DISCOURSE EFFECTS ON THE PROSODIC PROPERTIES OF REPETITIONS IN HUMAN-COMPUTER INTERACTION

Kerstin Fischer

University of Hamburg

ABSTRACT

Repetitions may occur in human-computer interaction for various reasons; in this paper the constraints on the use of repetitions and their prosodic realization in the communication with a (simulated) automatic speech processing system which is not functioning properly will be analysed. It will be shown that repeats may have certain phonetic and prosodic properties which the respective original utterances do not necessarily display; however, besides these local changes depending on the immediate sequential context, the use of linguistic strategies such as repetitions changes globally throughout the dialogue. Thus, both the occurrence of repeated utterances and their prosodic realization depend on the relationship to global properties of the discourse structure which is partially determined by changes in the speakers' attitude towards the system.[1]

1. INTRODUCTION

Automatic speech processing systems do not always work as they should. Irritations caused by the system may lead to reactions by the speakers which are emotional or which are meant to increase the understanding for the system but which are actually even more difficult to process for automatic speech processing systems; recent studies (Levow (1998), Huber et al. (1998)) indicate that, because of increasing recognition error rates, research is necessary which deals with the linguistic features of repetitions and corrections uttered by dissatisfied or even angry users. In particular, there are discourse effects on the prosodic properties of such utterances in German human-computer dialogues, which include the occurrence of repetitions with respect to the global discourse structure on the one hand and the phonetic and prosodic realization of the repeated utterances on the other.

2. DATA AND METHOD

In order to determine how speakers react if the system repeatedly misinterprets their utterances, Wizard-of-Oz dialogues have been recorded, i.e. dialogues in which speakers believe to be talking to a machine while the system output is actually manipulated by a human 'wizard' (see Fraser and Gilbert (1992)). In these experiments, different kinds of system malfunction are being simulated. As a methodology for controlling inter- and intrapersonal variation, a fixed dialogue schema has been created which determines the utterances made by the system; thus certain sequences of system output have been defined which are combined in a fixed order, all phases appearing at least twice. These recursively recurring dialog phases make it possible to analyse the same sequences of utterances in different phases of the dialogue. The system output is thereby completely independent of the users' utterances. For instance, the system may ask the user to make a proposal for a day when to meet. Irrespective of the contents of the user's reaction, the system will then utter that the first of January is a holiday, simulating a speech recognition error. After the next speaker utterance, the system will assert that it is impossible to meet at four o'clock in the morning. This sequence may occur four times in each dialogue. The speakers' reactions to these utterances may change during time.

Speakers are instructed to schedule ten appointments with the system. Each of the recordings is ended by a sequence of system output 'I did not understand' and is then interrupted by the experimenter with the comment that the machine is obviously frozen. The speakers are then asked to answer some questions about their satisfaction with the system, whether they believe to have been emotionally engaged and whether they have believed to be talking to a computer. Afterwards they are informed about the real purpose of the recording.

The data are currently 26 dialogues of between 18 to 33 minutes length. They consist of 248 turns on the average,[2] 124 of which are uttered by the speaker. The relationship between female and male speakers is about equal; participants are between 20 and 38 years old and all native speakers of German. None of them has reported that s/he has realized that the system output was created by a human 'wizard', but all of them stated that they have reacted emotionally during the recordings (mostly angrily, some reported to have been amuzed).

The dialogues have been annotated regarding the macro-structural position of each turn, such that for each turn the position with respect to the dialogue phase is immediately identifiable. Thus, e002, for instance, is the identification of the speaker, while the other four digits represent the position of the respective utterance in the macro-structure of the dialogue; thus the first digit describes the subdialogue which is marked by the system's acceptance of a date suggested by the speaker. Altogether there

[1] The research for this paper was supported by the German Federal Ministry of Education, Science, Research and Technology (BMBF) in the framework of the Verbmobil project under grant number 01 IV 701 F7. The responsibility for the contents lies with the author.

[2] The variation, in spite of the fixed dialogue schema, is due to fact that speakers sometimes wait until "the system has recovered", and thus the 'wizard' has to initiate a nonscheduled request to propose a date in order to continue the dialogue.

are eight subdialogues while the speakers' task is to schedule ten appointments with the system, which means that it is impossible for them to fulfil the task. The next digit refers to a particular sequence of system reactions which is repeated several times in the dialogue. The last two digits count the utterances within such a phase. Table 1 shows an extract from the fixed schema which determines the order of (simulated) system output; phase 2101-2103, for instance, which is also instantiated in example (1), occurs five times per dialogue, one further occurrence being 3101-3103, only the phrasing having been altered slightly. Likewise, the sequence of utterances by the system in phase 2301-2308 occurs four times per dialogue. The speakers' reactions to these sequences can therefore be compared over time.

In addition to the macro-structural annotation shown in Table 1, fourteen of the dialogues have already been annotated regarding lexical, conversational, and prosodic peculiarities. Thus it is possible to retrieve automatically turns which constitute, for instance, a metalinguistic utterance or a repetition and which display phonetic features such as hyperarticulation, syllable lengthening, or distortion by laughter.

Analysing the differences between the same sequences in different structural positions of the dialogue allows to determine those differences which are created by different attitudes towards the system such as increasing dissatifaction. Thus if the speaker reacts differently although the situation is constant and the sequences of system output are actually the same, the differences can be attributed to a change in attitude towards the system and to emotional invlovement.

The methodology employed, however, allows the comparison not only of the prosodic, conversational and lexical properties of pairs (or triplets etc.) of utterances from different dialogue phases, but also of interpersonal variation. Thus, different styles of reacting to the system's malfunctions may appear. Using this corpus therefore allows to compare the linguistic behaviour of different speakers and to detect general tendencies.

3. REPETITIONS IN HUMAN-COMPUTER INTERACTION

The domain on which the influence of discourse categories on the prosodic properties of speech will be exemplified are repetitions. Repetitions may occur in all kinds of sequential contexts, for instance, after statements of failed understanding or after incomprehensible utterances by the system, and after utterances which show that the system has misunderstood. These sequential contexts occur repeatedly in the dialogues. Altogether, there are three aspects of the discourse structure which influence the occurrence of repeats in the dialogues under consideration:

First of all, while in human-to-human communication speakers generally acknowledge their understanding of their partner's utterances by means of display of continued attention, initiation of relevant next contribution, acknowledgment, etc. (Clark and Schaefer (1989: 267)), speakers do this to varying degrees in human-computer interaction. In the current dialogues, repeats are conditionally relevant when the system claims not to have understood. Furthermore, it may happen that the system asks the speaker to make a proposal even though the speaker has just made one. In such contexts, it is not only relevant but also preferred to meet the request by making a proposal (Levinson (1983: 336)), and the easiest way to do so is to repeat one's previous proposal. However, repetitions may also occur in positions in which they are not directly conditionally relevant, for instance, when the system utters something which is not understandable, or when it says something which is unrelated to the speaker's previous utterance, that is, in the context of simulated recognition errors.

Consequently, depending on the type of utterance which precedes a repeat, it is more or less conditionally relevant and it thus may indicate differing degrees of awareness to 'normal' sequential organisation. Thus, repeats can be classified according to their relevance in the sequential structure of the dialogues. In example (1), the speaker is confronted with two utterances of nonsense by the system. The first time she initiates a clarification sequence. The second time she repeats her utterance; this proposal is conditionally not relevant. After this sequence, the system requests the speaker to propose something and the speaker repeats her utterance; this time her proposal can be regarded as conditionally relevant. We can therefore distinguish two types of repetitions, conditionally relevant and conditionally not relevant, with respect to their position in the sequential structure.

(1) s0222101: bla irgendwas Termin wotsoewer?
e0222101: nochmal. (*again*)
s0222102: was soll Termin blablurb wot sehell bla?
e0222102: Donnerstag, einundzwanzigster erster, zehn bis sechzehn Uhr. (*Thursday, 21rst of January, 10 am to 4 pm*)
s0222103: bitte machen Sie einen Vorschlag. (*please make a proposal)*
e0222103: DONnerstag, einundzwanzigster erster, zehn bis sechzehn Uhr. (*Thursday, 21rst of January, 10 am to 4 pm*)[3]

As a second influence on the occurrence of repetitions in the dialogues under consideration, it has to be taken into account that although the utterances by the system and their order of occurrence are identical in all dialogues, the speakers react to the system output very differently; regarding repetitions, there are speakers (e004, e018) who use less than five repetitions throughout the dialogue while others use more than fifty (e016, e022). The average number of repetitions per dialogue is 29.7, of which an average of 14.4 are conditionally relevant.

As Figure 1 shows, speakers can be distinguished according to which communicative strategy they prefer: There are some who prefer to reformulate and to use metalanguage, and there are others who repeat their utterances very often instead. There is a negative correlation of -0.6 between the use of reformulations and metalanguage on the one hand and repetitions on the other. Thus, the number of repetitions occurring is dependent on who is speaking, and repetitions have to be seen as only one out of a number of strategies possible which are in linguistic opposition to the use of repeats in these dialogues.

[3]Capital letters are used to show when speakers realize a syllable with exceptionally strong emphasis. The convention <B> symbolizes a breathing event, <P> a pause, <L> syllable lengthening, and for those cases for which their phonetic realization is under discussion, the phonetic representation is given in square brackets according to the conventions developed by Wells et al. (1992).

Table 1: Part of the Fixed Dialogue Scheme

	dialogue act	actual utterance
...	...	...
2101	nonsense	bla irgendwas Termin wotsoewer?
2102	nonsense	was soll Termin blablurb wot sehell bla.
2103	request proposal	bitte machen Sie einen Vorschlag. '*please make a proposal*'
2201	reject proposal	dieser Termin ist schon belegt. '*this date is already occupied*'
2202	misunderstanding	die Urlaubszeit ist vom zehnten sechsten bis fünfzehnten siebten. '*vacation time is from 10.6.-15.7.*'
2203	request proposal	bitte machen Sie einen Vorschlag. '*please make a proposal*'
2301	misunderstanding	der erste erste ist ein Feiertag. '*the first of January is a holiday*'
2302	failed understanding	ich habe Sie nicht verstanden. '*I did not understand*'
2303	misunderstanding	der Wochenendtermin ist schon belegt. '*the weekend is already occupied*'
2304	misunderstanding	ein Termin um vier Uhr morgens ist nicht möglich. '*an appointment at 4am is not possible*'
2305	reject date	dieser Termin ist schon belegt. '*this date is already occupied*'
2306	misunderstanding	Freitag, der achte erste paßt sehr gut. '*Friday 8.1. suits me well*'
2307	misunderstanding	die erste Märzwoche ist schon belegt. '*the first week in March is already occupied*'
2308	accept proposal	ich habe den Termin für Sie notiert. '*I have noted the appointment*'
3101	nonsense	bla irgendwas Termin wotsoewer?
3102	nonsense	bla rapartibla Termin blurb.
3103	request proposal	bitte machen Sie einen Vorschlag. '*please make a proposal*'
...	...	...

A third influence on the occurrence of repetitions that needs to be considered is that the likelihood that the speaker repeats her utterance, rather than, say, reformulates it, increases during the dialogues, due to changing speaker attitude. Thus, when we consider the reactions to a particular utterance, for example, the system's statement that holidays will be in June and July (when the task is to find a date for an appointment in January), the likelihood that a speaker reacts by means of a repetition increases from 14% when this utterance occurs for the first time to 43% when it is uttered a third time towards the end of the dialogue. Likewise, if the system produces a sequence of incomprehensible utterances, the likeliness that the speakers will only repeat their utterances is five times higher when it occurs for the fifth time than when speakers are confronted with it for the first or even the second time. While in early phases of the dialogues speakers react directly to the system's output, that is, acknowledging what has been said and reacting relevantly, for instance, by means of reformulations and meta-communicational statements, they cease to try out different conversational strategies when they are more frustrated. Thus, the use of repeats after system utterances other than explicit statements of the failure to understand is a characteristic feature of later phases of the dialogue when the speakers are already emotionally engaged.

The following example shows how speakers' strategies may change in contexts in which the system states complete recognition failures. While in turn e0112302 speaker e011 reformulates her utterance to make it more precise, she makes the proposal more explicit in utterance e0114302. In turn e0115201, several dialogue phases later, her conversational strategy consists in repeating her original utterance with the inclusion of pauses between the individual phrasal constituents of her utterance. Finally, in turn e0118207, she does not even attempt to repeat her utterance; while her original utterance is realized by means of a number of particular prosodic properties such as syllable lengthening and pausing between the phrasal constituents, her strategy after the utterance by the system that it has not understood is to propose a new date. Thus, the occurrence of repetitions has to be seen as a strategy which speakers may use to solve particular problems arising in the interaction with the computer and which is more likely to occur towards the end of the dialogues rather than at the beginning.

(2) e0112301: <Laugh> <P> ich hätte gerne ein' Termin im Januar. <P> paßt es Ihnen am Dienstag, dem 19.1., von 8 bis 14 Uhr? (*I'd like to have an appointment in January. <P> Does it suit you Tuesday the 19th from 8am to 2pm?*)

s0112302: ich habe Sie nicht verstanden. (*I did not understand*)

e0112302: ich hätte gerne einen Termin am Dienstag, den 19.1., <P> von 8 bis 14 Uhr. (*I'd like to have an appointment on Tuesday the 19th <P> from 8am to 2pm.*)

e0114301: <Swallow> Montag, 18.1., von 8 bis 12 Uhr. (*Monday 18th from 8 to 12 am*)

s0114302: ich habe Sie nicht verstanden. (*I did not understand*)

e0114302: ich hätte gerne ein' Termin, am Montag, dem 18.1., von 8 bis 12 Uhr. (*I'd like to have an appointment on Monday the 18th from 8 to 12 am*)

e0115103: <B> Freitag, der 22.1., von 8 bis 12 Uhr. (*Friday the 22nd from 8 to 12 am*)

s0115201: ich habe Sie nicht verstanden. (*I did not understand*)

e0115201: <B> Freitag, der 22.1., <P> von 8 bis 12 Uhr. (*Friday the 22nd <P> from 8 to 12am*)

e0118206: am Mo<L>ntag, <P> dem 4.1., <P> von 12 bis 14 Uhr. (*on Mo<L>nday <P> the 4th <P> from 12 to 2pm*)

s0118207: ich habe Sie nicht verstanden. (*I did not understand*)

e0118207: <B> am Dienstag, dem 5.1., von 12 bis 14 Uhr. (*on Tuesday the 5th from 12 to 2pm*)

Especially in later phases of the dialogues, speakers may repeat their utterances irrespective of the speech act uttered by the system. A statistical analysis shows that 35% of all repetitions occur in the last 34 turns. Consequently, the occurrence of conversational strategies such as repetitions has to be interpreted in the context of the macro-organisation of the dialogue.

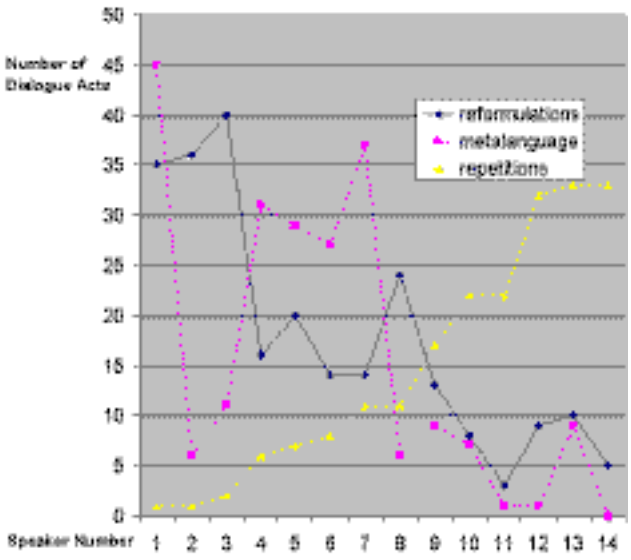

Figure 1: Number of Reformulations, Metalanguage, and Repetitions by Speaker

To sum up, the use of repetitions has to be seen in relation to other linguistic strategies, such as metalanguaging and reformulating, for which furthermore preferences by different speakers may be different. The use of repetitions is also dependent on the speakers' attitude towards their communication partner which may change during the interaction. Finally, repetition is not like repetition: Some are conditionally more relevant than others; therefore it may be concluded that different amounts of attention are being devoted to the sequential organisation of the dialogues, depending on the speakers' attitude as a result from the ongoing interaction with the system.

4. THE PROSODIC PROPERTIES OF REPEATS

The different types of repetitions in their various sequential contexts can now be investigated according to their prosodic properties. The dimensions along which their prosodic features can be described include the following:

- the variation of stress patterns;
- the inclusion of pauses;
- the variation of speed;
- the variation of loudness;
- the variation of intonation contours;
- the duration of syllables and of particular consonants;
- the inclusion of audible breathing and laughter.

The following examples show repeats in different structural contexts where phonetic and prosodic properties of the repeated utterance differ from those in the original utterance:

(3) e0021104: ich meine MONtag, den elften ERSten neunzehnhundertneunundneunzig [tsIC]. (*I mean Monday the eleventh of January, 1999.*)
s0021105: ich habe Sie nicht verstanden. (*I did not understand*)
e0021105: ich meine MonTAG, den elfTEN ersTEN neunzehnhundertneunundneunzig [tsik]. (*I mean Monday the eleventh of January, 1999.*)

In the original utterance, the speaker stresses the major first syllables of the content words *Montag, elften* and *ersten.* The weak syllables of *elften* and *ersten* are reduced to nasal alveolars, which is the standard pronunciation for these syllables. Also according to German standard pronunciation, the final consonant in *-zig* is realized by a palatal fricative. In contrast, in the repetition the non-prominent syllables *-tag* of *Montag* and the *-en*-syllables of *elften* and *ersten* are emphasized. The latter are not reduced but realized by /-En/. The *-zig* ending is realized by means of a /tsik/ in the repetition. Thus, we find variation of emphasis and instances of hyperarticulation in the example.

Likewise, in the repetition of the next example, the speaker (hyper-) articulates the formerly reduced syllable *-ter* of *vierunzwanzigster*, whose pronunciation is /st6/ in the standard pronunciation, as /stER/:

(4) e0027301: mein VorSCHLAG, SONNtag, vierundzwanzigster [st6] JANuar [janu'a:6], neunzehnhundertneunundneunzig. (*my proposal, Sunday, 24th of January, 1999.*)
s0027302: ich habe Sie nicht verstanden. (*I did not understand.*)
e0027302: <B> Sonntag, VIERundzwanzigster [stEr] JAnuar ["ja:nu?a:r], neunzehnhundertneunundneunzig. (*Sunday, 24th of January, 1999.*)

Furthermore, in the repetition a glottal stop is inserted between the two vowels of *Januar*, and the length of *Sonntag* increases by 50%.

In example (5), the repetition is more than a second longer than the original utterance, an increase of 17%. Furthermore, the duration of individual consonants, as for instance the duration of /S/ in *sechsstündig* which is more than doubled in duration from 120msec to 250msec in the repetition, is increased as well:

(5) e0198210: rei<L>cht nicht, wir suchen einen sechsstündigen Termin. DIENstag, neunzehnte erster, ACHT bis vierzehn Uhr. können Sie da? (*not enough, we are looking for an appointment of six hours. Tuesday, 19th, 8am to 2pm. does that suit you?*)
s0198301: ich habe Sie nicht verstanden. (*I did not understand*)
e0198301: einen SECHSs:tündigen Termin. DIENstag, NEUNzehnte erster, ACHT bis vierzehn Uhr. KÖNnen Sie da?

(*an appointment of six hours. Tuesday, 19th, 8am to 2pm. does that suit you?*)

Duration is thus one of the linguistic properties speakers attend to in the realization of repetitions.

In the previous examples, the utterances by the speakers were rejected by the speaker as not understandable on the whole. In the following example, the system utters something which indicates that it has only misunderstood aspects of the proposal by the speaker. Thus in example (6), the speaker stresses the name of the day and the date very much in the repeat, i.e. those words which carry the main informational load and of which she believes that they had been 'misunderstood' the first time. Thus, the speaker has hypotheses about what may have gone wrong and she tries to react accordingly, that is, she attempts to increase the understandability of those words which had not been previously understood.

(6) e0118204: am MONtag, dem VIERten ersten, von ZWÖLF Uhr bis vierzehn Uhr. (*on Monday, the 4th of January, from 12am to 2pm.*)

s0118205: Mittwoch, der sechste erste, von acht bis zehn Uhr ist schon belegt. (*Wednesday, the 6th of January, from 8 to 10 am is already occupied.*)

e0118205: <B> <:<very loud> AM MONtag:>, dem VIERten ERSten, von ZWÖLF bis VIERzehn UHR. (*on Monday, the 4th of January, from 12am to 2pm.*)

s0118206: Donnerstag, von acht bis zehn Uhr ist schon belegt. (*Thursday, from 8 to 10am is already occupied.*)

e0118206: am Mo<L>nta<L>g, <P> dem VIERten ERSten, <P> von ZWÖLF bis VIERzehn UHR. (*on Mo<L>nda<L>y, <P> the 4th of January, <P> from 12am to 2pm.*)

Regarding prosodic peculiarities, this sequence is an example for how a speaker uses loudness to increase understandability. Besides the varying loudness, in this example, the length of *Montag* ('Monday') increases from .95 msec to 1.15 sec to 1.3 sec in the second repetition. Furthermore, the stress pattern changes between the first proposal of this date and the first repetition. Finally, pauses are introduced between each phrasal constituent in the second repetition.

The example also shows that also the 'original' utterances do not need to be *normal* utterances, but that they display, for instance, a strong emphasis on particular syllables. These properties are, however, enforced when speakers repeat their utterances.

In example (7), which occurs quite at the beginning of the dialogue, the repeated utterance, in particular the content word *Termin* (*appointment*), is distorted by laughter; furthermore, unlike in the previous utterances, emphasis is reduced, and the speaker utters an interjection at the beginning of the repetition which indicates that she is dissatisfied with the system's behaviour:

(7) e0052202: ich hätte gerne einen TerMIN <P> von ACHT bis zwölf UHR, am ZWEIundzwanzigsten JANuar. (*I'd like to meet <P> from 8 to 12am, on the 22nd of January.*)

s0052203: bitte machen Sie einen Vorschlag. (*please make a proposal*)

e0052203: hm <Lip Sound> ich hätte gerne ein' <:<Laugh> TerMIN:> am zweiundzwanzigsten Januar, von acht bis zwölf UHR. (*I'd like to meet on the 22nd of January from 8 to 12am.*)

To sum up, there is a large number of linguistic properties which especially repeats display and which may cause problems for automatic speech recognition.

However, the phonetic and prosodic realization of repetitions has also to be seen in the context of the development through the dialogue. Thus, the use of the strategies observed varies throughout the dialogues. For instance, hyperarticulation has been found to occur mainly in the later phases of the dialogues, as Figure 2 shows. More precisely, 70% of all instances of hyperarticulation occur in the second half of the dialogues, and 37% can be found within the last 34 speaker turns. Consequently, the prosodic properties of utterances have also to be regarded with respect to the macro-structure of the dialogues.

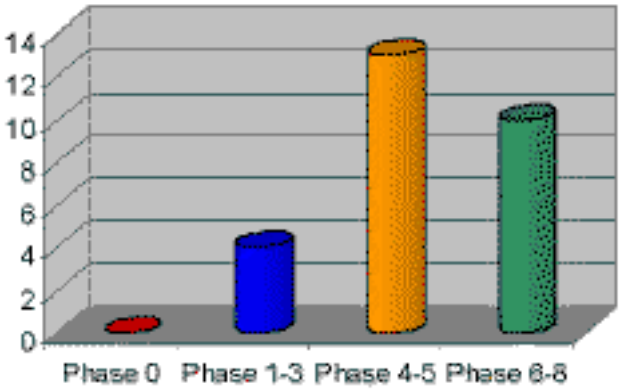

Figure 2: Distribution of Instances of Hyperarticulation in the Dialogues

Finally, it needs to be considered that there are repeats in which few or no peculiar linguistic properties occur, even in later phases of the dialogue. For instance, the only linguistic property speaker e012 employs in her first repetition is the inclusion of short pauses at different places in her utterance, and only in the second repetition she uses emphasis to increase understandability:

(8) e0128103: <B> einen sechsstündigen Termin, <P> am Sonntag, dem zehnten ersten, <P> von acht bis vierzehn Uhr. (*an appointment of six hours <P> on Sunday, the 10th <P> from 8am to 2pm.*)

s0128104: bitte machen Sie einen Vorschlag. (*please make a proposal*)

e0128104: <B> einen sechsstündigen Termin, <P> am Sonntag, dem zehnten ersten, von acht bis vierzehn Uhr. (*an appointment of six hours on Sunday, the 10th <P> from 8am to 2pm.*)

s0128201: ich habe Sie nicht verstanden. (*I did not understand*)

e0128201: <B> einen SECHSstündigen Termin, <P> am Sonntag, dem zehnten ersten, von ACHT <P> bis vierzehn Uhr. (*an appointment of six hours <P> on Sunday, the 10th from 8am <P> to 2pm.*)

Thus, like the occurrence of repetitions itself, the prosodic properties of repeats seem to depend on the same constraining factors: the changing speaker attitude, and individual differences. Although for the fourteen linguistically annotated dialogues statistically no significant complementary use of different prosodic strategies could be found, it can be shown that individual speakers differ very much regarding their use of these strategies. For instance, while speakers e006, e007, and e010 use many instances of strong emphasis (about 50) and very few instances of syllable lengthening (less than 10), the relationship is reversed for speakers e002 and e011. Therefore, at least individual preferences have to be taken into account in the analysis of the prosodic properties of repetitions in (simulated) human-computer interaction.

5. CONCLUSION

There are basically three results from the discussion in this paper; the first is a methodological one: The particular phonetic and prosodic properties of repetitions have been shown to be conditioned by the local sequential organisation as well as by global aspects of the discourse structure. Consequently, it does not make sense to study an isolated linguistic strategy such as repetitions without recourse to their role in the larger discourse structure and without respect to the other linguistic strategies to which they are in linguistic opposition.

Secondly, and more concretely, a list of phonetic and prosodic properties could be identified which characterize German repeated utterances in the communication with an artificial communication partner. These findings match on the whole with the acoustic properties of repeats found for English by Levow (1998).[4]

Thirdly, a number of factors which determine the occurrence of repetions and their particular prosodic properties have been determined. In particular, the occurrence of repetitions is determined by speakers' changing attitude towards their artificial communication partners. Furthermore, different styles in dealing with the problems of communicating with a malfunctioning automatic speech processing system could be identified. Finally, the occurrence and the role of repetitions in a dialogue also depend on the contributions by the system itself.

Likewise, the prosodic properties of repeats themselves have been found to be constrained by the same conditioning factors which influence the occurrence of repetitions: The employment of linguistic strategies such as stress placement, syllable lengthening, and hyperarticulation, which all cause problems for automatic speech recognition, has been found to be speaker dependent on the one hand and to be determined by the speakers' changing attitude towards the system on the other. The use of repetitions with their peculiar prosodic features therefore has not only to be seen as only one strategies among others to reach particular goals, it is also part of a complex network of relations between micro- and macro-structure of the respective dialogue, as well as individual speaking styles and attitude towards the communication partner.

On the whole, however, it seems clear that repetitions constitute a severe problem for the development of automatic speech processing systems (see also Fischer (1999)), and this paper presents only one step towards determining what aspects influence the linguistic behaviour of speakers in human-computer interaction.

[4] The variation of intonation contours, however, seems to be not just towards an increasing F0 range in repeats, as suggested for English by Levow (1998), since, for instance, in example (5), the speaker's F0 range even decreases in the repetition. However, no statistical analysis in this respect has so far been carried out on the German data under consideration.

6. REFERENCES

1. Clark, H.H. and Schaefer, E.F., "Contributing to Discourse" *Cognitive Science* 13, 1989, p. 259-294.
2. Fischer, K. "Repeats, Reformulations, and Emotional Speech: Evidence for the Design of Human-Compuer Speech Interfaces." *Proceedings of HCI International '99*, Munich, Germany, August 1999.
3. Fraser, N.M. and Gilbert, G.N. "Simulating Speech Systems." *Computer Speech and Language* 1991, p. 81-99.
4. Huber, R., Nöth, E., Batliner, A., Buckow, J., Warncke, V., and Niemann, H. "You BEEP Machine - Emotion in Automatic Speech Understanding Systems." In Petr Sojka, Václav Matousek, Karel Pala, and Ivan Kopecek, editors, *Proceedings of the First Workshop on Text, Speech, Dialogue-TSD'98*, Brno, Czech Republic, September 1998. Masaryk University Press, 1998, p. 223-228.
5. Levinson, S. *Pragmatics.* Cambridge University Press, Cambridge, 1983.
6. Levow, G.-A. "Characterizing and Recognizing Spoken Corrections in Human-Computer Dialogue." *Proceedings of Coling/ACL '98*, Montréal, Canada, 1998, p. 736-742.
7. Wells, J., Barry, W., Grice, M., Fourcin, A. and Gibbon, D. "Standard Computercompatible Transcription." Phonetics and Linguistics Department, UCL, *Esprit-Project 2589, SAM-UCL_037*, 1992.