MEDICAL INTERVIEWS QUANTIFIED: THE HARDER YOU BRAKE, THE LONGER IT TAKES: INTERRUPTIONS, STATUS, AND GENDER.

Prof. Dr. Florian Menz, M.A.
Associate Professor
Dept of Linguistics
University of Vienna, Austria
Berggasse 11
A-1090 Vienna
Austria
E-Mail: florian.menz@univie.ac.at

Dr. Ali- Al-Roubaie
Lecturer
Faculty of Psychology
University of Vienna, Austria
ali.al-roubaie@univie.ac.at

Abstract

Discourse research has repeatedly dealt with interruptions as indicators of power and dominance, especially in the context of gender and status. But only little research has doctor-patient-communication as a database. Out of a total of 576 medical interviews 48 were selected for a qualitative context-bound in-depth analysis of interruptions and subsequently evaluated statistically. Our findings support a differentiated interpretation: Regardless of gender, physicians show more non-supportive interruptions than patients (p=0.000); patients failed more frequently to interrupt physicians than vice versa (p=0.034), and even more with senior physicians than with doctors-in-training (p=0.003). Here, status seems to be the decisive variable. On the other hand, both female patients and doctors produced more supportive interruptions than male persons (p=0,013). Furthermore, contrary to intuition we could show that interviews take the longer the more physicians interrupt. The article closes with a critical discussion concerning the conjunction of qualitative and quantifying methods within discourse analysis.

Key words: gender, interruption, doctor-patient communication, power, status

MEDICAL INTERVIEWS QUANTIFIED: THE HARDER YOU BRAKE, THE LONGER IT TAKES: INTERRUPTIONS, STATUS, AND GENDER.

1. INTRODUCTION: INVESTIGATION-LEADING QUESTIONS

Our considerations are based upon a study carried out at a Viennese hospital's outpatient ward[1]. One of the main goals regarding the actual study was to audit to what extent doctors can be sensitized for certain annoying or conducive conversation styles by short trainings (cf Lalouschek, Menz & Gstettner, 2006). However, the following discussion focuses upon a related, though somewhat different question: how do interruptions influence the quality of medical interviews? Discourse research has repeatedly dealt with interruptions as indicators of power and dominance in interaction. In particular, this topic has been discussed - in part, quite controversially – by feminist linguistics when it comes to typical gender differences.[2]

With regard to the ongoing discussion, our investigations hence rely upon three batteries of questions concerning medical interviews: Firstly, do interruptions correlate - as spread in literature - with a person's social gender, namely within a doctor's and patient's group respectively?

Secondly, is there any relationship between interruptions and a doctor's level of training, i.e. does a professional status represent some influence within a doctor's group?

Thirdly, two conversational parameters will be compared: is there any dependence between frequency and type of interruptions and the length of the interviews? This question's background relates to an apprehension repeatedly expressed by doctors, namely that patients would take up too much time when not interrupted in their telling urge. In contrast to this

view, qualitative discourse-analytical findings indicate that interviews will definitely become complicated if patients were constrained in describing their concern (the chief complaint) coherently, i.e. the reason, why they call at a health care center.

While chapter 2 briefly describes the data material, upon which this investigation relies, chapter 3 summarizes the discussion regarding interruptions on which the development of the actual category grid is based. In chapter 4 results are presented and discussed. Chapter 5 summarizes our contribution, supplemented by a critical view concerning possibilities and limits of such analyses (chapter 6).

2. DATA MATERIAL

In a 2-phase period of 14 days each a total of 576 interviews have been recorded in an outpatients ward (March and June 2002 respectively). Thereof, 48 interviews were transcribed according to the GAT transcription system (Selting et al. 1998) and entered into the actual analysis. Significant screening factors were kind of illness (namely heart diseases) and the fact that the patients had contact both with some intern[3], as well as a senior physician[4].

The conflict between an outpatient clinic's productive and reproductive level seems to be reflected namely by the interns' status (Menz 1991): On the one hand, a patient should take center stage; on the other hand it is one of the educational tasks of interns to train and perfect an anamnesis dialogue. In other words, interns often find themselves in some sort of a pupil-teacher-relationship towards senior physicians and as such are consequently advised to provide diagnostic and therapeutic suggestions. Thus, an intern-driven anamnesis virtually
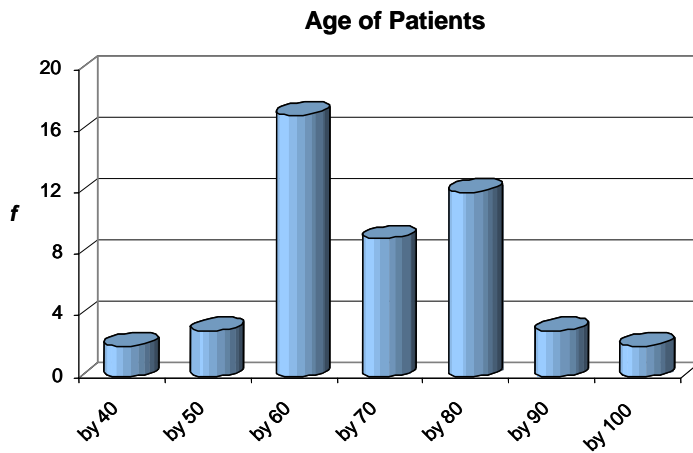
Age of Patients

Figure 1: Distribution of Patients Age

reflects a medical hierarchy - a weighty status difference to the decisive-entitled senior physicians, mostly opaque to the patients[5].

Our sample's attributes differ in terms of their distribution's homogeneity, depending on the specific parameters. For instance, women and men were almost equally represented when it comes to the medical staff (N=7).

Position	
senior physicians	3
interns	4

Gender	
Female	4
Male	3

Table 1: Position and gender of the medical staff: frequencies

While in medical interviews (N=48) patients gender was equally proportioned, significant divergences from population's age distribution were observed due to the symptom-based likelihood of cardiovascular disorders (cf fig. [1]): Accordingly, roughly 80% of the patients were between 50 and 80 years old at consultation.

Due to more methodical reasons the duration of a single medical interview was calculated by indirect estimation. For this purpose, we drew on the number of transcript lines per conversation. This measure seems to represent the real interview's duration more adequately

than a pure physical time line would do (e.g. minutes). Such a bias is mainly caused by the fact that every intern-driven anamnesis additionally includes a longer physical examination (as opposed to a senior physician's interview), usually carried out more or less quietly. Thus, any comparison of duration regarding possible conversation types based on professional levels would consequently fail. By grouping the line's frequencies into classes of 100 lines each, a central tendency of the distribution classes clearly appears around 300 to 400 transcript lines, supporting distribution's goodness of this parameter:

below 100	up to 200	up to 300	up to 400	up to 500	up to 600
1	6	17	13	8	3

Table 2: amount of transcript lines per frequency class

3. INTERRUPTIONS: DEVELOPMENT OF A CATEGORY GRID

Very early, interruptions have also been associated with social power and status, focused above all by gender research. A crucial starting point has been Lakoff's (1975) introspective assumptions and assertions. But the early work of West & Zimmermann (1975) also has to be mentioned, being quoted over and over again as an evidence for differences specific to gender-related interruptions. However, other studies (cf e.g. Samel's overview, ibid. 2000) countered that status, rather than gender differences would be responsible for interruptions. Thus, higher-status persons interrupt more often than lower-status persons. Besides, in a metastudy, James & Clarke (1993) evaluated more than 40 singular studies published between 1965 and 1991, where gender differences regarding the use of interruptions could not be confirmed, as most investigations were of little significance due to certain methodical flaws. In particular, the authors criticize diffuse operationalization, experimental conditions instead

of natural communication, as well as missing or faulty statistical computation. Hence, in our investigation these complained methodical criticisms were especially taken into consideration.

In the course of the turn-taking-discussion (Jefferson 1973), interruptions have come very early to the fore of discourse research; particularly conversation analysis. While earlier studies initially tended to distinguish between interruptions as some violation of the 'turn-taking' system and overlappings as "backfires" within right that system (Schegloff 1987, p. 85) - thus ensuring compliance with the „one speaker only at the same time"-rule by so-called 'return-taking repair mechanism' - this distinction became a questionable matter in subsequent works, consequently leading to a synonym application of both terms, meanwhile (e.g. Yieke 2002).

However, regarding the narrower field of medical communication, Holmes (1992) and West (1998) report that – regardless of professional status – female doctors were interrupted more often by patients regardless of their gender than their male colleagues. Hence, gender would be of stronger influence on communication patterns than status, contradicting the findings of James & Clarke (ibid.). But West's study (ibid.) likewise seems signified by a very limited number of examined conversations, the absence of readings regarding statistical significance and the purely formal operationalization of interruptions. On account of these contradictory results we evaluated doctors' interruptions as well as those done by patients. In addition, status was operationalized within doctor's group by degree of education.

One of the main criticisms of early investigations within the scope of the American conversation analysis has been the orientation towards a rather formal definition of interruptions as overlappings that happen two or more syllables (sometimes even words) before the end of an entire statement of the interaction partner (Jefferson 1973, Esposito

1979). In contrast, further investigations have shown that purely formal based investigations[6] of interruptions are not effective due to the polyfunctionality of such overlapping sequences. Consequently, a broader context has to be taken into account (Kotthoff 1993a, Tannen 1995). Menz et al. (2001) and Vodopiutz et al. (2001) for instance suggest that gender-specific differences can only be can be examined adequately in a comprehensive contextual analysis, encompassing conversation purposes and tasks. Goldberg (1990) arranged interruptions on a continuum between "relationship driven" and "power driven" poles and therefore suggested a functional, purpose-orientated interpretation frame. So did James & Clarke (1993), alternatively recommending a functional approach to the analysis of interruptions or overlappings that differs between cooperative and dominance-related speech acts. In these investigations the initial, purely formal distinction between overlappings and interruptions as it was still performed by Schegloff (1987) and West & Carpenter (1975) was overturned in favour of a more functional concept. Kunsmann (2000) takes the same line when arguing gradual crossings between interruptions, overlappings and listener's signals (minimal responses). However, in order to allocate adequate categorizations of interruptions, it seems advisable to take a more comprehensive conversational context into account.

Resting upon these critical considerations, we hence included the sociological variables *status* and *gender* in our investigation and categorized the linguistic variables according to their functional and contextual aspects, thereby taking into account the conversational context. To this end we adapted Yieke's (2002) systematic and applied overlappings and interruptions synonymously, classifying them according to a functional differentiation as supportive, non-supportive and neutral ones.[7] In our description such a distinction inherently makes sense, as a number of possible situations can be demonstrated, where interruptions do not represent a vio-

lation of the speaker's right to debate. This is for instance the case when clarifying inquiries, the ratification of explanations once they are comprehended (even if the explicating person possibly has not come to an end at all (Testa 1988)), the announcement of immediate danger („fire! ") or (minor and immediately corrigible) misunderstandings concerning a speaker's turn.

In addition to these types another two groups have to be distinguished, namely failed interruption attempts and minimal listener's signals (Kunsmann 2000), which are usually likewise expressed simultaneously, though should not be classified as interruptions due to their shortness.

1.1   Supportive interruptions[8]

According to Yieke (2002, 179f) this category is considered as a listener's statement primarily signalling interest and attention to the spoken (cf also Coates 1996), albeit such manifestations of interest do not necessarily have to be expressed simultaneously (Yieke, ibid.).

However, for the purpose of an unequivocal, narrowly defined operationalization, we categorized only those supportive statements, which were expressed simultaneously and borne by a cooperative and supportive intention. Other than Yieke[9] (who likewise takes listener's signals into account) we categorized these signals separately by virtue of their specificity (cf below). Hence, supportive interruptions can be of completing, clarifying or mending mode as demonstrated by the following examples:

Clarifying interruptions

```
Example 1: AM01 (D: doctor, P: patient)

    14   D:   Sie ham also [SCHMERZEN?
    15   P:               [aha\/ aha\/ (und da [vorne hat    ) mhm\/
    16   D:                                    [i- in=der BRUST

    17   D:   ja,=

    18   P:   <da HINTEN auf=da Rücken auf=da Schulter und da [(    )
    19   D:                                                   [mhm\/
```

In this example the relevant simultaneous speech sequence can be found in line 15/16. It is

clarifying insofar, as the patient's localization "in front" becomes more precisely specified by

the doctor („in the chest").

Completing interruptions

```
Example 2: AM01 (D: doctor, P: patient)
    50   D:   mhm\/ DORT in [diesem [RUNDEN
    51   P:                 [tz]    [RUNDE GROSSE. [(KAISERpalais da)
    52   D:                                       [ja, ja?  ja?
```

```
Example 3: DT01 (D: doctor, P: patient, N: nurse)

    401  N:   BITTEschön. (---) wie sind Sie hergekommen jetzt?

    402  P:   auch mit=m (.) [Roten Kreuz. ja.
    403  N:                  [Roten KREUZ?=nehmens draussen PLATZ?

    404  N:   [ich ruf sie Ihnen an.
    405  P:   [bitte.
```

Examples 2 and 3 illustrate how the interrupters complement and elaborate on the speaker's

statements.

Mending clarifications

```
Example 4: (P: patient, D: intern)

    64  D:   des (Plavix) nehmens scho wieda,=oda ab wann dürfens as
             wieda
    65       nehmen,
    66  P:   äh des is=es derzeit (-) STEPPED;=
    67  D:   =na=ge- (.) PAUSING is nur;
    68  P:   PAUSING; jojo.
```

It is quite characteristic for mending clarifications that the speaker's statements are corrected

in some detail, without implying a further change of turn. They rarely occur in our material.

1.2   Non-supportive interruptions

Yieke (2002, 180f) refers to overlappings as a violation of a speaker's rights and/or that are

dominance-related due to their non-supportive interruption style. In order to hedge such

statements, we defined and operationalized them likewise rather narrowly, as this category

implies dominance-related speech: simultaneous speech sequences accompanied by a sub-

ject's or addressee's change were categorized as non-supportive interruptions.

Non-supportive interruptions with subject change

```
Example 5: (P: patient, D: intern)

    64  P:   (--) muss i STEHN bleibn, (-) wenn i a bissl LÄNGA
    65       steh, das da=dann gehts wieda WEG?=a donn tuats
    66       obKLINGEN=net? .h (.) oda oda i nimm an SPRAY?=
    67  D:   =<mhm\/ <p>>=

    68  P:   =donn is [(AUCH etwas    ).
    69  D:            [mit=n MAGEN hams=NIE zu tun ghabt.

    70  P:   <bitte? <all>>
```

```
71  D:   mit=n MAGEN\/ hams=[NIE zu tun ghabt-
72  P:                     [NA=i  (.) gut i=man i nimm genug
```

In this example, the intern does not only interrupt the patient's statement at a moment when no turn change was intended, but also changes the subject jumping from breathing-related problems to a possible stomach trouble. The lack of understanding of the question by the patient (line 70) is another indication of the fact that an unexpected change has occurred.

Non-supportive interruptions with addressee's change

```
Example 6: (P: patient, D: doctor, N: nurse)

295  P:       [und da H sogt a:, sind Sie gsch=gscheit, Sie ham
296  P:   sovü scho durchgmocht,=und Sie ham sovü übastandn,=und
         Sie ham
297  P:   so ankämpft, (.) äh des [wern Sie a no;
298  D:                     [gut. (---) <mach ma an Termin
         aus

299  D:   Montag Aufnahme;=ja? <to nurse>>
300  P:   ((groans quietly))
301  N:   mhm\/
```

Other than in the preceding transcript excerpt, there is also a change in addressing a turn in this category: the doctor interrupts the patient and at the same time in his statement he turns to the nurse as an addressee.

1.3   Back Channels

Unlike Yieke (2002) we encoded back channels specially because they are, on the one hand, normally shorter than two syllables and hence do not fit the current definition of a

simultaneous speech. On the other hand, though, they can be documented by means of recording so-called back channel behavior (in the sense of active listening). In German, such listener's signals are normally expressed by „mhm∨" and „aha∨", accompanied by falling-rising intonation in each case.

1.4 Failed interruption attempts

Another category not covered in Yieke's pattern concerns failed interruption attempts. However, right these attempts do qualify very well as a proper criterion for asymmetrical conversational relations since they reflect a certain dominance divide in case somebody in vain attempts to attain the right of speech by interruptions in a conversation.

The schematical category grid looks as follows:

Physician- and/or patient-initiated interruption						
Supportive inter-ruption			Non-supportive interruption		Failed inter-ruption at-tempt	Listener's signal
KU	VU	RU	Topic change	Addressee's change		

Table 3: category grid. KU: clarifying interruption, VU: supplementing interruption, RU: mending interruption

As a matter of fact, supportive interruptions could not always be distinguished unambiguously from non-supportive ones. To meet a comprehensible and interpretable assignment for all cases, we therefore regarded the local course of conversation as a broader context of interest. In particular, operationalization via subject's or addressee's change turned out quite promising.

4. QUANTITATIVE RESULTS

1.5   Data structure

While empirical-explorative approaches - such as corpus analysis - can be seen as central components of quantitative linguistics, corresponding procedures in terms of some flanking supplement regarding discourse-analytic investigations are rather rare. Nonetheless, there are good reasons for a synthesis as reflected particularly by this specific investigation:

Questions as outlined at the beginning of this article suggest descriptive as well as inferential analysis of the codified data, given a qualitative or quantitative parameterization of initial variables can be done.

Last but not least, this approach gave us an opportunity to perform a comprehensive and statistically grounded prognosis regarding efficiency and goodness of medical interviews. According to our hypotheses, variables derived from both the characteristics of the involved persons (setting-related function, gender etc.) and the characteristics regarding communication (interruption, support, duration of interview etc.).

As aforementioned, empirical methods of testing and analysis are here to be taken merely as flanking procedures depicting a quantified map of the events, thus reflecting some specific methodically inherent parts of the research interest on the whole. Altogether, data are based on 48 different transcripts that were abstracted by means of a multistage compression into a set of 20 attributes (i.e. basic variables) for each case. Besides some more formal parameters, such as a transcript- and person-related flag, we coded person-specific attributes (variables) such as age, gender or status, as well as criteria characterizing a conversation's quality (failed

attempts of interruption, supportive interruptions, duration of anamnesis, etc.). Here, the terms "function" and "position" are meant to distinguish between medical positions on the one hand (senior physicians and specialists[10] versus interns) and doctors versus patients on the other hand.

We focused our interest upon the respective medical interviews whose measurable nature particularly resulted from the coactions of the involved persons and situational circumstances. Consequently, the statistically isolatable case corresponds to the conversational act which in turn was determined by a number of other characteristics (see above).

This implies, among other things, that the specific characteristics and patterns of the interviews were partially caused by the same persons. As a result, a priori partial and unsystematic dependencies within data had to be taken into account. In view of that, we therefore resorted to those statistically robust procedures whose methodical assumptions are distribution-free and/or show low requirements regarding the variables scale level.

One of the crucial questions was whether doctors and patients differ significantly in degree of interruptions. Hence, according to the distinction in supportive and non-supportive interruptions as well as failed interruption attempts, both groups were compared separately in each category.
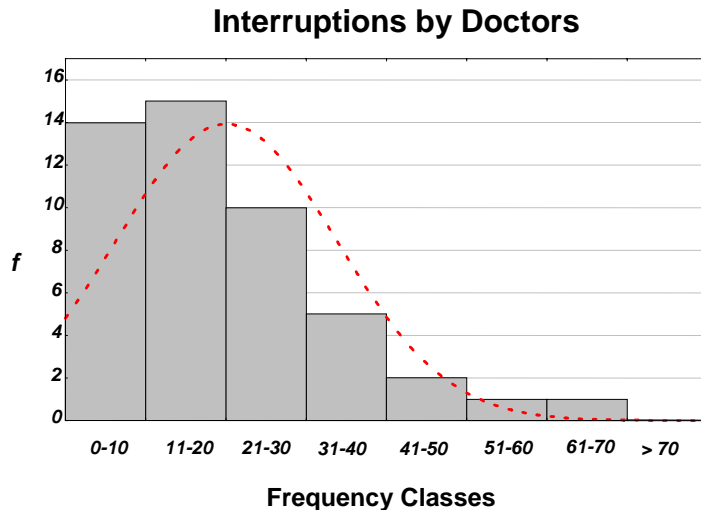
1.6   Evaluation

Interruptions by Doctors

Figure 2: Distribution of interruptions
initiated by doctors as frequency classes.
Vertical axis represents frequencies and the
following, faded in the suitable normal
distribution respected from the data (dotted

The distribution of interruptions effected by doctors turned out left skewed in case of both sexes (fig. [2]). Most interviews were interrupted about 20-30 times (Median: 17.5), however, up to 65 interruptions per interview were also recorded. Test statistics[11] (table 4) confirm a distributional skewness (1.39), showing a considerable divergence from

normal distribution regarding all three types of interruption:

K-S	D=.11412	P>.20
Lilliefors		P<.15
Shapiro-Wilk	W=.89002	P<.0001

Table 4: test statistics on
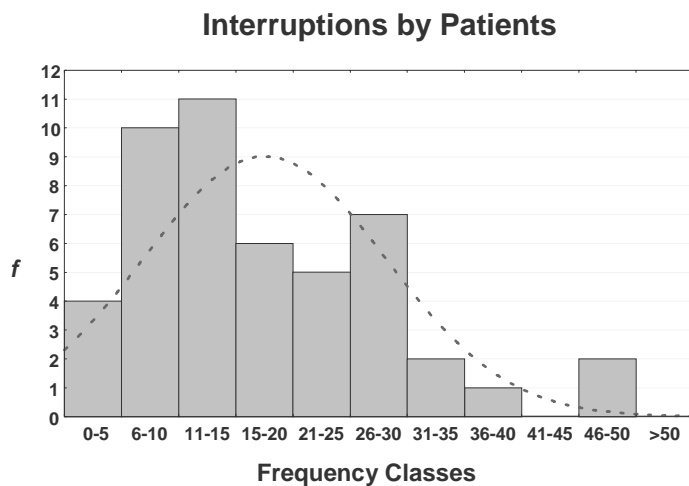normal distribution for the
interruptions initiated by
doctors[12]

Interruptions by Patients

The distribution of interruptions effected by patients is comparatively flat (Kurtosis 2.346: 0.523), showing a broader

Figure 3: Distribution of interruptions initiated
by patients.

distribution over the frequency classes - which is not further surprising in view of a more than fourfold number of participants (fig. [3]):

Interruptions by patients were non-normally distributed as well, so that ordinal-related and distribution-free methods respectively were applied in the following analyses (cf table 5):

K-S	D=.11942	P>.20
Lilliefors		P<.10
Shapiro-Wilk	W=.92319	P<.0042

Table 5: test statistics on normal distribution for interruptions initiated by patients.

1.7   Results

As an initial step we contrasted the total of doctors and patients in respect to possible differences in each category (supportive, non-supportive and failed interruptions). Do doctors and patients show different behavior patterns concerning their *supportive interruptions*? And if so, does this behavior considerably change during the course of the interview?

Since we were already aware of the low distributional goodness of fit from preceding tests, a variant of the mentioned Kolmogorov-
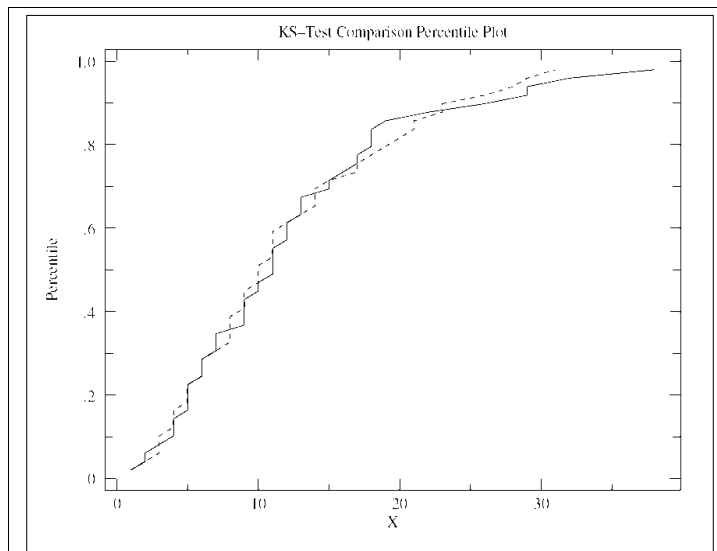
Figure 4: Cumulatative distribution of doctors' and patients' supportive interruptions

Smirnov test (this time comparing two samples) was applied. We opted for the K-S as the method of choice, as no prerequisites concerning any specific distributional quality (besides unimodality) are required. Test statistics are merely based on the difference of the cumulated frequencies of both groups. Figure [4] illustrates that there are barely any differences between both groups:

As can be seen, doctors and patients seem to show a similar type of communication when it comes to the category of supportive behavior. Both groups would respond equally often to their opponent and this result remains likewise significant once any "outliers"[13] are removed by means of the Tukey elimination (Tukey 1977) before analysis (cf, for instance fig. [5], upper right corner). Does this result equally apply to all interviews regardless of their duration or do we have to differentiate longer interviews concerning their supportive quality?

This question can roughly be answered by plotting the interviews – ranked by the amount of interruptions – on a pseudo time axis representing the amount of lines of the transcripts. To point out the tendencies in both groups, we superimposed the intervention frequencies by a linear trend broken down by doctors and patients.
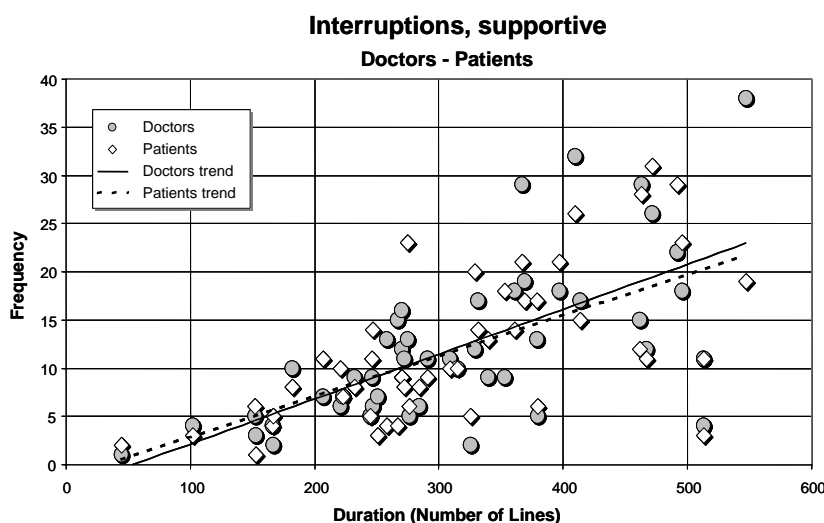

Figure [5] depicts two different references: on the one hand the number of supportive interruptions rises the longer the interview lasts; on the other hand there is no pronounced

Figure 5: Frequencies of positive interventions initiated by doctors and patients respectively. Duration of conversation vs. frequency (vertical axis)
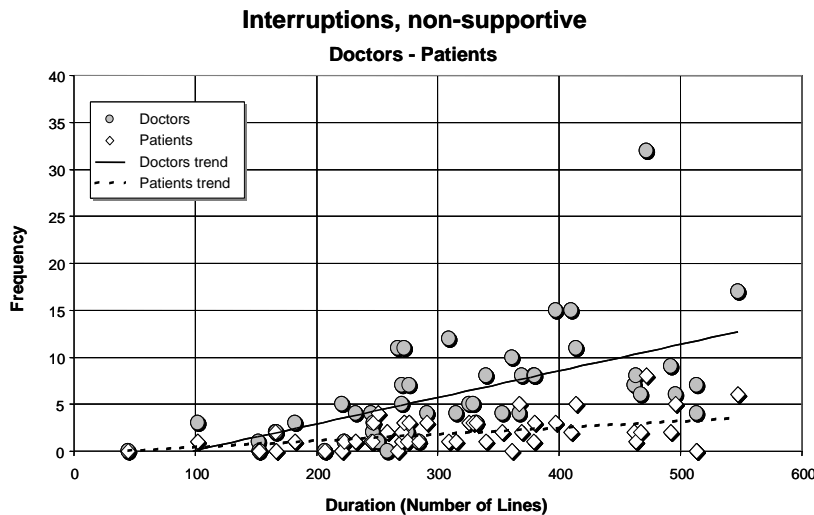
Interruptions, non-supportive

Doctors - Patients

Figure 6: Distribution of non-supportive interruptions initiated by doctors and patients.

reference to differently frequent interventions of one group. To a large extent, both group trend lines run quite identically.

Can this result be replicated regarding the two other interruption types (non-supportive, failed attempts)? Again, for each type we examined possible differences between the distributions via the Kolmogorov Smirnov test.

As expected, this time significant differences between the professional subgroups became evident, emerging further on even more clearly as discussion duration increases (fig. [6]). Now, first to the question of whether non-supportive interruptions might randomly be transacted by doctors and patients:

The answer seems clear-cut: doctors were demonstrably more frequently involved in non-supportive interruptions than were patients (p=0.000).

As indicated by figure [6], this phenomenon becomes more apparent the longer the discussion lasts. A dominance of doctors' non-supportive interruptions is already obvious starting from approximately 250 transcript lines. Corresponding interruptions transacted by patients accumulate between 200 and 400 transcript lines, whereby the number of non-supportive interruptions hardly exceeded five.
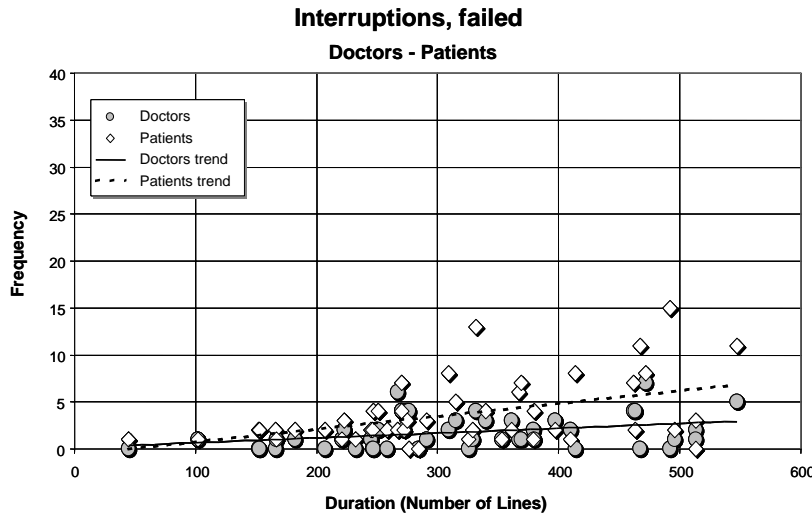
Figure 7: Distribution of faild interruptions initiated by doctors and patients respectively. Duration vs. frequency (vertical axis)

According to our objectives, the connection between duration of anamnesis and extent of respective interruptions was to be represented in the sense of a correlation. Again we had to consider the comparatively low distributional quality of the actual data.

Spearman's rank correlation coefficient[14] provides a reasonably robust estimation of the relationship between interruptions and duration (see Zimmerman /Williams 1997).

A noteworthy correlation was observed in the category of doctors non-supportive interruptions (see also fig. [6]). Doctors seem to interrupt patients the more frequent in a non-supportive way, the longer an anamnesis lasts ($R=0.702$). The reverse case arises however when we look at the failed interruption attempts: patients try to interrupt more frequently in vain than doctors do ($p=0.034$).

Again, the relevance of the discussion's duration seems evident (fig. [7]), although the effect is not that obvious due to the comparatively smaller volume, in particular since the vertical scaling (frequency of failed interruptions) conforms to the preceding diagrams (fig. [5,6]) in height.

Regarding the doctors different status (seniors vs. interns), we additionally examined whether they interrupt or were interrupted by patients respectively more frequently in the three categories mentioned above. Statistics revealed that patients significantly more often failed to interrupt seniors than interns (p=0.003). Status seems to have a certain influence, here.

However, the interactants' gender seemed pretty irrelevant except for one category: both female patients and doctors initiated more frequently supporting interruptions than did male persons (p=0.013); see also (fig. [8]).
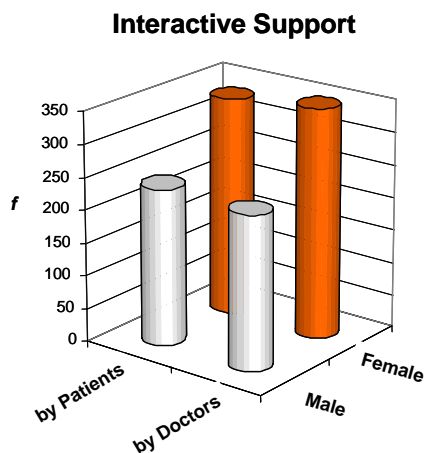
Interactive Support

Figure 8: Frequencies of supportive interruption: crosstabulated relationship between doctors

5. DISCUSSION

Our actual investigation comprises some quantitative issues deriving from a more comprehensive survey on medical conversation behavior and concerns several fields of research: first, results relate to investigations dealing with asymmetrical communication, in particular to those postulating interruptions and overlappings as indicators for dominance and/or status-oriented linguistic behavior. Beyond that, however, they also allude to questions dealing with the conjunction of qualitative and quantitative methods in discourse research (see also Gruber & Menz eds 2001). Further respective considerations will be presented in chapter 6.

The explanatory power of gender and status regarding different behavior of verbal interruptions has extensively been discussed in literature (see also chapter 3 and

aforementioned publications). The results of our analysis support additional evidence for a differentiated interpretation.

Distributions of non-supportive interruptions (i.e. interruptions in a narrower sense) bespoke among other things, that interruptions may very well indicate asymmetrical communication situations:

Physicians do so - regardless of their gender - far more frequently than do patients. The asymmetry of the linguistic type "anamnesis" has been examined likewise in detail regarding numerous other dimensions (see recent studies e.g. Menz & Lalouschek 2006 and Neises, Ditz & Spranz-Fogasy eds 2005), likewise indicating that interruptions are status-oriented. This view is furthermore backed by one of our actual results: patients fail more often in interrupting doctors, particularly when talking to physicians of higher status (specialists or seniors). Aside from a minor significance, however, our data do not reflect any statistically significant difference regarding both sexes. This suggests the conclusion that - regarding dominant linguistic behavior - status (or function) obviously shows stronger influence than gender. It seems, as if "doing gender" (sensu establishing asymmetry, cf Schoenthal 1998: 165ff) in physician-patient interaction rather does not take place by increased interruptions, but by means of other linguistic strategies.[15]

However, the necessity for a functional and contextual differentiation within the range of interruptions seems evident when we take a closer look at the following result: both female physicians and female patients show more supportive interruptions than do their male counterparts. As this category's definition implies, the fact that a speaker's interests were maintained despite interruption, results support other findings that assign women a more relationship-oriented linguistic style (see e.g. Tannen 1990 as well as her discussion in Talbot

1998). If interruptions only were formally defined as longer overlappings preceding situations with possible change of speaker (see section 3), such results would not be interpretable, but rather of distorting evidence, in case.

Finally, we resort to an almost paradoxical result, which could be of some interest particularly regarding communication trainings of physicians:

As mentioned above, we observed a significant correlation between the duration of the interview and the number of the non-supportive interruptions. Doctors tend to interrupt patients frequently before a contextual relation is recognizable. According to different studies, doctors would do so to stop the (alleged) loquacity of patients in time. Already more than 30 years ago, Rohde (1974) criticized this symptom as "preventive emergency behavior". Discourse-analytic views however stress that interruptions in medical interviews prevent patients primarily from stating their main grievance (the specified chief complaint). Consequently, patients attempt over and over again, thereby unwillingly complicating medical dialogue (Lalouschek 1995) and, as our results show, even extend them. To sum up: interruptions performed in order to shorten a discussion, will rather extend it! Frequent non-supportive interruptions thus cause not only qualitatively, but also quantitatively unwanted, yet paradoxical results.

6. CRITICAL VIEW: DISCOURSE RESEARCH AND QUANTITATIVE METHODS OF COMPLEX ORDER

Diverse variants of linguistic analysis usually reflect the acquisition and quantification of data by providing statistically descriptive characteristics. However, what if we examine more

complex hypotheses, probe assumptions or want to explore cross linked interconnections? Distributional examinations, significance tests of central tendency as well as analyses of variances yet limit the information obtained this way.

Regarding the actual study, a possible question might be for instance, whether abstract dimensions are conceivable "beyond" the given variables (interruptions, discussion duration etc.) thus allowing for grouping the involved variables or persons. Since these variables only represent intuitive or theory-imminent estimations based upon underlying linked structures, another resuming analysis would be of additional benefit.

This leads us straight to a cluster of so called "dimension analyses", whereof two typical approaches are worth mentioning. Any of them lead to a new, more abstract, combination of data. Hence, asking "which variables, cases, persons etc. are related more than others?", explorative dimension analyses might facilitate the answer every now and then. The requirements regarding the quality of data (distribution characteristic, scale level) are yet inconsistent. However, in any case the application of such methods is meaningful only given a sufficient data volume. Adding to a broader – more network driven - view, two other approaches originating from analysis methods in cognitive studies are briefly discussed in the sequel (chapter 6.4, 6.5).

1.8   Cluster Analysis

All kinds of cluster analyses attempt to identify homogeneous subsets within a random quantity of items. Thereby, the goal is to separate a certain number of objects into distinct groups (clusters), so that these objects resemble each other - if possible - within groups, whereas the differences between groups remain as large as possible.

By analogy, here the appropriate question therefore is, whether the anamneses (the cases) can be grouped meaningfully: Which of the interviews are more similar - e.g. regarding interruptions - than others? Are there different types of interview in this regard? Cluster analysis hence would allow re-evaluating conversations regarding their similarity and possible types of conversation in terms of a linguistic view.

1.9   Factor Analysis

A completely different view derives from the question of whether there are abstract dimensions, represented by variables. Chances are that dimensions like "discussion experience", "education" or "harmonization tendency" – just to name a few – contextually superpose raised variables to a certain extent. Factor analysis examines, whether some variables are more related than others. Thus, while cluster analysis looks for conversational types, factor analysis attempts to group variables into more comprehensive dimensions. As with cluster analysis (or any other statistical method), factor analysis does not provide any a-priori information regarding possible meaningfulness or prominence of the calculated factors, so rather some methodical and content-oriented foreknowledge of appropriate interpretation is demanded. Such approaches, however, may lead to possible new questions of interest that in return have to be explored by a qualitative driven discourse-analysis. Synoptically taken, such statistical research methods thus might not only be of some ex-post facility (in order to supplement qualitative analyses), but also ex-ante in the course of a mutual replenishing process for the generation of (so far unconsidered) hypotheses.

1.10   Time Line Analysis

The temporal changes in language and communication lead to the question, whether they are to be described and generalized formally by coincidental nature. Timeline analysis typically addresses two different aspects: prognoses in the context of some constant change (trend) as well as the discovery of periodic fluctuation (season). This attempt usually concerns regression analyses. Some simple variants are pictured in figures [4, 5, 6]: we wanted to know whether - during a certain conversation period - doctor and patient behavior differs regarding certain aspects of trend and thus being computed as a linear trend for both groups. Prognosticly seen, we therefore could expect that the trend differences would strengthen according longer conversations or, pointedly spoken, along more interruptions (than in the discussions raised).

1.11   Dynamic Models

A rudimentary way of quantifying the significance of topics is to assess the amount of textual references regarding a certain topic in a given corpus. In such a case it is often of interest to determine the extent and direction of transverse references: which topics refer to others, and to what extent? Quantitative content analysis usually reflects this question by tabulating static frequencies.

A completely different approach, however, arises from a probabilistic view when seeking for some (conditional) probability that interconnects topics and respective references in comparable corpora. This leads to the assumption that a certain topic may disclose *a-posteriori* probabilities for a further cluster of topics, the given frequencies of the respective references presupposed. In contrast to a more static analysis of frequencies, we thus deal with

a dynamic model. Such processing models - probably better known as belief systems - are a useful tool to reflect causal relations within a model based upon nominal or ordinal data. We preferably think it advisable in all those cases referring to textual and contextual categories, where representation as a simple quantity does not seem promising. Respective computations reveal (conditional) Bayesian probabilities of the discrete states of a model, so that targeted questions referring to interesting constellations can be addressed to the system. Additionally, the advantage of the adaptability comes, as new information improves the prediction accuracy. Furthermore, such hybrid text analyses might provide additional insight, particularly in conjunction with fuzzy logic and neural networks (Al-Roubaie 1998).

1.12   Network Analysis

Taking the corpus topics as narrative agents allows for another view, where right these topics represent the nodes of an n-dimensional network, connected by certain relations. Such relations (also noted as "edges" due to their origin from graph theory) represent the joint meaning of these topics within an entire net or a part of it (e.g. a subnet). One prominent goal of network analysis is to explore dis/similarity regarding the involved objects (or actors) and to point out the meaning of possible subgroups. In the consequence, such an approach additionally allows for identifying key objects (e.g. topics whose loss would lead to the isolation of entire subgroups) or to illustrate the dynamics of groupings ("which topical categories are more significant, more powerful than others"?).

Formally, networks equal matrixes, being condensed for the purpose of visualization by means of dimension-reducing procedures (e.g. multi-dimensional scaling, principal main component analysis, block clustering etc.) into a two-dimensional space. Depending upon

interest of research, single nodes or changes regarding the "power" of subgroups come to the fore. A variety of topics can serve as network nodes, such as organizations, opinions, dates or even grammars. Hence, surprisingly a many linguistically relevant aspects can be analyzed and interpreted as a network, as long as respective relations are kept identifiable. Again, an example from our investigation: let the medical personnel be the actors, being assigned certain attributes (e.g. status, position, gender). Accordingly, do they reflect certain networks; do interns – for instance – act as an identifiable subgroup reflecting homogeneous interests, such as the amount of interruptions or conversational duration? Which other subgroups are close to them, is there any cooperation? In summation, is there any particular effect caused, for example, by seniors and specialists, females or men? Which "key nodes" might be capable of weakening or even crashing the entire network? Such questions typically are addressed by network analyses, so again there is some further potential capability of answering one question or the other.

The long-term development of statistical methods and approaches regarding qualitative data in particular, as well as an increasing receptiveness and disposition for quantifying procedures by discourse research - as they are presented and applied in this article - once again raise hope of a mutual enrichment.

7. REFERENCES

Al-Roubaie, A. (1998). How to isolate fuzzy patterns: a hybrid approach in text analysis. *21st Biannual Conference of the Society for Multivariate Analysis in the Behavioural Sciences*. University of Leuven, Belgium. July 13–15, 1998

Coates, J. (1996). Women talk: conversation between women friends. Oxford: Blackwell.

Esposito, A. (1979). Sex differences in children's conversation. *Language and Speech* 22:213-220.

Fisher, S. (1993). Reflections on gender, power, and discourse. In A. Todd & S. Fisher (eds) *The Social Organisation of Doctor-Patient Communication* (287-299). Norwood, NJ: Ablex.

Goldberg, J. (1990). Interrupting the discourse on interruptions: An analysis in terms of relationally neutral, power and rapport oriented acts. *Journal of Pragmatics* 14: 883-904.

Gräßel, U. 1991. Sprachverhalten und Geschlecht. Eine empirische Studie zu geschlechtsspezifischem Sprachverhalten in Fernsehdiskussionen. Pfaffenweiler: Centaurus Verlagsgesellschaft (*Aktuelle Frauenforschung Band 13*).

Gruber, H. & Menz, F. (eds)(2001). Interdisziplinarität in der Angewandten Sprachwissenschaft. Methodenmenu oder Methodensalat? Bern: Lang.

Holmes, J. (1992). Women's talk in public contexts. *Discourse & Society* 3(2): 131-150.

James, D. & Clarke, S. (1993). Women, men, and interruptions: a critical review. In D. Tannen (ed) *Gender and Conversational Interaction* (231-280). Oxford: OUP.

Jefferson, G. (1973). A case of precision timing in ordinary conversation: Overlapped tag-positioned address terms in closing sequences. *Semiotica* 9: 47-96.

Kotthoff, H. (1993a). Kommunikative Stile, Asymmetrie und „Doing Gender". *Feministische Studien 11*: 79-95.

Kotthoff, H. (1993b). Unterbrechungen, Überlappungen und andere Interventionen. Vorschläge zur Kategorienunterscheidung und kontextorientierten Interpretation. *Deutsche Sprache 2*: 162-185

Kowal, S. et al. (1998). Unterbrechungen in Medieninterviews. Geschlechtstypisches Gesprächsverhalten? In G. Schoenthal (ed) 279-299.

Kunsmann, P. (2000): Gender, status and power in discourse behaviour of men and women. *Linguistik Online* 5/00 (http://www.linguistik-online.com/1_00/KUNSMANN.HTM).

Lakoff, R. (1975): *Language and Women's Place*. New York: Harper and Row.

Lalouschek, J. (1995). Ärztliche Gesprächsausbildung. Eine diskursanalytische Studie zu Formen des ärztlichen Gesprächs. Opladen: Westdeutscher Verlag.

Lalouschek, J., Menz, F. & Gstettner, A. (2006): *Ärztliche Gesprächsführung in der ambulanten medizinischen Versorgung*. Münster: LIT-Verlag.

Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 64, 399-402.

Menz, F. (1991). Der geheime Dialog. Institutionalisierte Verschleierungen in der Arzt-Patient-Kommunikation. Bern: Peter Lang.

Menz, F., Lalouschek, J., / Stöllberger, C. & Vodopiutz, J. (2002). Geschlechtsspezifische Unterschiede bei der Beschreibung von akutem Brustschmerz: Ergebnisse einer medizinisch-linguistischen transdisziplinären Studie. *Linguistische Berichte* 191: 343-366.

Neises, M., Ditz, S. & Spranz-Fogasy, Th. (eds): Patientenorientiert reden – Beiträge zur psychosomatischen Grundkompetenz in der Frauenheilkunde. Stuttgart: Wissenschaftliche Verlagsgesellschaft.

Rohde, J.J. (1974). *Soziologie des Krankenhauses.* Stuttgart: Enke.

Samel, I. (2000). Einführung in die feministische Sprachwissenschaft. Berlin: Schmidt.

Schegloff, E. (1987). Recycled turn beginnings. In G. Button & J. Lee (eds) *Talk and Social Organization* (70-85). Clevedon: Multilingual Matters.

Schoenthal, G. (ed). (1998). Feministische Linguistik - Linguistische Geschlechterforschung. Ergebnisse, Konsequenzen, Perspektiven. (*Germanistische Linguistik* 139-140). Hildesheim: Olms.

Schoenthal, G. (1998). Geschlechtstypisches Kommunikationsverhalten: Ergebnisse, Konsequenzen und Perspektiven. In G. Schoenthal (ed) 155-174.

Selting, M./Auer, P./Barden, B./Bergmann, J et al. (1998): Gesprächsanalytisches Transkriptionssystem (GAT). *Linguistische Berichte* 173: 91-122.

Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality, *Biometrika*, 591-611.

Shapiro, S. S., Wilk, M. B., & Chen, H. J. (1968). A comparative study of various tests of normality. *Journal of the American Statistical Association*, 63, 1343-1372.

Talbot, M. (1998). *Language and Gender: An Introduction*. London: Polity Press.

Tannen, D. (1990). *You just don't understand.* New York: William Morrow.

Tannen, D. (1995). Talking from 9-5: Women and Men in the Workplace. Language, Sex and Power. New York: Avon.

Testa, R. (1988). Interruptive strategies in English and Italian conversation: Smooth versus contrastive linguistic preferences. *Multilingua* /: 285-312.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.

Vodopiutz, J., Poller, S., Schneider, B., Lalouschek, J., Menz, F. & Stöllberger, C. (2002). Chest pain in hospitalized patients: Cause specific and gender-specific differences. *Journal of Women's health & gender based Medicine*. Vol. 11, No. 8: 1-9.

West, C. & Zimmerman, D. (1975). Sex roles, interruptions and silences in conversation. In B. Thorne & N. Henley (eds) *Language and Sex: Difference and Dominance* (105-129). Rowleyy, M.A: Newbury House.

West, C. (1998). When the doctor is a lady: Power, status and gender in physician-patient encounters. In J. Coates (ed) *Language and Gender: A Reader* (127-138). Oxford: Blackwell.

Yieke, Felicia A. (2002): Language and discrimination: A study of gender and discourse in workplaces in Kenya. University of Vienna (unpublished dissertation).

Zimmerman, D. W., & Williams, R. H. (1997). Properties of the Spearman correction for attenuation for normal and realistic non-normal distributions. *Applied Psychological Measurement*, 21, 253-270.

NOTES

1   The project "Institutional and discussion-structural influences on physicians communicative skills in an internal outpatient clinic" was promoted by the Felix Mandl-fund and the City of Vienna mayor's fund. Project management: Univ. Prof. Dr. Florian Menz (linguistics), Univ. Prof. Dr. med. Claudia Stoellberger (medical science); Project cooperation: Mag. Andreas Gstettner (linguistics), Univ. Lecturer Dr. Ali Al-Roubaie (statistics, methods).

2   A good overview is offered e.g. by Gräßel (1991), Talbot (1998), Schoenthal (ed) (1998) or Samel (2000).

3   According to the Austrian medical training scheme, interns (or first year residents) are referred as doctors-in-training who - after having successfully completed their studies – have to absolve their training as a general practitioner (3-4 years) or a specialist (6 years). This can be seen as a prerequisite to receive their *ius practicandi*.

4   Medical specialists, including seniors

5   Cf. results regarding orientation in Lalouschek, Menz, Gstettner (2006).

6   I.e., by the number of simultaneously spoken syllables. Kowal et al. (1998: 292) even mention the category "interruptions without simultaneous speech", preferably exerted by media professionals regarding "precision timing" (Jefferson 1973)

7   The actual investigation afforded a thorough analysis of the discussions by means of a given category raster. Definitions of the individual kinds of interruption in use had to be coordinated on the one hand with the field of investigation (physician-patient conversations in an outpatient clinic) and accordingly specifically distinguished from other models. On the other hand, definitions that did not differentiate that well were to be combined into meta-categories in order to create a manageable grid for statistical evaluation. Besides, the (statistically) small sample of conversations forced us to restrict the categories to a considerable number.

8   Sincere thanks to project collaborator Mag. Andreas Gstettner for categorization of statements.

9   Yieke (2002) subsumes mainly *back channels* under such signals. We still coded them seperately, as – from a functional and contextual point of view - they can be meant supportive or non-supportive (the latter was frequently encountered in case of impatience or an intentional take over of a turn). This actual example again illustrates the need of a context-sensitive manifestation's interpretation for a reliable and valid assignment.

10   In order to warrant statistical significance, senior physicians and specialists were merged into one single group. In so doing, an agreeable size of respective subsamples could be accomplished.

11   A clear knowledge of data characteristics (type of distribution, scale level etc.) is essential for choosing the appropriate method. Subsequent examinations test for possible deviations from a universal set or from another sample, resulting in so-called test statistics.

12   "K-S" here stands for a Kolmogorov-Smirnov test variant that examines a sample's possible origin from a normal-distributed population. However, since K-S implicitly assumes some knowledge of the underlying population parameters (actually not available), the Lillifors test (Lilliefors 1967) - which carries out a suitable check with unknown average and standard divergence - was additionally applied. In order to finally safeguard both tests, also results of the newer Shapiro-Wilk test (Shapiro 1965, Shapiro et al. 1968) were consulted.

13   Outliers are data that do not fit the distributional pattern as indicated by the majority of the remaining data. In any case, test statistics rely on adjusted data (i.e. after elimination of outliers). None of these corrections affected the results tendency, though helped to achieve significance every now and then.

14 To put it simply, a correlation describes the magnitude of the relation between two parameters. Spearman's rank correlation coefficient R places comparatively low demands on the data's statistical goodness.

15 Menz et al. (2002) have identified such practices, e.g. different focusings, different self-representations regarding coping with pain, upgrade or demotion and quality of pain description.