

Analysis of heart stress response for a public talk assistant system

Martin Kusserow, Oliver Amft, and Gerhard Tröster

Wearable Computing Lab., ETH Zurich, Switzerland
{kusserow, amft, troester}@ife.ee.ethz.ch
<http://www.wearable.ethz.ch>

Abstract. Conference presentations are stressful communication tasks for many speakers. This mental stress inhibits the speaker’s ability to recall information and perceive the audience. Moreover, stress deteriorates linguistic and paralinguistic capabilities of the speaker. This paper proposes a wearable talk assistant to monitor mental stress and provide relaxation feedback during public speaking. The assistant senses the speaker’s body stress by means of heart activity. With this data the system recognises stressful talk phases. We evaluate the approach in authentic conference talks. The talk assistant was worn by 5 speakers before, during, and after giving a 20 minute talk. Our results demonstrate that it is feasible to distinguish the talk period from the surrounding periods and detect talk phases. These findings show that heart activity provides vital information to estimate the speaker’s body stress. Moreover, we outline ways to proactively support a speaker non-disruptively while talking in order to maximise the presentation performance.

Key words: Speaker monitoring, mental stress, heart rate variability

1 Introduction

Giving a public talk is generally considered as a challenging task. It requires a cognitive performance to communicate expert knowledge of a particular domain. Moreover, public talking is a stressful social situation that stimulates the autonomous nervous system (ANS). In turn, most unexperienced speakers observe symptoms, such as cold hands, sweat, and increased heart rate. Nevertheless, even professional speakers may perceive such body stress, for example when talking in front of an unknown audience.

While ANS activation over a resting level may positively stimulate cognitive capabilities of a speaker [1], elevated levels of stress and anxiety restrain cognitive and communicative capabilities. Consequently the presentation may become unclear, the speaker may be unable to interact with the audience and the talk will miss its goal to transfer vital content. Hence it is desirable to control stress level during public talks.

We believe that a talk assistant system could be deployed that provides relaxation feedback stimuli and automatically adapts feedback during a talk. Several

feedback strategies, applicable even during a talk were discussed in the literature, see Section 1.3. Technically, however, a talk assistant requires an appropriate stress monitoring solution in order to adapt and personalise the feedback.

In order to adapt feedback, a first vital information is to identify stressful talk phases. A talk assistant could fade in a relaxation stimulus during a stressful talk phase and remain quite otherwise. Moreover, the talk assistant feedback may incorporate the stress level during the stressful phase to fine-tune feedback.

1.1 Stress phases in public talks

Body stress is a natural response of the ANS to manage critical situations, due to physical or mental load. Historically, ANS response supported the fight or flight reaction needed when facing an enemy. However physical activity, movement in particular, is not relevant during public speaking. The ANS response for presentations is primarily related to mental effort [1] and anxiety [2] and is reflected in physiological patterns. These sources of ANS response cannot be differentiated [1] and are referred to as body stress in this paper.

Behnke and Sawyer [2] analysed speaker reports after a talk and found that anxiety due to public speaking follows a temporal stimulation sequence. In an anticipatory phase before the talk, anxiety increases, e.g. during preparation and imagination of the performance. At the time being called upon and during the first minute of speaking, anxiety peaks (confrontation phase) and decays thereafter (adaptation phase). However the exact phase timing remains unclear, since a post-hoc analysis based on the speaker's memory was used. Moreover, magnitude and timing of the body reaction depends on speaker and particular situation, e.g. stress symptoms may recur even in a later stage of the talk. Coping strategies to manage these stressful talk situations require long and laborious training.

1.2 Paper contributions

Our goals in this paper were to (1) develop a stress monitoring solution for public talks and (2) evaluate physiological responses as indicators for stressful talk phases.

Regarding the first goal, the challenge of talk stress evaluation is to non-invasively monitor the speaker. In particular, the speaker should not be hindered or influenced by the monitoring approach during a presentation. Hence, classical stress assessments, such as taking saliva samples, are not feasible during a talk. Neither the speaker could be asked during the talk, to rate momentary body stress level.

While public speaking is a general stress stimulus [1], identification of robust physiological indicators applicable for a talk assistant is an open research field. Towards this second goal, we monitored speakers during actual conference talks

In this setting, this paper makes the following contributions:

1. We utilise a commercial heart monitoring chest-belt to analyse seven physiological features from time and frequency-domain. The physiological features included heart period, variables of heart rate variability (HRV) and respiration, extracted from the R-wave signal. Most of these features were reported to indicate body stress when recorded from Electrocardiogram recorders in laboratory settings [3]. We evaluate their capability to distinguish the talk situation from anticipation and post-talk relaxation using the chest-belt device.
2. We analyse actual talks to identify confrontation and subsequent adaptation phases using heart period. We then evaluate a phasing for all physiological features by pattern classification. To this end, our evaluation is a first attempt to confirm the sequence of anxiety reporting provided by speakers after a talk [2]. Finally, we compare the physiological response to talk anxiety self-ratings of the speakers.

Our analysis is based on 7 hours of data recorded from real conference talks of five different speakers. On each occasion, we additionally recorded at least 30 min before and after the talk to monitor anticipation and relaxation.

Vital aspects of our system are simplicity and comfort: the chest-belt is unobtrusive, easy to use, and does not interfere with the speaker’s performance. We used a commercial off-the-shelf device to permit quick deployment and repeatability. Comparable systems typically require electrode attachments, such as for Holter monitors (see Section 1.4).

Below, we present our talk assistant concept. In Section 1.4 we review related works on stress analysis during public speaking. Section 2 presents the recording procedure utilised during the real talk situations and Section 3 outlines the applied data analysis procedures. The results of our work are detailed in Section 4 for the talk situation and Section 5 for the phases within the talk. A discussion of the results is given in Section 6. Section 7 concludes the findings of this work.

1.3 Talk assistant concept

Eventually the talk assistant shall support speakers in their performance, reducing the need for long and laborious training by the speaker. Even during a talk, such feedback can be deployed to support relaxation.

More than 35 years ago researchers first thought of using technical devices to reduce public speaking anxiety and to improve public speaking experience [4, 2]. Therefore, veridical feedback was found to be essential to change physiological responses into the desired direction [5]. Talk stress feedback is related to Biofeedback that is considered a particular useful technique to reduce heart rate due to mental stressors [6, 7]. Nevertheless, feedback during talks must occur without attracting attention of the audience. Examples include dedicated messages to the speaker and relaxation screens in the back of the audience [2]. In addition, particular ways of breathing [8] or speaking [9] can promote parasympathetic activation and reduce stress during talks.

Figure 1 shows our concept for the basic operation of a talk assistant. The ANS can respond with body stress symptoms to a talk situation (stressor). These symptoms include changes in physiological signals and heart activity in particular. The talk assistant records heart period data and adapts feedback systems embedded in the environment. In an initial implementation, feedback could be related to the identified confrontation phase.

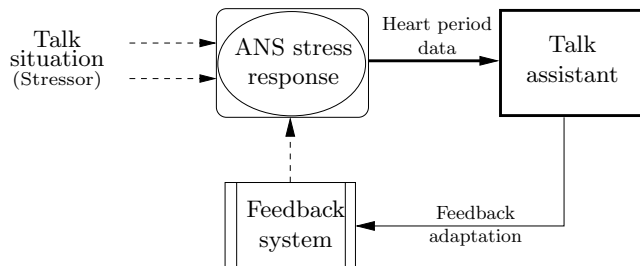


Fig. 1. Basic schematic of the talk assistant operation. This paper focuses on the recording of heart period data and analysis of heart-related physiological stress indications during actual public talks.

1.4 Related work

Dickens and Parker [10] first assessed stage fright in 100 male and female college students. Blood pressure and pulse rate samples were taken directly before and after classroom presentations. Significant increase in pulse rate and blood pressure between pre-talk and post-talk period was observed for over 90% of the speakers. However, no measurement was taken during the talk period. Behnke [4] reported 4 stable characteristic events in cardiac patterns of 24 male college students during public speaking: anticipation, confrontation, adaptation, and release. Confrontation and adaptation were assigned the highest levels of physiological arousal. Later Behnke [2] established two general patterns of public speaking anxiety, habituation and sensitization. They differ in the level of anxiety during the anticipation phase. The degree of anxiety reflected in the pattern of sensitization was further supported by the work of Booth-Butterfield [11] and Pörhölä [1] who categorized heart rate patterns of college students during public speaking.

Different methodologies exist in monitoring physiological signals during public speaking. The duration of the public speaking task varied from 2 minutes [11] up to 7 minutes [1]. Typically, the talk situation was manually partitioned into anticipation, confrontation, adaptation, and release phase each lasting from 30 seconds [11] up to 2 minutes [12].

Audience size and composition varied from 15 classmates [13] to 150 college students [10].

Topics of the speech were hobby or favorite activity [2], topic of own choice [1] or any other informative topic [4] rather than a scientific topic. Investigations were solely done in university classroom environments and not in front of a scientific conference audience. A comprehensive summary of these methodologies can be found in the work of Pörhölä [1].

Recording of heart rate during public speaking was done by hand [10], paper-based physiograph [4], heart rate monitors attached to the index finger [11], ear-worn device [1], and an ECG system [14].

Compared to blood pressure and skin conductance level (SCL) only heart rate yielded significant effects related to the talk situation itself [5]. Cardiac response was found to clearly differentiate between anticipation, confrontation and release phase in comparison to SCL [14].

While heart activity was confirmed to be a primary stress measure during talks, cardiac features were not analysed in detail. Most related works assumed a static phase or moment to determine stress levels of anticipation, confrontation and release. Moreover, it remains open, whether HRV or respiration features provide a robust indication for the talk situation and the talk phase timing. Moreover, a chest-belt measurement towards a talk assistant system was not investigated.

2 Conference talk recording

A scientific talk of five PhD and Masters students (aged 23 to 30 years, 1 female) was included in our investigation. The talks were given at three research conferences (ISWC 2007, IOT 2008, and an ETH Electronics institute colloquium). In all situations, the speakers presented their research results to an audience of 30–50 experts. The experts were of mostly unknown identity to the speakers. The speaker’s performance was neither recorded on audio or video nor rated by members of the audience. All speakers have had public speaking experience with comparable audiences. None of the speakers had known clinical anxiety disorders or cardiovascular diseases.

Figure 2 illustrates the talk recording procedure. At start of the conference session speakers sat amongst the audience and followed the presentation of preceding talks (pre-talk period). Subsequently, speakers delivered their talk (talk period). A question and answer (Q&A) period followed the talk. Finally, speakers returned to their seat in the audience to listen to remaining talks of the conference session (post-talk period).

For each speaker, a pre-talk period of at least 30 min before begin of the actual talk was recorded. After the talk and Q&A period, recording was continued for another 30 min (post-talk period). The average talk duration was 21 ± 7 min, without Q&A period. After attaching the monitoring system, speakers were asked to follow their normal talk preparations. An observer followed the speaker’s activities and annotated the recordings.

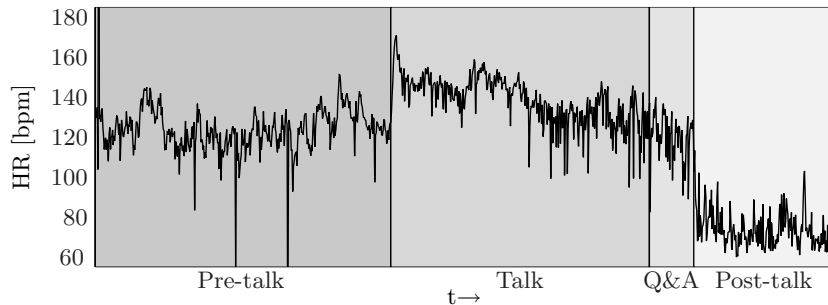


Fig. 2. Heart rate (HR) signal during the conference talk of a representative speaker. Pre-talk, talk, question and answer (Q&A), and post-talk period. Notice that peak HR is at ~ 170 bpm at beginning of the talk.

The wearable recording system consisted of a commercially available Suunto heart rate monitor chest-belt^a and the Q-Belt Integrated Computer (QBIC) [15]. Figure 3 depicts the wearable sensor system. The heart monitor chest-belt provides heart period data (RR intervals) by measuring the time between consecutive QRS complexes. Data was wirelessly transmitted to the QBIC using the ANT communication protocol^b. For data recording on QBIC the Context Recognition Toolbox (CRNT) [16] was used.

For our research work, QBIC offered the flexibility to replace chest-belts and add further sensors without redesigning the recording and analysis procedure.

For the speakers, this procedure did not require interaction with the recording system, except to put it on and off. In addition all speakers answered the personal report of confidence as a speaker questionnaire (PRCS) [17]. PRCS measures general anxiety, subjectively perceived by the speaker during public speaking. Scores range from 0 (not anxious) to 30 (highly anxious).

3 Analysis methods

HRV analysis attempts to separate the types of ANS activation using features of the time and frequency domain. In particular, two frequency bands are associated with ANS activation: low-frequency (LF) band (0.04–0.15 Hz), reflecting parasympathetic and sympathetic activation, as well as high-frequency (HF) band (0.15–0.4 Hz) reflecting parasympathetic activation [18]. As there is no consensus on the optimal approach to extract HRV frequency domain features from RR intervals, we used a non-parametric FFT-based method according to guidelines established in [19].

Heart activity features were derived for the analysis and comparison of (1) all recording periods (pre-talk, talk, post-talk) and (2) the analysis of phases within the talk period. The first analysis shows the relation of heart activity to the entire

^a <http://www.suunto.com>

^b <http://www.thisisant.com>



Fig. 3. Wearable system used for the recording of heart period data during the conference session. Left: Heart monitor belt attached to the thorax and QBIC belt computer for data storage. Right: Speaker wearing the system during the conference talk.

talk situation. In the second analysis, we investigated whether the confrontation phase can be identified and discriminated from the remaining talk.

3.1 Heart period signal preprocessing

The signal processing procedure was adapted to RR interval data provided by the chest-belt. In particular, signal preprocessing needed to compensate for errors in the RR interval detection.

Heart period was recorded as timestamped RR interval. Non-parametric analysis of heart period data requires an evenly sampled signal and stable first and second order moments across time [19]. Thus, to compute HRV features the raw RR intervals were filtered, interpolated, and detrended. Figure 4 summarises the signal processing flow.

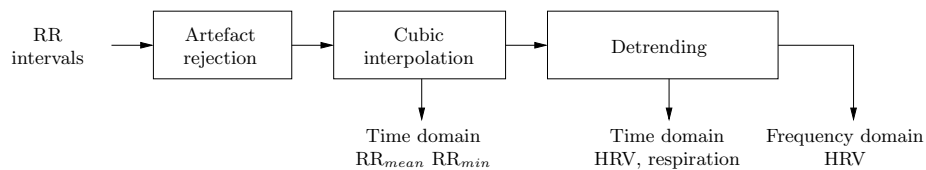


Fig. 4. Feature extraction process. Irregularly sampled RR intervals were filtered, interpolated and detrended.

In a first step, spurious RR intervals caused by added or missed R-peaks during QRS detection were filtered from the irregularly sampled time series.

Spurious RR intervals that differed by more than 20% compared to their predecessor were removed [20].

Subsequent cubic interpolation was applied to obtain an evenly sampled time series at 8 Hz sampling frequency that permits the analysis of high heart rate levels [21].

After interpolation, detrending was applied to obtain a weakly stationary time series. As detrending method we used a modified smoothness priors approach [22], where the cut-off frequency can be adjusted by a single parameter λ and data end-point distortion is avoided. With our modification using a sliding window approach, detrending can be performed online.

The smoothing parameter λ was set to 1600 which equals a cut-off frequency of ~ 0.04 Hz at a sampling frequency of 8 Hz. This prevents attenuation of the relevant LF frequency band, starting at 0.04 Hz.

Spectral power in LF and HF was computed from the estimation of the power spectral density (PSD), for which Welch's method was used [23]. A common sliding window size of 1024 samples (128 s, Hanning window) and a step size of 512 samples (64 s) was used. These settings provide a frequency resolution of ~ 0.0078 Hz and satisfy the recommended length of at least 2 min for analysis of the LF frequency band [19].

3.2 Heart activity feature extraction

Features of HRV in time and frequency domain were selected according to the recommendations made in [24]. In the frequency domain, the following HRV features were computed: low-frequency spectral power (LF), high-frequency spectral power (HF), and LF/HF ratio. The following time domain features were computed using the same sliding window configuration: standard deviation of all RR intervals (SDNN)^c, and standard deviation of differences between adjacent RR intervals (SDSD).

In addition to the HRV features the minimum RR interval (RR_{min}), time of occurrence of RR_{min} (τ_{min}), and respiration frequency (f_{resp}), calculated from respiratory sinus arrhythmia (RSA), were computed. For calculation of f_{resp} the advanced counting method [25] was used. For that method, performance was found to be superior compared to other RSA techniques.

4 Talk situation analysis

In the talk situation analysis we investigated the relation of heart activity features between anticipation (pre-talk), relaxation (post-talk) and the talk period. With this approach we identified features that change significantly during the talk period.

We computed features (see Section 3) for the following time segments: last 10 min before talk begin (pre-talk period), first 10 min after talk begin (talk

^c RR intervals are also referred to as NN (normal-to-normal) intervals [24].

period), and first 10 min after end of the Q&A period (post-talk period). These segment durations were chosen to provide an acceptable number of observations for analysis and compare features across the three periods. The Q&A period was not included in the talk period as the direct interaction of speaker and audience may raise different physiological responses.

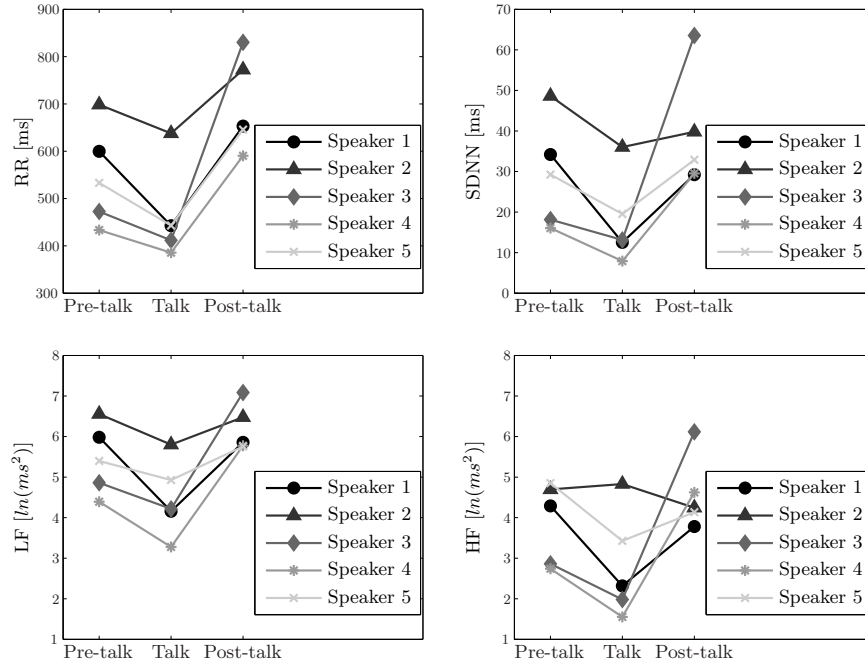


Fig. 5. Mean RR interval (top left), mean SDNN (top right), mean LF power (bottom left), and mean HF power (bottom right) of the pre-talk, talk, and post-talk period for all five speakers.

Figure 5 shows four selected time and frequency domain features (mean RR interval, mean SDNN, mean LF power, and mean HF power) for all subjects across the three periods.

Mean RR interval decreases from pre-talk to talk period by 50 to 150 ms and significantly increases by 200 to 400 ms from talk to post-talk period. For all speakers, this result confirms the strong variations in heart rate shown in the example heart rate plot in Figure 2. Lowest mean RR interval occurs in the talk period, highest mean RR interval in the post-talk period. This pattern conforms to previous findings [2].

Mean SDNN shows a similar pattern however, in contrast to mean RR interval, pre-talk and post-talk levels were similar. Mean LF power is higher than mean HF power across all periods and speakers. Mean LF power drops during

the talk period and recovers to pre-talk level in the post-talk period. Except for speaker 2, mean HF power follows a similar pattern.

Speaker 2, in contrast to the other speakers, shows a slight increase in mean HF power during the talk and a decrease in the post-talk period. Only speakers 3 and 4 exceed their pre-talk level in mean LF and mean HF power during the post-talk period.

Table 1. Paired t-test for talk vs. pre-talk and talk vs. post-talk periods at the 5% significance level for all features; (n.s. = not significant).

Talk vs. pre-talk p-values							
	RR	LF	HF	LF/HF	f_{resp}	SDSD	SDNN
Speaker 1	< 0.001	0.009	0.002	n.s.	0.027	< 0.001	< 0.001
Speaker 2	< 0.001	0.012	n.s.	0.002	n.s.	n.s.	0.004
Speaker 3	0.002	n.s.	n.s.	n.s.	n.s.	n.s.	0.021
Speaker 4	0.001	0.036	n.s.	n.s.	n.s.	0.033	0.002
Speaker 5	< 0.001	0.035	0.037	n.s.	0.037	0.037	0.015
# sig. results	5	4	2	1	2	3	5

Talk vs. post-talk p-values							
	RR	LF	HF	LF/HF	f_{resp}	SDSD	SDNN
Speaker 1	< 0.001	0.004	0.007	n.s.	n.s.	< 0.001	< 0.001
Speaker 2	< 0.001	0.018	n.s.	0.035	n.s.	0.041	n.s.
Speaker 3	< 0.001	0.001	< 0.001	0.005	0.036	< 0.001	< 0.001
Speaker 4	< 0.001	0.004	< 0.001	0.011	n.s.	< 0.001	< 0.001
Speaker 5	< 0.001	0.011	0.048	n.s.	n.s.	0.005	0.001
# sig. results	5	5	4	3	1	5	4

Paired t-tests were used to analyse the differences in feature means of the talk period with both surrounding periods. The analysis was made for a hypothesis that two matched samples come from distributions with equal means at a significance level of 5%. Table 1 details the speaker-specific p-values and number of significant results per feature for talk vs. pre-talk and talk vs. post-talk periods.

For all speakers RR intervals were significantly different between pre-talk, talk, and post-talk periods. Analysis of pre-talk and talk periods showed signif-

icant differences across all speakers in SDNN. For talk and post-talk periods, features LF and SDSD showed a significant difference in addition to RR for all speakers. Speaker 3 showed a significant difference for talk and post-talk periods across all features. All other features showed sparse significance for individual speakers only.

Lilliefors’ goodness-of-fit test of composite normality was applied to prove normal distribution of the features prior to the t-test. Nevertheless it must be noted that the eight observations available per feature in each period restricts the result interpretation. We assume it here as a preliminary indication of informative features that distinguish the talk period from the surrounding periods.

5 Talk phases analysis

We investigated whether the confrontation and adaptation phases within a talk can be identified and separated.

Without considering a static partitioning of the talk period into confrontation and adaptation phase, we determined three distinct points in time with respect to talk start time: (1) time τ_{min} until minimum RR interval RR_{min} occurs, (2) time τ_{ant} until the RR interval first reaches mean RR interval RR_{ant} of the anticipation phase (pre-talk), and (3) time τ_{rel} until the RR interval first reaches the mean RR interval RR_{rel} of the release phase (post-talk). Table 2 details the speaker-specific results.

Table 2. Talk characteristics analysis: speaker-specific minimum RR interval RR_{min} and time of occurrence τ_{min} , time τ_{ant} until recovery to mean RR interval of anticipation phase, and PRCS score. τ_{min} , τ_{ant} are relative to talk start time.

	RR_{min} [ms]	τ_{min} [s]	τ_{ant} [s]	PRCS Score
Speaker 1	404.7	218.9	n.a.	8
Speaker 2	500.1	12.6	262	3
Speaker 3	348.8	18.1	629	11
Speaker 4	349.9	81.8	580	20
Speaker 5	397.0	38.5	n.a.	2

Except for speaker 2, RR_{min} ranged from 348 ms (172 bpm) to 404 ms (146 bpm). Speaker 2, similar to the talk situation analysis, showed a higher RR_{min} of 500 ms (120 bpm). The minimum RR interval occurs within the first 12 to 82s for all five speakers^d.

^d In fact, speaker 1 also showed a local minimum RR interval of 406 ms at 37s after talk begin. Due to a failing technical demonstration during the talk, the global minimum RR interval was found 219s after talk begin.

Except for speaker 1 and 5, τ_{ant} is within 262 to 629 s. For speakers 2 to 4 this indicates that lower RR_{min} is related to longer τ_{ant} . Although in range of the other speaker's, RR_{min} of speaker 1 and 5 did not recover to RR_{ant} . None of the five speakers recovered to RR_{rel} within the talk period.

Moreover, table 2 shows the PRCS scores of all speakers. We used PRCS as an indicator of how the measured physiological response related to generally self-perceived stress during a talk. We observed that lower RR_{min} is reflected by higher PRCS scores (more anxious). However, speaker 5 reported the lowest PRCS score that does not correspond to the RR_{min} result. It would rather corresponds to RR_{min} of speaker 1 having a PRCS score of 8.

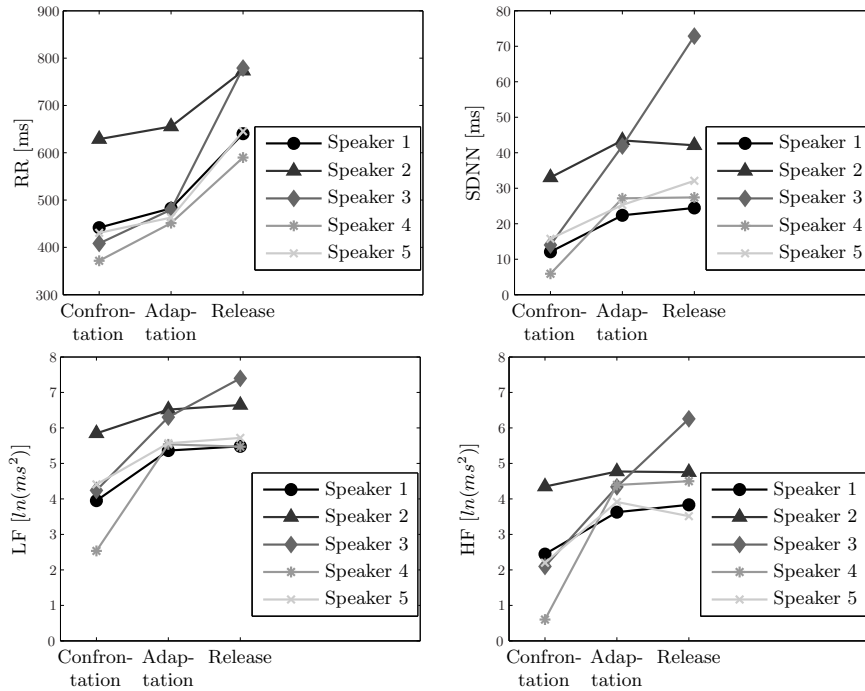


Fig. 6. Mean RR interval (top left), mean SDNN (top right), mean LF power (bottom left), and mean HF power (bottom right) of the confrontation, adaptation, and release phase for all speakers.

In order to analyse the hypothesis of two talk phases (interpreted here as confrontation and adaptation) for all features, we computed features (according to Section 3) for the first 5 min and last 5 min of the talk period. Similar to the talk situation analysis, we excluded the Q&A period in this investigation. Ideally, the talk assistant should require measurement data from the talk period only in order to adapt feedback. Moreover, the durations represent a tradeoff between number of observations and assumed stationarity of the physiological state. For

comparison to the relaxation effect after the talk period we additionally derived features for the first 5 min of the post-talk period (release phase).

Figure 6 visualises the speaker-specific results for confrontation, adaptation, and release phases. Results are represented for mean RR interval, mean SDNN, mean LF power, and mean HF power.

Mean RR interval, mean SDNN, mean LF power, and mean HF power showed an increase from confrontation to adaptation phase. Increase in mean RR interval is even larger from adaptation to release phase. Except for speaker 3, whose features continue to increase in the release phase, mean SDNN, mean LF power, and mean HF power remain similar between adaptation and release phases. These results indicate an adaptation trend during the talk.

In order to determine whether the two talk phases could be discriminated, we applied a Naïve Bayes classifier. Moreover, we estimated in this analysis those features that are particular informative to indicate a phase structure.

From the computed feature set of each speaker we included 3 observations of the confrontation and adaptation phase respectively. We performed a leave-one-out cross-validation to determine training and testing set for the classification. For classifier training 2 observations were used, 1 observation was used for testing. Figure 7 shows the speaker-specific classification performance for all features.

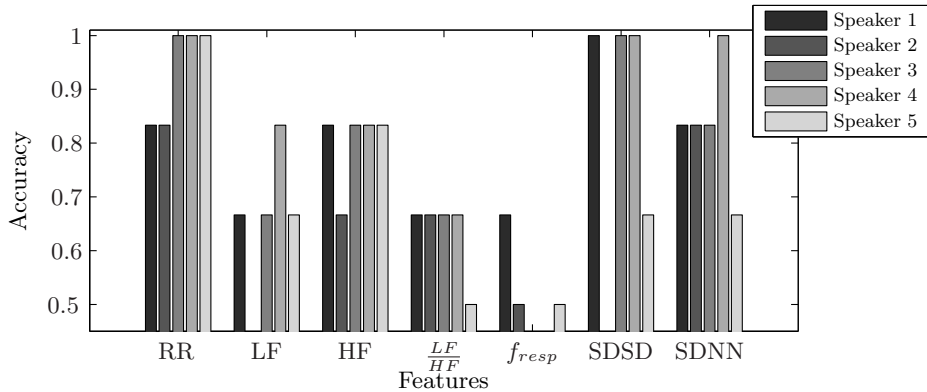


Fig. 7. Speaker-specific leave-one-out discrimination test of confrontation and adaptation talk phases for all features.

Among the features mean RR interval showed the best classification performance along with mean SDDS, mean LF power and mean SDNN. Mean HF power, mean LF/HF ratio, and mean respiratory frequency f_{resp} performed less well. Except for mean RR interval, features showed high variations between speakers.

6 Discussion

6.1 Talk situation analysis

Both, the visual analysis and the t-test of heart activity features showed that the talk period was best identified from time-domain features. In particular, the mean RR interval showed significant differences when compared to pre-talk and post-talk periods. In contrast, the averages for LF/HF and respiration frequency f_{resp} did not change in a similar way. This is an interesting observation, since the LF/HF ratio is frequently cited as stress indicator in laboratory studies [3]. Nevertheless, public speaking is a clear cue for mental stress and hence is reflected in the body stress response.

In our analysis, we found that in particular the HF feature is less stable. We assume that these results are related to the talk scenario: speaking can influence the LF and HF power since irregular breathing can affect both frequency bands [26]. Moreover, using the chest-belt device may deteriorate frequency-domain features due to deviation in RR interval measurement compared to a wet electrode ECG devices. However, a classic wet electrode ECG device would have been more cumbersome for a speaker and is less acceptable for a talk assistant system. We plan to investigate this effect in the future.

Nevertheless, these results do indicate that the chest-belt device provides relevant information in temporal features under real conference talk situations. These findings were confirmed by our subsequent analysis of talk phases.

We observed that one speaker (speaker 2) had larger values for the RR interval in all periods while maintaining the same overall pattern, compared to others. These results were attributed to the high level of physical fitness of that speaker.

6.2 Talk phases analysis

The talk phases analysis confirmed a confrontation phase including τ_{min} for all speakers during the first minutes of the talk. After this distinct time, RR interval increased differently for each speaker. Two speakers did not reach the anticipation level, all speakers missed the average post-talk level. Hence the talk was activating for all speakers, while some could not recover from the confrontation level heart activity.

To this end, our analysis procedure can be directly applied for a talk assistant system. We considered the time from talk begin to τ_{ant} , including τ_{min} as most challenging moment during the talk. Whereas, when assuming a steady process, the two remaining points τ_{ant}, τ_{rel} are indicators of a recovery activity from this particular activation state of the body. Thus, assuming two phases during a talk period, the goal of a talk assistant system should be to minimise τ_{ant} . The reference measure RR_{ant} is available to the system when the assistant is switched on before the talk. This procedure requires no further calibration to RR_{rel} or mean resting RR interval. Nevertheless, the speaker should have

the option to adjust the feedback online. This can be achieved by adapting confrontation phase identification threshold incorporating individual heart activity limits [1] corresponding to either favourable arousal or unfavourable stress.

The visual feature analysis showed a most consistent relation between assumed talk phases for the RR interval. This observation was confirmed by the classification test. It is important to note that the two phases classification is an inappropriate relaxation feedback to the speaker. Rather it should trigger relaxation cues incorporating the temporal trend of the features.

The PRCS scores of speaker anxiety confirmed the estimated RR_{min} result, with one exceptional speaker. We attributed this observation to an inaccurate questionnaire response.

Monitoring of behaviour and physiological response in a conference environment is clearly limited in data size and constant conditions: talk length is fixed by the conference session constraints. Artificial extension under laboratory condition may not elicit the same physiological responses as the field study since social implications and talk conditions change.

A limitation for this investigation was the low number of observations available for each talk phase, constrained by the tradeoff between resolution of the frequency-domain features and assumed stationarity in physiological responses. For the spectral features, a time window of 2 min was needed to analyse the spectral LF band, while a duration of 5 min was assumed as upper limit for a stable physiological state. Physiological response in this particular scenario, however, may contain a number of short time phenomena, too short for sufficient frequency resolution using FFT-based PSD estimation [27]. We plan to investigate this issue in the future.

7 Conclusion

In this paper we analysed heart stress response during actual conference presentations using a body-worn monitoring system. We expect that this system can be used in a talk assistant system that supports a speaker during stressful talk situations with automatic relaxation feedback. Different options for such feedback have been proposed in the literature.

Towards the talk assistant system we addressed two most critical challenges. Firstly, a comfortable speaker monitoring system is needed to measure body stress of the speaker. For this purpose, we deployed a commercial heart monitor chest-belt that can transfer readings wirelessly to an on-body or room-installed base system. Secondly, relevant physiological information must be obtained from the speaker in order to adapt and personalise assistant feedback. To this end we investigated seven features of heart activity from time- and frequency-domain that were reported to indicate body stress in laboratory investigations [3].

Our investigations showed that time-domain features, in particular heart period, can provide robust information for the talk situation. Moreover, these features help to discriminate the talk phases. Using this phasing information a talk

assistant system could adapt feedback during the confrontation phase. Minimising the duration that the speaker stays in this phase is the primary goal of the talk assistant.

Classical HRV stress indicators did not respond as expected in our analysis. This finding requires further investigation and validation of a chest-belt device, in particular for public talk situations. Currently, we investigate the use of complementary sensor modalities to include additional information in the speaker's body stress estimation. To this end, we like to investigate whether a simple notification or reminder is an acceptable and effective speaker feedback solution.

Acknowledgements

We would like to thank the conference speakers for participating in the talk recordings.

References

1. Pörhölä, M.: Arousal styles during public speaking. *Communication Education* **51**(4) (2002) 420–438
2. Behnke, R.R., Sawyer, C.R.: Public speaking anxiety as a function of sensitization and habituation processes. *Communication Education* **53**(2) (2004) 1164 – 1173
3. Pagani, M., Lucini, D., Rimoldi, O., Furlan, R., Piazza, S., Biancardi, L.: 20. In: *Heart Rate Variability*. Futura Publishing Company, Inc. (1995) 245–266
4. Behnke, R.R., Carlile, L.W.: Heart rate as an index of speech anxiety. *Speech Monographs* **38** (1971) 65–69
5. Rohrmann, S., Hennig, J., Netter, P.: Changing psychobiological stress reactions by manipulating cognitive processes. *International Journal of Psychophysiology* **33** (1999) 149–161
6. Sharpley, C.F.: Biofeedback training versus simple instructions to reduce heart rate reactivity to a psychological stressor. *Journal of Behavioral Medicine* **12**(5) (1989) 435–447
7. McKinney, M.E., Gatchel, R.J.: The comparative effectiveness of heart rate biofeedback, speech skills training, and a combination of both in treating public-speaking anxiety. *Biofeedback and Self-Regulation* **7**(1) (1982) 71–87
8. Sakakibara, M., Takeuchi, S., Hayano, J.: Effect of relaxation training on cardiac parasympathetic tone. *Psychophysiology* **31** (1994) 223–228
9. von Bonin, D., Frühwirth, M., Heuser, P., Moser, M.: Effects of speech therapy with poetry on heart rate variability and well-being (in German). *Research in Complementary Medicine* **8**(3) (2001) 144–160
10. Dickens, M., Parker, W.R.: An experimental study of certain physiological introspective and rating-scale techniques for the measurement of stage fright. *Speech Monographs* **18**(4) (1951) 251–259
11. Booth-Butterfield, S.: Action assembly theory and communication apprehension - a psychophysiological study. *Human Communication Research* **13**(3) (1987) 386–398
12. Behnke, R.R., Beatty, M.J.: A cognitive-physiological model of speech anxiety. *Communication Monographs* **48** (1981) 158–163

13. Beatty, M.J., Behnke, R.R.: Effects of public speaking trait anxiety and intensity of speaking task on heart rate during performance. *Human Communication Research in Complementary Medicine* **18**(2) (1991) 147–176
14. Croft, R.J., Gonsalveza, C.J., Gandera, J., Lechema, L., Barry, R.J.: Differential relations between heart rate and skin conductance, and public speaking anxiety. *Journal of Behavior Therapy and Experimental Psychiatry* **35**(3) (2004) 259–271
15. Amft, O., Lauffer, M., Ossevoort, S., Macaluso, F., Lukowicz, P., Tröster, G.: Design of the QBIC wearable computing platform. In: *ASAP 2004: Proceedings of the 15th IEEE International Conference on Application-specific Systems, Architectures and Processors*. (2004) 398–410
16. Bannach, D., Amft, O., Lukowicz, P.: Rapid prototyping of activity recognition applications. *IEEE Pervasive Computing* **7**(2) (2008) 22–31
17. Paul, G.L.: Insight versus desensitization in psychotherapy: An experiment in anxiety reduction. PhD thesis, Stanford University, Palo Alto, CA (1966)
18. Malik, M., Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology: Heart rate variability - standards of measurement, physiological interpretation, and clinical use. *European Heart Journal* **17** (1996) 354–381
19. Berntson, G., Jr, J.T.B., Eckberg, D.L., Grossman, P., Kaufmann, P.G., Malik, M., Nagaraja, H.N., Porges, S.W., Saul, J.P., Stone, P.H., van der Molen, M.W.: Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology* **34** (1997) 623–648
20. Kleiger, R.E., Miller, J.P., Jr., J.T.B., , Moss, A.J., the Multicenter Post-Infarction Research Group: Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *The American Journal of Cardiology* **59** (1987) 258–262
21. Singh, D., Vinod, K., Saxena, S.: Sampling frequency of the RR interval time series for spectral analysis. *Journal of Medical Engineering and Technology* **28**(6) (2004) 263–272
22. Tarvainen, M.P., Ranta-aho, P.O., Karjalainen, P.A.: An advanced detrending method with application to HRV analysis. *IEEE Transactions on Biomedical Engineering* **49**(2) (2002) 172–175
23. Welch, P.D.: The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics* **AU-15**(2) (1967) 70–73
24. Malik, M., Camm, A.J.: *Heart Rate Variability*. Futura Publishing Company, Inc. (1995)
25. Schäfer, A., Kratky, K.W.: Estimation of breathing rate from respiratory sinus arrhythmia. *Annals of Biomedical Engineering* **36** (2008) 476–485
26. Beda, A., Jandre, F.C.: Heart-rate and blood-pressure variability during psychophysiological tasks involving speech: Influence of respiration. *Psychophysiology* **44**(5) (2007) 767–778
27. Cerutti, S., Bianchi, A.M., Mainardi, L.T.: 5. In: *Heart Rate Variability*. Futura Publishing Company, Inc. (1995) 63–74