems are designed to assist humans with various kinds of physical or mental deficits in order to promote their autonomy in every-day activities (e.g. [3, 25, 38]). Assisting entails – among a range of other activities – actions such as *requests, offers, proposals, suggestions,* from both user or assistive system. Research in the field of human-agent interaction (HAI) has considered such actions primarily in terms of *system-initiated request behavior,* e.g. see [43] for human compliance with a robot-initiated request versus human-initiated request, see [44] for offering assistance in multi-party interaction, see [10] for robot-initiated interaction for getting help by humans. Kimoto et al. [16] also investigate into so-called "explicit and implicit suggestions" by a robot in object reference conversations. Their study design and analysis show, however, a *substantial difference in the use of the term 'suggestion'*: Kimoto et al. (2016) investigate into linguistic forms of *requesting* users to choose one object among several objects, and into different ways of *requests-for-confirmation* after users' selection, which differs substantially from the technical definition of suggestions used in the presented work.

Suggestions in Human-Human Interaction (HHI)

Conversation analytic informed research on human-human interaction (HHI) on 'making suggestions' serves as an inspiration for our dialog design in HAI [see for previous HHI-inspired work e.g. 23, 24, 35]. Among the substantial amount of empirical studies focusing on social actions that 'recruit' another person in some kind of temporally immediate or remote collaboration or assistance [15, 34], such as *requests, offers, proposals and suggestions*, Couper-Kuhlen [5] has carried out an interactional linguistic [6] corpus analysis. She presents a technical definition[2] of Proposals*, Offers*, Request* and Sugges-tions* that is based on two questions which co-participants orient to during interaction: (i) Who is the agent of the future action? (ii) Who will benefit from the promoted action [p. 625f.]? Given that the co-participant who initiates such social action is called 'self' and the recipient 'other', a taxonomy of four sub-classes is presented as follows:

Action term	***Agent of future action***	***Beneficiary of future action***
Proposal*	self & other	self & other
Offer*	self	other
Request*	other	self
Suggestion*	other	other

Table 1. Empirically based distinctive dimensions of the presented social actions according to Couper-Kuhlen, p. 634

A suggestion in the sense of an interactional-linguistic technical term is thereby defined as "an action type advocating a future action or activity to be carried out by the recipient that will benefit the recipient" [5, p. 633]. This applies exactly to the virtual agent's appointment sugges-tions as their enactment relies entirely on the user and is assumed to promote his/her leisure time activities.

Furthermore, CA studies provide following findings about the interactional features of suggestions:

1. Social (A)Symmetry: As a suggestion constitutes a social action in which "a future action is being promoted for the benefit of the recipient" [5, p. 623] and the promoted action is to be executed by the recipient solely [p. 633f.], it might put the suggestion producer in danger of placing him-/herself in a superior position to his/her interlocutor which might evoke rejections [5] or other actions with respect to the marked social status.

2. Recipient's Willingness & Ability: The interactive establishment of a suggestion presupposes the recipient's willingness and ability to carry out the suggested action [13, p. 3f.]. This might lead to marking or discussing potential problems and uncertainties before turning the initial sugges-tion into an arrangement [13, p. 2].

3. Deontic authority: The *right to decide what action is to be brought about in the future by oneself and others* is part

[2] The asterisks in the following action terms indicate that these are tech-nical terms and not lay terms.

of the interactive establishment of suggestions. Especially the initiation of a suggestion marks the *deontic status and stance* of the producer in relation to the recipient [33], which might lead to interactive negotiation concerning rights and obligations [33].

4. Reaction Structures: Suggestions have inherent conversational structures concerning the response-relation to the initiating action[3] [5, 27, 29]. This means that the initiating action like a suggestions-question has *alternative responses*, i.e. (at least) an *acceptance* or a *resistance* (see also Figure 2). The response types are differentiated as follows:

- **(+)-type response:** responsive action that aligns with the projected action, e.g. acceptance to a suggestion.
- **(–)-type response:** responsive action that disaligns with the projected action and establishes a distancing from it, e.g. resistance to a suggestion.

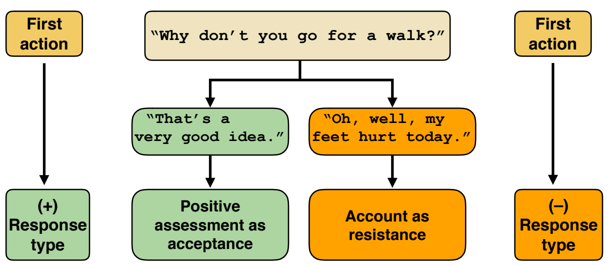

Figure 2. Conventional human reaction structures to suggestions.

Conventionally, *response types are communicatively marked in their type-relation to the initiated action.* (+)-types often come quickly and in an explicit manner. (–)-types are marked by different conversational practices, like delay of reply by longer pauses and hesitations ("uhm"), interjections ("oh") or explanations for resistance, i.e. so-called 'accounts'. Thus, type-related structured social actions like suggestions entail a wider variety of expectable and alternative response types in HHI which makes dialog modeling of them in HAI more challenging.

METHOD

Reiterative System Design

Similar to other HHI-inspired HAI studies [e.g. 23, 24, 35], the presented work was carried out in a research cycle. First, an HHI-inspired dialog model was tested using the Wizard-of-Oz method [26]. Second, based on these findings, an enhanced dialog model was developed. Third, the enhanced dialog model was programed for and tested with an autonomous system prototype, resulting in first initial observations that lay ground for further optimization.

Analytical Method

Analysis draws on ethnomethodological CA [32]. This entails the data-based sequential analysis of social actions of the observed participants (*turn taking and forming of actions*) on the grounds of the multimodal resources they use for producing their actions (*turn construction*). In contrast to other empirical approaches, CA informed analysis focuses on examining participants' observable actions as they are captured in audio- or video-recordings, and aims at revealing the participants' own orientation to what they think they do "displayed in their own conduct" [31, p. 79].

STUDY I: SYSTEM'S TURN DESIGN AND DIALOG STRUCTURES

Basic Dialog Model of Appointment Suggestions

In order to investigate making appointment suggestion (AS) in a setting with a virtual agent and its virtual calendar, it was first considered what turn design and dialog structures would be apt with respect to the known human interactional features of suggestions, and the technical requirements.

- The form and structure of the system's suggestion should be in a polite and most facticity-oriented way in order to avoid user impressions of experiencing a social put-down or patronizing. Therefore, the main suggestion sequence was designed as a *two-part sequence*, consisting first of an *information delivery* of the suggested event, entailing all necessary information items like day-of-week, start time and activity. After that, the user is *asked for* his/her *commitment*, i.e. whether he/she would like to participate in the suggested event.
- The linguistic form of the system's turns should be shaped in a way that would elicit *most simple and explicit user responses* in order to enhance natural language understanding (NLU) and improve user intention interpretation by the SDS. Therefore, the system's main functional turns are designed in form of *yes/no-questions* that make relevant as next action a type-specific and linguistically conforming yes/no-response.
- The suggested leisure time event ought to be an appealing and doable activity in order to avoid out-of-domain user responses, like trouble-telling or positioning activities [2, 17]. Thus, the more or less group specific appointment suggestions were known social events offered by the institution from which the participants were recruited.

Besides these reasons for the specific design decision, it was an open question if users tend to produce their answers to the system's yes/no-questions within typical human interactional structures, or if they would adopt a more technical stance towards the virtual agent and use a simplified register [8] for answering the system's questions.

In sum, the first basic dialog model for AS was organized in its potential course of interaction according to Figure 3.

Procedure, participants, data

Study

To investigate the functionality of the dialog design for appointment suggestions, a Wizard-of-Oz (WOz) study was

[3] In CA terminology, this topic is called 'preference structure' and is extensively explained in its technical definitions in [27] and [29].

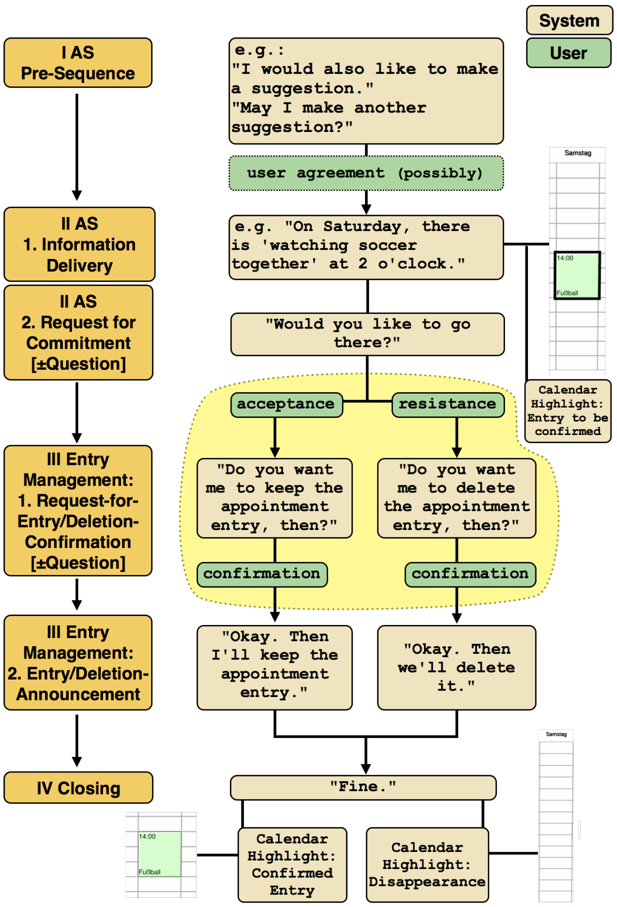

Figure 3. Basic dialog model of appointment suggestion (AS) and its entry into calendar. Left: functional description of the sub-tasks I-IV. Right: expected dialog traversals; brown: possible system actions, selected by wizard; green: expected (high-level) user actions. The yellow dotted outline reflects the region of interest for the present analysis.

conducted in which two schedule management tasks were explored: appointment entries (hereafter AE) and the interactive establishment of appointment suggestions (AS) of which only AS will be focused on in this paper. The AS-tasks were of two kinds: three complete appointment suggestions were issued in which the system provided all basic information items of the suggested event, like day-of-week, start time, end time and activity (see Figure 3, II AS, part 1). One AS was initiated by the system by suggesting an activity only (ASI for AS-initiation). In case of acceptance, the missing information items were asked for and entered [see 20 for a detailed description and findings]. The interaction was video-recorded with three external HD-cameras, a screen capturing and a stationary eyetracker (Tobii X2-60) beneath the monitor. Two wizards were deployed who had been given minimal instructions: to follow the designed interaction script and otherwise to act according to their natural interaction competence. Participants were advised to enter up to 10 appointments. They had not been told that the system would make appointment suggestions. Therefore, the introductory sequences were formulated in different ways as each AS was issued at different points during the interaction, but in the same order to all participants (see Figure 4).

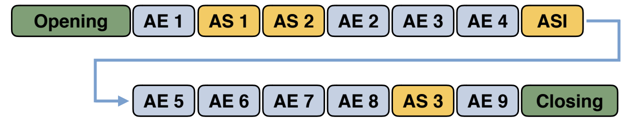

Figure 4. Study I: Chronological course of schedule management tasks in interaction with virtual agent; order and number of free appointment entries (AE) by user and appointment suggestions (AS & ASI) by system according to WOz script.

Participants

The ECA system used in study I is aimed at *being potentially deployed* in an institution that is specialized in providing assistive services for citizens with various degrees of need of assistance. This institution is project partner within the cooperative project KOMPASS. Therefore, 'senior' (SEN) and 'cognitively impaired' (CIM) persons were recruited, and 18 SENs and 19 CIMs participated in the study. In addition, 16 students participated as controls (CTLs). All 53 subjects were paid volunteers. The study was authorized by an independent ethics committee for research.

Concerning the type of mild cognitive impairment (MCI) of the participants, there were no clinical information provided by the institution or the participants themselves. While clinical definitions and assessments of MCI are formulated in Neurology and Psychology [e.g. 19, 39], we treat it as an empirically open question for our analysis whether and how cognitive deficits are reflected *in the communicative conduct* by the diagnosed person. While there is sample CA research on distinctive interactional features in interactions with aphasic [11, 37] and autistic persons [e.g. 7, 18] which are *not* part of the investigated user groups in our studies, we cannot draw on apt literature concerning our *very heterogeneous group of participants*. Therefore, if data analysis shows instances of interactional challenges for a specific SEN- or CIM-participant, analysis will only be valid within the limits of the specific case analysis.

Data

The data corpus of study I encompasses about 18 hours of video recordings. All verbal exchanges between virtual agent and users are transcribed according to the conversation analytic transcription system GAT 2 [30].

FINDINGS

As a first basic result the qualitative interaction analysis reveals that all hereafter presented recurrent user practices were found in all three groups of participants. Concerning the assumed cognitive and eventually interactional impairments of some participants, there were no 'atypical' interactional practices [1] observed *within the specific context of appointment suggestions*. This means, that the analyses of AS did not reveal communicative activities that might be seen as 'deviating' from 'typical' human interaction competences by the elderly and so-called mildly cognitively impaired participants in comparison to students in study I.

Concerning the actual user reactions to the system's yes/no-turn design in the suggestion sequence (II-2, see Fig.3) and entry management sequence (III-1), our analyses show that most users produce linguistically conforming responses, i.e. simple yes- or no-answers. Some yes/no-answers have verbal extensions that show linguistic alignment [22] (e.g. "Yes, I **would like to**." or "Yes I would **go there**", see verbal design of suggestion turn (III-1) in Figure 3).

Besides such technically manageable user responses, other types of user responses occur that might constitute greater challenges for an autonomous spoken dialog system (SDS) with respect to user intention interpretation. The following section A and B focus on *implicit versions of acceptance or resistance* that also fulfill other interactional functions, like the marking of the user's knowledgeable status, deontic status or the (–)-type of response. Section C shows a typical case of *further information request*. Section D shows a case of user acceptance with *wish for entry modification*. Types A and B were manageable within the first dialog model due to the human wizard's understanding competence. C and D were not really manageable due to the model's interactional limits. All types of critical user responses show substantial need of optimizing the basic dialog model.

A) Implicit Versions of Acceptance

The following two extracts show how users produce implicit versions for confirmation as (+)-type response to the system's entry management request (III-1) (Extract 2) or for acceptance to the system's suggestion (II-2) (Extract 3). Both focal user responses are understood and treated by the human wizard as (+)-type replies, as can be seen in the wizard's subsequently issued next utterance. They might pose difficulties for an autonomous system's processing.

```
01   SYS_ver   on SAturday,                          II-1
02             there is watching SOCcer together
             at two o'clock;
03             (1.0)
     CIM_hea         |↓↑
04   CIM-044   that's |RIGHT;
05             (0.2)
06   SYS_ver   would you like to go THERE-           II-2
07             first name*;
08             (0.5)
     CIM_hea                 |↓↑
09   CIM-044   °h (2.2) me |YES,
10             (1.3)
11   SYS_ver   do you want me to KEEP the            III-1
             appointment entry then-
12             (0.9)
     CIM_hea   |↓↑
13   CIM-044   |TWO o'clock;
14             in institution* in the café name*
15             (1.9)
16   SYS_ver   o↓KAY; (-)                            III-2
17             then i'll KEEP the entry;
18             (1.5)
```

Extract 2. Implicit user response in case of confirmation, marking being "in the know" (CIM-044, AS 1)

In this case, the focal moment of analysis is in lines 13 to 14: participant CIM-044 produces a linguistically non-conforming (+)-type response to the system's yes/no-question for entry management in line 11 (III-1). His answer constitutes a confirmation and takes the form of a *repeat of the start time* of the suggested event in line 13 (`two o'clock`) that is accompanied by an affirmative head nod (see `CIM-hea`). Subsequently, CIM-044 extends the turn incrementally by naming the institution and the café where the event will take place. With this turn extension, CIM-044 *demonstrates understanding* [28] and knowledge about the suggested event, as these two information items had not been presented by the system in the information delivery in line 01[4]. Next to confirming the appointment suggestion, CIM-044 also *positions himself as a knowing and informed person*. This had first been displayed in line 04 in reaction to the previous information delivery where CIM-044 produces a confirmation with respect to the presented information (`that's RIGHT`).

While the human wizard understands and treats CIM-044's format of an answer as a confirmation, the question remains how an autonomous spoken dialog system could manage this type of implicit confirmation. The same question applies for the following case.

```
01   SYS_ver   on THURSday,                          II-1
02             there is a computer course at the
             name of location*;
03             (1.2)
04   SYS_ver   would you like to go THERE;           II-2
05             (0.3)
06   CIM-062   good iDEA-=well:-
07             i had planned THAT anyw-
08             i wanted to be taught anyway
             something by:-
09             (0.9)
10   CIM-062   per comPUter;
11             i wanted to learn a little bit
             Anyway;
12             (1.2)
13   SYS_ver   do you want me to KEEP the            III-1
             appointment entry then-
```

Extract 3. Implicit user response in case of acceptance, marking deontic status vis-à-vis system (CIM-062, AS 2)

In this case, CIM-062 produces a (+)-type, but linguistically non-conforming response in form of a positive assessment (`good iDEA-`) in line 06. Such an alternative (+)-response type to an explicit yes-response might be manageable by an autonomous system by building and implementing an acceptance-lexicon. However, CIM-062's subsequent actions might be a problem for NLU and user intention interpretation: CIM-062 extends her response turn in lines 07 to 11 by elaborating the reasons for her positive assessment and acceptance in line 06. First, she explains for her acceptance by telling that the suggested activity had already been planned by herself (line 07), thereby claiming not only to understand the suggestion and to be "in the know", but also claiming to be ahead of the virtual agent's moves. After that, this elaboration is expansively rephrased twice in two following turn extensions: CIM-062 elaborates on her wish to learn and have computer lessons (lines 08 to 11). In this vein, CIM-062 positions herself as a person who is informed and eager to learn. These elaborations and

[4] Events for AS were, however, all from the institution's newsletter of leisure time events (see previous section Basic Dialog Model) so that the suggested event was actually identical with the participant's references.

positioning activities [2, 17] might also be interpreted as CIM-062's intent of marking her superior deontic status [33] vis-à-vis the suggester.

B) Implicit Versions of Resistance

In cases of users' resisting the system's suggestion, some users deploy the typical human practices for (–)-type responses.

```
01   SYS_ver   on TUESday,                             II-1
02             there is SEniors' group at three
               o'clock;
03             (0.5)
04   SYS_ver   would you like to go THERE-             II-2
05             first name*;
06             (2.0)
07   SEN-022   on TUESday (.) i have something to
               do,
08             i have visitors at three o'clock.
08             (1.1)
09   SYS_ver   do you (.) want me to ^deLETE the       III-1
               appointment then;
10             (1.1)
11   SEN-022   yes? (.) at THIS point yes.
12             (0.8)
13   SYS_ver   o↓KAY; (-)                              III-2
```

Extract 4. Linguistically non-conforming user response in case of resistance (SEN-022, AS 1)

In line 07, SEN-022 produces a resistance to the suggestion. This is done by first stating her *inability* to participate in the suggested event. After that, SEN-022 extends her turn by stating the reason for her inability. In line 08, she *gives an account* for not being able to attend the suggested event and names the reason, which is an alternative appointment at the given time. This case demonstrates two of the known interactional features of suggestions: (i) the question of willingness and/or ability is discussed [13]; (ii) the principle of type-related responses is applied, i.e. the (–)-type of the reply is marked by a verbose turn design, in this case by giving an account. The fact that the human wizard interprets SEN-022's account correctly as resistance to the suggestion (line 09) is seen in line 11 in SEN-022's next confirmation.

C) Insertion of Information Requests

Instead of producing a reply to the system's asking for the user's commitment (II-2), some users request further information related to the suggested event (line 06 in Extract 5), necessary for the user's decision.

```
01   SYS_ver   on SUNday,                              II-1
02             there is name of serial* at a
               quarter past eight;
03             (1.3)
04   SYS_ver   would you like to go THERE;             II-2
05             (1.8)
06   SEN-023   who are the actors,
07             (2.5)
08   SYS_ver   sorry i don't know;
09             (0.9)
10   SEN-023   u HU-
11             (--) then deLETE it please;
12             (1.2)
```

Extract 5. Insertion of user request for further information before producing response to AS (SEN-023, AS 2)

While the wizard understands the user turn as a deviation from the main activity, i.e. the sequence of suggestion and acceptance or resistance, and produces an adequate response to it (see line 08), an autonomous SDS might struggle here.

D) Modified Acceptance

Some users accept the suggested event but request a modification of a specific information item. CIM-031 e.g. first produces a conforming yes-response in line 07, extending this by representing his decision as having been taken before the suggestion was made (`i wanted to do that TOO`). After a repeat of the day of concern (line 08), CIM-031 further extends his turns (line 09) by requesting to change the start time to one hour earlier than suggested. As the wizard did not have any resources for fulfilling this request, she could only continue with asking the relevant next question (line 11). That this kind of user request might be persistently relevant for the user, can be seen in CIM-031's next response: after a confirmation, CIM-031 repeats the request-for-modification for a second time (lines 14-16).

```
01   SYS_ver   on SAturday,                            II-1
02             there is watching SOCcer together
               at two o'clock;
03             (0.7)
04   SYS_ver   would you like to go THERE-             II-2
05             first name*;
06             (1.8)
07   CIM-031   (m/n)yes=i wanted to do that TOO-
08             on SAturday-
09             but from about ONE o'clock to
               FOUR o'clock;
10             (3.5)
11   SYS_ver   do you want me to KEEP the
               appointment entry then-                 III-1
12             (2.5)
13   CIM-031   uh- yes;
14             I would prefer to- (.)
15             m SAturday;
16             to enter it from one o'clock to
               FOUR o'clock;
17             (3.8)                                   III-2
```

Extract 6. Acceptance with request for modification of specific information item (CIM-031, AS 1)

DISCUSSION: IMPLICATIONS & ENHANCED MODEL OF APPOINTMENT SUGGESTIONS

Implications for System Design

The analysis has revealed that the features of the interactive establishment of suggestions, as known from human-human interaction [5, 13, 33], are also observable in human-agent interaction. Some users deploy the known practices of producing *type-relatedly marked responses* (Extract 4). Others produce longer and verbose responses when *marking their knowledgeable status* (Extract 2), marking their *deontic status* (Extract 3) or when *positioning* themselves as being ahead of the virtual agent (Extract 3 & 5). These kinds of responses show that those users treat the system as a socially equal interlocutor [9]. With respect to the system's task to process such user responses and to ascribe a distinct user intention of agreement or disagreement in the context of yes/no-questions, securing understanding seems to be challenging. Therefore, we consider the following practical implications of the results:

(i) Implicit versions of acceptance (A) or resistance (B) could become manageable for an autonomous system by adding a *clarification request* (CR) sequence to the first dialog model. If user responses to a yes/no-question are not parsable as distinct values of dis-/agreement, the system initiates a request-for-clarification, i.e. a repair initiation. The CR-sequence is designed as a three-step sequence (see Figure 5), combining a non-understanding notification, typically used in SDS [e.g. 4] with a yes/no-question representing the system's candidate understanding of the user's previous answer and a closing in form of an epistemic marker [12, 14] in case of understanding (otherwise, the CR is reprompted).

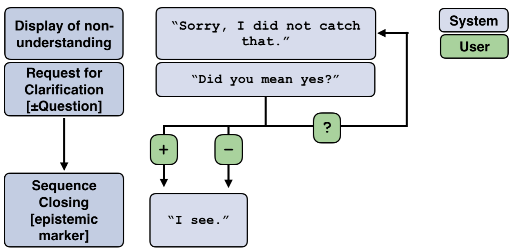

Figure 5. Clarification Sequence in cases of system's non-understanding in yes/no-contexts.

(ii) When users produce their (–)-type response according to HHI practices in form of an *account* (B), i.e. giving reasons for not being able to commit to the suggested event by telling an *alternative appointment*, this is implicative for a topic change. The same holds for user requests for entry modifications (D) or potential requests for already entered appointments which might conflict with the AS. With the dialog manager `flexdiam` [42], such a switch or topic change is manageable.

(iii) Requests for further information (C) might be tackled by starting a web search and implementing further sequence structures for securing information transfer.

Enhanced Model of Appointment Suggestions

The findings and implications suggest that human interactional structures revolving around suggestion turns and entry management turns make relevant to implement additional sequential structures dealing with user reactions formats like implicit acceptance and resistance (A & B) and user requests dealing with personal aspects of dispositions (C & D) that might be incomprehensible for an autonomous SDS. The enhanced dialog model in Figure 6 represents dialog structures tackling especially cases A & B (potentially C & D in the sense of making the user abandon his project of further requests and priming him/her for explicit yes/no-replies). It entails the implemented CR-sequence and additionally a truncated, but multi-conditional entry management sequence that is empirically grounded on the analysis of user confirmation practices in study I[5]: The blue dotted outline represents a pre-action grace period, i.e. a pre-commitment strategy of expecting either a refuting other-correction in case of a system's misunderstanding of the prior user reply, or an explicit ratification – while assuming implicit agreement after a set time of silence or the user proceeding to the next topic.

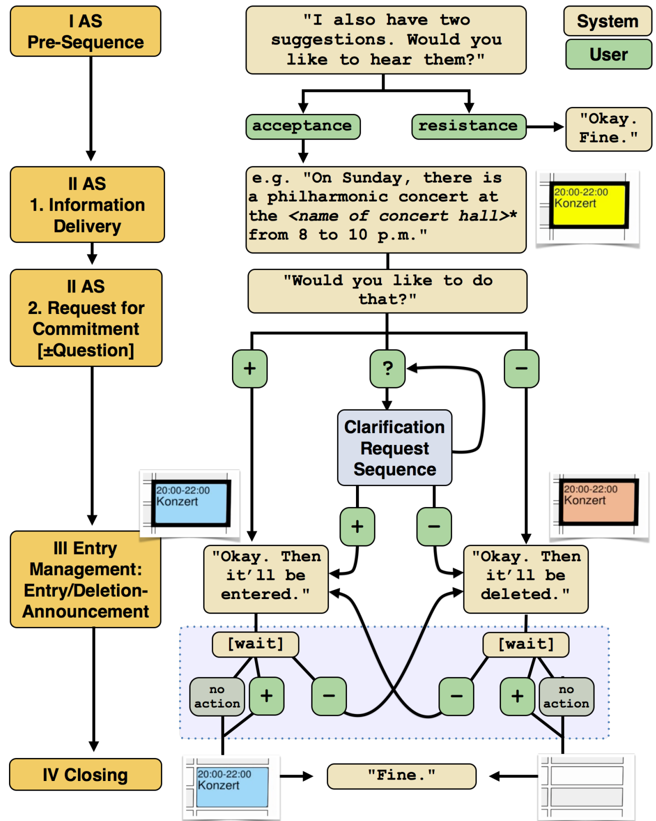

Figure 6. Enhanced dialog model of appointment suggestions. Left: functional description of sub-tasks I-IV. Right: expected dialog traversals; brown: possible system actions issued by DM; green: expected (high-level) user actions.

STUDY II: AUTONOMOUS SYSTEM WITH ENHANCED SUGGESTION MODEL

In order to test the functionality of the enhanced model of AS and its dialog structures (see Figure 6), a study was set up with the *autonomous* agent system. The technical set up for data collection was identical with the previous WOz study. Prior to the experiment a training of the ASR of the autonomous system was conducted, by either having the participants read well-known proverbs off the screen, or having them repeat the proverbs in the case of dyslexic participants. Participants were recruited via the same institution. 19 SENs, 16 CIMs and 10 CTLs participated which results in a total of 45 participants. The participants were told to enter at least about 5 to 6 personal or fictitious appointments, but they were free to enter as many as they wanted. To avoid puzzling effects on users by interrupting their AE by AS, it was decided to have the AS issued after a participant had signaled to have finished entering appointments. For this and for the closing of the interaction, an

[5] Analysis showed that the two-part entry management sequence was too explicit as users seldom produced confirmations to the entry/deletion announcement (III-2) after having confirmed the previous request (III-1). This is not shown in this paper, but documented by the first and second author in the first internal research report of 2016.

experimenter in the adjacent room was able to observe the participants and to initiate the strategy change in the dialog manager to go to fully-formed suggestion and afterwards to the closing of the interaction.

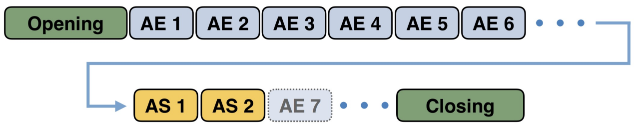

Figure 7. Study II: Chronological course of scheduling tasks in interaction with Billie; number of free AE by user varied according to user preference; two system-initiated AS after user-initiated ending of AEs; final AEs possible.

AUTONOMOUS SYSTEM

The system was driven by the incremental mixed-initiative dialog manager `flexdiam` [42]. The virtual agent was controlled by ASAPRealizer [36], using CereVoice speech synthesis. The displayed schedule items and highlights were controlled by the dialog manager (DM) directly. ASR was performed by a standalone machine running a leading brand dictation software SDK, served by an array microphone. NLU employed a contextualized template-matching parser; POS-based heuristics were used to reduce open dictation of activities to suitable candidates for short representations for verbal summaries and graphical display.

The DM was able to react to the participants' attempts to enter new appointments and their replies to system-initiated suggestions. This could also include topic shifts, which could be effected by the user entering a new appointment while the agent presented a suggestion. For the initiation of clarification requests respectively repair in the appointment dictation phase (analog to presented suggestion phase above), the estimated entropy of uncertain variables was matched against fixed thresholds to either reformulate a request for information, or else enter a candidate understanding phase in which hypotheses were successively eliminated, while at the same time elicitating corrections.

FINDINGS: INITIAL OBSERVATIONS

Concerning the functionality of the enhanced dialog model with special focus on the CR-sequence in cases of system's non-understanding of user's intention, first case analyses show encouraging results. Extract 7 shows a system's non-understanding in face of a colloquial yes-response by user.

```
01  SYS_ver   would you like to do that;        II-2
02            (1.2)
03  CIM-122   yo:;
04            (0.8)
05  SYS_ver   sorry i did not catch that-
06            (0.3)
07  CIM-122   ye:[:s;
08  SYS_ver      [did you mean yes;
09            (0.3)
10  SYS_ver   i see;
11            (0.4)
12  SYS_ver   okay; then it'll be entered;      III
```

Extract 7. Clarification Request Sequence (CIM-122)

CIM-122 accepts the AS in line 03 by answering "yo" which is an incomprehensible word for ASR-software. After the system issues the first part of the CR, i.e. the display of non-understanding (line 05), CIM-122 rephrases the previous answer into the standard "yes". This is processed successfully as the system continues with the entry-announcement (line 12) after finishing the CR-sequence.

Overall, there was a tendency towards reduced necessity for CR-sequences during the second of the two suggestion instances (see Figure 8). For the second suggestion, only three CR-sequences were ever entered; no more than one per participant (the difference was however not statistically significant; paired t-test p=0.55, p=0.15, p=0.25 for CTL, SEN, CIM, respectively). We assume that the effect is best explained by successful (short-term) adaptation by the users to the system expectations in that specific action context.

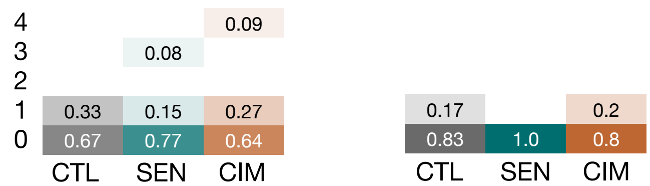

Figure 8. Histograms of the proportion of successive CR-sequence iteration counts (y-axis) by user group. Left: first AS, right: second AS. During the second AS, multiple repetitions of clarification requests were no longer observed.

CONCLUSION

A basic dialog model of system-initiated suggestions was tested in spontaneous human-agent dialog obtained in a WOz study. Analysis revealed several implicit forms of acceptance and resistance by users. In addition to direct and indirect declarations, some of the user responses reflect (quasi-)social reactions to the system. In particular, analysis showed that users' acceptance of system-initiated suggestions can be accompanied by the reassertion of their deontic primacy (e.g. their declaring that they thought about it all along anyway), while resistances can be produced in form of giving an 'account', attempting to justify or mitigate the rejection. For an enhanced dialog model for an autonomous SDS, we combined selective spotting of several known verbal acceptance and resistance patterns with a repair strategy that explicitly primed for yes/no replies, as well as a pre-action grace period that allowed for late repairs. This was suitable for operation by our user groups, i.e. older adults and people with cognitive impairments, and first results of the second study allow for the tentative assumption that basic in-situ adaptation to system limitations can also be expected from them. Further work will concentrate on an extensive analysis of the data stemming from the autonomous study. In a follow-up long-term study, the question will be if the initially observed short-term user adaptation to the system's priming to yes/no-answers might also become a persistent learning effect.

ACKNOWLEDGMENTS

All authors gratefully acknowledge the financial support from the German Federal Ministry of Education and Research in the project "KOMPASS" (FKZ 16SV7272).

REFERENCES

1. Charles Antaki and Ray Wilkinson. 2013. Conversation Analysis and the Study of Atypical Populations. In *The Handbook of Conversation Analysis*, Jack Sidnell and Tanya Stivers (eds.). Wiley-Blackwell, 533 – 550.
2. Michael Bamberg. 1999. Is there anything behind discourse? Narrative and the local accomplishments of identities. *Challenges to theoretical psychology*, 220 – 227.
3. Jenay Beer et alii. 2012. The Domesticated Robot: Design Guidelines for Assisting Older Adults to Age in Place. In *Proceedings of HRI'12*, 335 – 342.
4. Dan Bohus and Alexander Rudnicky. 2009. The RavenClaw dialog management framework: architecture and systems. *Computer Speech and Language* 23: 332 – 361.
5. Elizabeth Couper-Kuhlen. 2014. What does Grammar tell us about Action? *Journal of Pragmatics* 24, 3: 623–647.
6. Elizabeth Couper-Kuhlen and Margret Selting. 2001. Introducing Interactional Linguistics. In *Studies in Interactional Linguistics*, Margret Selting and Elizabeth Couper-Kuhlen (eds.). John Benjamins Publishing Company, Amsterdam, 1 – 22.
7. Paul Dickerson, John Rae, Penny Stribling, Kerstin Dautenhahn, and Iain Werry. 2005. Autistic Children's Co–ordination of Gaze and Talk: Re–examining the 'Asocial' Autist. In *Applying Conversation Analysis,* Keith Richards and Paul Seedhouse (eds.). Palgrave Macmillan Ltd, 19 – 37.
8. Kerstin Fischer. 2010. Why it is interesting to investigate how people talk to computers and robots. *Journal of Pragmatics* 42: 2349 – 2354.
9. Kerstin Fischer. 2011. Interpersonal Variation in Understanding Robots as Social Actors. In *Proceedings of HRI'11*, 53 – 60.
10. Kerstin Fischer, B. Soto, C. Pantofaru, and L. Takayama. 2014. Initiating Interactions in Order to Get Help: Effects of Social Framing on People's Responses to Robots' Requests for Assistance. In *Proceedings RO-MAN 2014*, 999 – 1005.
11. Charles Goodwin (ed.). 2003. *Conversation and Brain Damage*. Oxford University Press.
12. John Heritage. 1984. A change-of-state token and aspects of its sequential placement. In *Structures of Social Action. Studies in Conversation Analysis*, J. Maxwell Atkinson and John Heritage (eds.). Cambridge University Press, 299–345.
13. Hanneke Houtkoop. 1987. *Establishing Agreement. An Analysis of Proposal-Acceptance Sequence*. Foris Publications.
14. Wolfgang Imo. 2009. Konstruktion oder Funktion? Erkenntnisprozessmarker („change-of-state tokens") im Deutschen. In *Grammatik im Gespräch. Konstruktionen der Selbst- und Fremdpositionierung*, Susanne Günthner and Jörg Bücker (eds.). Walter de Gruyter, 57–86.
15. Kobin H. Kendrick and Paul Drew. 2016. Recruitment: Offers, Requests, and the Organization of Assistance in interaction. *Research on Language and Social Interaction* 49,1: 1 – 19.
16. Mitsuhiko Kimoto, Takamasa Iio, Masahiro Shiomi, Ivan Tanev, Katsumori Shimohara, and Norihiro Hagita. 2016. Alignment Approach Comparison between Implicit and Explicit Suggestions in Object Reference Conversations. In *Proceedings of HAI'16*, 193 – 200.
17. Gabriele Lucius-Hoene and Arnulf Deppermann. 2004. Narrative Identität und Positionierung. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 5 (2004): 166 – 183.
18. Tom Muskett, Mick Perkins, Judy Clegg, and Richard Body. 2009. Inflexibility as an interactional phenomenon: Using conversation analysis to re-examine a symptom of autism. *Clinical Linguistics & Phonetics*, 24, 1: 1 – 16.
19. Ziad Nasreddine et alii 2005. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool for Mild Cognitive Impairment. *Journal of the American Geriatrics Society* 53: 695 – 699.
20. Christiane Opfermann and Karola Pitsch. 2017. Reprompts as Error Handling Strategy in Human-Agent-Dialog? User Responses to a System's Display of Non-understanding. Accepted for *RO-MAN 2017*.
21. Han-Saem Park and Sung-Bae Cho. 2012. A modular design of Bayesian networks using expert knowledge: Context-aware home service robot. *Expert Systems with Applications* 39: 2629 – 2642.
22. Martin Pickering and Simon Garrod. 2006. Alignment as the Basis for Successful Communication. *Research on Language and Computation* 4: 203 – 228.
23. Karola Pitsch, Anna-Lisa Vollmer, and Manuel Mühlig. 2013. Robot feedback shapes the tutor's presentation. *Interaction Studies* 14, 2: 268 –296.
24. Karola Pitsch, Anna-Lisa Vollmer, Katharina Rohlfing, Jannik Frisch, and Britta Wrede. 2014. Tutoring in adult-child interaction. *Interaction Studies* 15,1: 55 – 98.

25. Martha Pollack. 2005. Intelligent Technology for an Aging Population. *AI Magazine* 26, 2: 9 – 24.

26. Laurel Riek. 2012. Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines. *Journal of Human-Robot Interaction* 1, 1: 119 – 136.

27. Harvey Sacks. 1987. On the preference for agreement and contiguity in sequences in conversation. In *Talk and Social Organisation*, Graham Button and John R. E. Lee (eds). Multilingual Matters LTD, Clevedon, England, 54 – 69.

28. Harvey Sacks. 1992. *Lectures on Conversation*. Blackwell.

29. Emanuel A. Schegloff. 2007. *Sequence Organization in Interaction. A Primer in Conversation Analysis*. Cambridge University Press.

30. Margret Selting et al. 2009. Gesprächsanalytisches Transkriptionssystem. In *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 2009, 10: 353 – 402.

31. Jack Sidnell. 2013. Basic Conversation Analytic Methods. In *The Handbook of Conversation Analysis*, Jack Sidnell and Tanya Stivers (eds.). Wiley-Blackwell, 77 – 99.

32. Jack Sidnell and Tanya Stivers (eds.). 2013. *The Handbook of Conversation Analysis*. Wiley-Blackwell.

33. Melisa Stevanovic and Anssi Peräkylä. 2012. Deontic Authority in Interaction: The Right to Announce, Propose, and Decide. *Research on Language and Social Interaction* 45, 3: 297 – 321.

34. Tanya Stivers and Jack Sidnell. 2016. Proposals for Activity Collaboration. *Research on Language and Social Interaction* 49, 2: 148 – 166.

35. Luise Süssenbach, Nina Riether, Sebastian Schneider, Ingmar Berger, Franz Kummert, Ingo Lütkebohle, and Karola Pitsch. 2014. A robot as fitness companion: towards an interactive action-based motivation model. In *Proceedings of RO-MAN 2014*, 286 – 293.

36. Herwin van Welbergen, Ramin Yaghoubzadeh, and Stefan Kopp. AsapRealizer2.0: The Next Steps in Fluent Behavior Realization for ECAs. In *Proceedings of IVA 2014*, 449 – 462.

37. Ray Wilkinson. 2015. Conversation and aphasia: advances in analysis and intervention. *Aphasiology*, 29, 3: 257 – 268.

38. Yorick Wilks, Jan Jasiewicz, Roberta Catizone, Lucian Galescu, Kristina Martinez, and Deborah Rugs. 2014. CALONIS: An artificial companion within a smart home for the care of cognitively impaired patients. In *International Conference on Smart Homes and Health Telematics*, 255 – 260.

39. B. Winblad et alii. 2004. Mild cognitive impairment - beyond controversies, toward a consensus: report of the International Working Group on Mild Cognitive Impairment. *Journal of Internal Medicine* 2004, 256: 240 – 246.

40. Ramin Yaghoubzadeh, Marcel Kramer, Karola Pitsch, and Stefan Kopp. 2013. Virtual agents as daily assistants for elderly or cognitively impaired people. In *Proceedings of IVA 2013*, 79 – 91.

41. Ramin Yaghoubzadeh, Hendrik Buschmeier, and Stefan Kopp. 2015. Socially Cooperative Behavior for Artificial Companions for Elderly and Cognitively Impaired People. In *Proceedings of the 1st International Symposium on Companion-Technology ISCT 2015*, 15 – 19.

42. Ramin Yaghoubzadeh and Stefan Kopp. 2016. flexdiam – Flexible Dialogue Management for Incremental Interaction with Virtual Agents. In *Proceedings of IVA 2016*, 505 – 508.

43. Yoshinobu Yamamoto, Mitsuru Sato, Kazuo Hiraki, Nobuyuki Yamasaki, and Yuichiro Anzai. 1992. A Request of the Robot: An Experiment with the Human-Robot Interactive System HuRIS. In *IEEE International Workshop on Robot and Human Communication*, 204 – 209.

44. Keiichi Yamazaki, Michie Kawashima, Yoshinori Kuni, Naonori Akiya, Matthew Burdelski, Akiko Yamazaki and Hideaki Kuzuoka. 2007. Prior-to-request and request behaviors within elderly day care: Implications for developing service robots for use in multiparty settings. In *ECSW'07: Proceedings of the Tenth European Conference on Computer Supported Cooperative Work*, 61 – 78.