Listening Effort: Cardiovascular Investigation Through the Point Process

Edoardo Maria Polo[1], Maximiliano Mollura[2], Alessia Paglialonga[3] and Riccardo Barbieri[2]

[1] DIAG, Sapienza University of Rome, Italy
[2] Politecnico di Milano, Milano, Italy
[3] National Research Council of Italy (CNR), Institute of Electronics, Information Engineering and Telecommunications (IEIIT), Milano, Italy

Abstract

In the context of quantifying listening effort, traditional hearing tests do not provide any information about the stress experienced during listening tasks. Although there have been attempts to quantify memory allocation during acoustics tests, there is no agreement in the literature regarding the role of physiological indices in characterizing different listening effort levels. To this extent, the aim of our study is to ascertain if cardiovascular measurements continuously recorded during the task can help in quantifying listening effort. In the presented protocol, 21 normal young hearing subjects performed a validated speech-in-noise test at two fixed effort levels, while electrocardiogram (ECG) and blood volume pulse (BVP) were continuously recorded. From these time series, the RR series, the amplitude difference between each systole and diastole and the pulse arrival time were extracted. In addition, the ECG-derived RR series were modelled trough a point process framework, yielding instantaneous cardiovascular and autonomic indexes to be considered in our statistical analysis. Overall, the average modelled RR intervals and the pulse arrival time were found effective in distinguishing the two different effort levels (p=0.031 and p=0.016). In addition, the amplitude difference between each systole and diastole was able to significantly separate high effort from both low effort and the initial resting period (p<e-3).

1. Introduction

Listening effort is defined as 'a specific form of mental effort that occurs when a task involves listening' [1]. Quantifying this kind of effort would be very important to understand the listeners' challenges that are usually not identified by traditional audiometric measures [2]. Usually to measure the listening effort, the amount of capacity allocated to perform acoustic tests is estimated by the number of items (words/sentences) recalled. Very few study investigated listening effort by means of physiological signals and moreover there is no a complete agreement in the literature about how our body reacts according to increased task demand in listening. The most accredited physiological measure in this field is given by the changes in the pupillary diameter [3]. Neuroimaging evidences, also, revealed an increased activity in the cingulo-opercular network for degraded vocal stimuli [4] and other brain regions were found to be more active during degraded speech with respect to normal one [5]. The major lack of findings is related to autonomic nervous system (ANS) indexes which are also less burdensome to compute with respect to central measures. Indeed, few study are present and of those, results are promising but still not conclusive. In [6] increased skin conductance amplitude and decreased high frequency power extracted from the R-R spectrum from the electrocardiogram (ECG) signal were found significantly different in both fast and normal speech but only with respect to baseline. In [7], instead, they found differences in the electrodermal response between two different degraded speech (two-talker babble vs. speech-shaped noise) and the pulse amplitude of the photoplethysmographic signal appeared to increase (but not significantly) when masked speech was proposed with respect to unmasked one. For audiologists, an objective index of listening effort (as a physiological one) could complement current clinical assessment tools to inform counselling sessions, provide information on intervention strategies and shed light on cases where there is uncertainty about the need for intervention [1]. The aim of the study is the investigation of cardiovascular features linked to ANS which can reveal physiological behaviors according to two different listening difficulty levels (low and high). In the present study, indeed, we used a validated adaptive speech-in-noise test while monitoring ECG and blood volume pulse (BVP) signals on 21 normal hearing subjects. In particular we applied the point process framework to analyze the ECG signal. The advantage of using the point process is the possibility to compute heart rate variability (HRV) features in short time windows. For the computation of HRV features with traditional methods, indeed, at least 5-minutes recording are recommended [8] in order to have reliable measures. By using the point process, instead, short-term changes in HRV can be monitored using short recording windows, thus limiting test duration and the associated fatigue.

2. Methods

2.1. Study Design and Data

In this study, we used a recently developed, adaptive speech-in-noise test which was validated on both normal hearing and hearing impaired subjects [9]. The speech stimuli used are Vowel-Consonant-Vowels (VCVs) which are degraded by a gaussian white noise superimposed. Starting from the first speech stimulus at +8 dB SNR, in each trial the subject has to choose among three possible VCVs. During the test the SNR of each VCV increases or decreases according to incorrect or correct responses, respectively. Test difficulty and duration depends this way on the ability of the subject in discriminating speech stimuli in noise. The full test and procedure is explained in [10]. Usually in speech-in-noise tests the outcome is the speech reception threshold (SRT) defined as the minimum SNR at which an individual can recognize a certain percentage of the speech material (i.e., 79.4% in the three-alternatives design here used). In order to find two different grades of difficulty we used the individual SRT to distinguish those two levels. In particular, since the just noticeable difference in speech for changes in SNR and for changes in intelligibility is around +3 dB SNR [11], we defined low difficulty trials those at which the SNR was higher than the individual SRT+2 dB SNR and high difficulty trials when the SNR was lower or equal to the individual SRT+2 dB SNR. In particular, for each subject we analyzed same length segments related to consecutive trials at two fixed levels, i.e. low (L) and high (H) difficulty. Our experimental protocol includes 13 females and 8 males (age: 26.18 ± 1.47 years). Before performing the test, all participants underwent pure-tone audiometry on both ears to ensure that their hearing thresholds were in the normal hearing range (pure tone average thresholds among 500, 1000, 2000 and 4000 Hz $<$ 20 dB HL). Moreover, in order to minimize behavioral heterogeneity, participants were asked to avoid coffee and smoke for two hours prior to the experiment. Just before the start of the speech-in-noise test, 2-min of baseline was recorded while subjects were looking at a grey screen. Each feature in the effort phases was then compared with the same feature computed on the same amount of time during the baseline (B). This protocol has been approved by the Ethical Committee of Politecnico di Milano.

2.2. Data and Feature Extraction

ECG and BVP signals were acquired by Procomp Infinity (sampling frequency = 2048 Hz) and they were then filtered by a 4th order zero-phase low-pass Butterworth anti-aliasing filter with a cut-off frequency at 125 Hz and 25 Hz, respectively. Subsequently, both signals were downsampled at 250 Hz. R-peak locations were extracted on the ECG signal by means of a Pan-Tompkins based algorithm and the location and related amplitude values of systolic, diastolic and onset events from the BVP signal were extracted and synchronized with the R-peaks. In more details, systolic and diastolic values were found as maxima and minima between two subsequent R-peaks while onsets were found by looking at the inflection points between each diastolic and the following systolic value. All fiducial points extracted on ECG and BVP signals were then manually checked to be sure of the right locations. The RR series computed from the ECG signal was modelled within a point process framework, which is fully explained in the next section, because of its ability in dealing with non-stationary signals and at the same time tracking fast changes induced by external stimuli [12]. From the resulting time-varying estimation we were able to extract the modelled RR series (μ_{RR}), its variability (σ^2_{RR}), the power spectral density of RR in very low (*RR VLF*), low (*RR LF*) and high (*RR HF*) frequencies, *LF/HF*, the normalized power spectral density of RR in low (*RR LFn*) and high (*RR HFn*) frequency ranges and the total power spectral density of RR (*RR TOT*). Systolic and diastolic values of the BVP signal were used to compute the blood volume amplitude (*VA*) as the average difference between each systole and diastole. Onset locations on the BVP signal were used to compute the pulse arrival time (*PAT*) as the average time interval between each pulse onset and the corresponding R-peak.

Figure 1 shows an example of the RR series and μ_{RR} of one subject during the test.

2.3. Point Process

A point process is a stochastic model underlying the occurrence of events in time or space. The intrinsic point process nature of the RR interval which relies on the R-wave events (R_k) in the ECG signal makes this model a well-suited framework for modeling the heartbeat. Specifically, we modelled the inter-beat-interval series according to the following history-dependent inverse gaussian distribution:

$$p(t) = \left(\frac{\theta_{p+1}}{2\pi(t - R_k)^3} \right)^{\frac{1}{2}} \exp\left(-\frac{\theta_{p+1}(t - R_k - \mu_{RR})}{2\mu^2_{RR}(t - R_k)} \right) \tag{1}$$

whose expected value is estimated with an autoregressive (AR) model according to the following formulation:

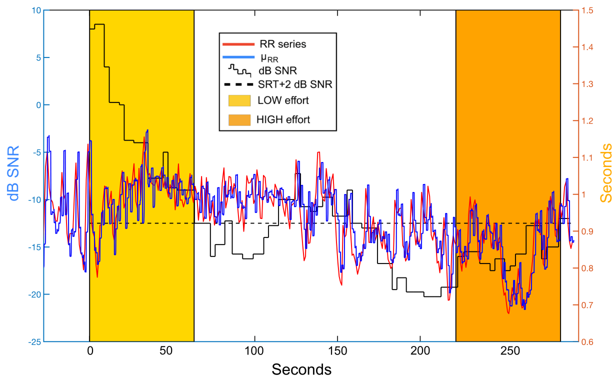

Figure 1. An example of the measured physiological response during the test is reported for one subject. RR series and μ_{RR} are reported in red and blue, respectively. The dark staircase represents the dB SNR outgoing during the test. The dark dotted line represents the threshold (SRT+2dB SNR) of the subject.

$\mu_{RR} = \theta_0 + \sum_{i=1}^{p} \theta_i(t)RR_{k-i}$, where RR_k represents the k-th RR interval closer to time t. AR parameters, $\theta_{0:p}$, and the shape parameter θ_{p+1} are continuously estimated through local maximum likelihood estimation according to [12]. Time-varying spectral indices are derived from the estimated AR coefficients.

3. Results

Statistical tests were performed to compare each feature extracted from ECG and BVP signals among the two stress levels (L and H) and the baseline (B). Specifically, Friedman's test was performed if at least one variable was not normal and two-way ANOVA test was performed if all variables were normal. In all comparisons Bonferroni correction was applied. Table 1 reports the median value and the median absolute deviation of all the features computed for all three levels. From the statistical analysis we found out that μ_{RR} is significantly different between L and H (p=0.031) with an observed decrease from L to H. This feature in B instead is not statistically different from L and H, even if an increase is present with respect to H. Looking at the frequency features, *RR VLF* is significantly lower in H with respect to B (p=0.026) and the decreasing trend is however maintained also with respect to L. The same happens for *RR TOT* (p=0.016). Although no significance is present for *RR LF*, *RR HF*, *RR LFn*, *RR HFn* and *RR LF/HF*, everything seems consistent with a greater sympathetic activation in the high effort phase. In particular, *RR LFn* and *RR LF/HF* show increasing trends from B and L with respect to H and consistently, *RR HF* and *RR HFn* show opposite trends. Only *RR LF* does not perfectly agree since the highest value is in B, but it is higher in H with respect to L. However, this is unconfirmed in *RR LFn* which is more related to the sympathetic system, since it refers to a proportion excluding *RR HF*, which conversely is related more to the parasympathetic system [13]. The same can be said for σ^2_{RR} which represents the variability of μ_{RR}: in the H phase a lower variability can be observed with respect to both B and L.

PAT follows the same behavior of μ_{RR}: decreasing values in H with respect to B and L with significant differences between L and H (p=0.016).

VA has shown to be the most effective feature for discriminating the three effort levels: a significant decrease, indeed, can be observed in H with respect to each of the other two conditions (p=4e-4 for L-H and p=6e-3 for B-H).

Table 1. Median and median absolute deviation for each feature computed in each level. Significant values are marked with * including Bonferroni correction.

	Baseline	LOW	HIGH
mu_{RR} [ms]	846(94)	836(88)*	818(85)*
σ^2_{RR} $[ms^2]$	879(785)	802(632)	624(611)
VLF $[ms^2]$	1100(2800)*	1400(1100)	700(736)*
LF $[ms^2]$	903(1400)	619(972)	786(780)
HF $[ms^2]$	415(602)	424(324)	311(369)
LF/HF	2.45(3.30)	1.71(1.62)	2.67(2.06)
LFn	0.64(0.17)	0.56(0.14)	0.73(0.15)
HFn	0.45(0.17)	0.39(0.23)	0.33(0.14)
TOT $[ms^2]$	3400(3900)*	3500(1400)	2100(1500)*
VA [a.u.]	5.92(3.05)*	5.82(3.72)*	3.96(3.09)*
PAT [ms]	299(31)	302(37)*	291(44)*

4. Discussions

Starting from time domain features μ_{RR}, *VA* and *PAT* have shown to be very effective in discriminating the high stress level with respect to the low one. In particular, μ_{RR} which represents the RR series showed an average lower in the high stress level, so an acceleration of the heartbeat was present in the most difficult phase of the test. This is also reflected in *PAT* where there is an acceleration of the pressure wave from the heart to the periphery. This reflects the high-stress sympathetic activation that is also clearly visible in the *VA* which was significantly different as well as from the high stress phase even during baseline.

The amplitude modulation of the BVP signal represents indeed the volume of blood on the periphery. A lower BVP amplitude modulation value is therefore linked to a greater peripheral blood pressure, which is associated with vasoconstriction. Our results point at this index as a reliable marker to significantly discriminate the high effort phase from the other two conditions. With regard to the frequency domain features, significantly lower values were found in *RR VLF* and *RR TOT* in the high stress phase. The total potency is now accepted to be an indication of a greater variability of the tachogram that reflects a prevalent parasympathetic activation. This is in agreement with what was found in the time domain. *RR VLF* is still a matter of debate and warrants further elucidation. Nevertheless, several studies have associated *RR VLF* power to parasympathetic activity [14], thus in accordance with the expected response to the elicited effort levels. These two variables, however, were significantly different between high stress and baseline, not emphasizing differences in the two phases of the test. Of note, even if the other frequency domain features were not significant, the trends were still consistent. From table 1, indeed, we can see how *RR LFn* and *RR LF/HF*, which are both proportional to the sympathovagal balance, are higher during the noise-induced degradation. *RR HFn*, on the other hand, showed lower values in the high effort phase where the degradation in speech stimuli increased.

5. Conclusion

In the present work we have analyzed the ECG and the BVP signal to extract effort-sensitive features which could help in understanding when subjects experience a greater stress during hearing tests. In particular, we introduce a point process model to characterize heart beat dynamics in association with hearing effort tests. In particular, we computed the average beat-to-beat interval modelled through the point process framework, the volume amplitude of the BVP signal and the pulse arrival time computed as time difference between R-peaks and the corresponding onset values on the BVP signal. These cardiovascular features revealed highly important associations of the autonomic state of each subject in the two defined effort states in 21 normal hearing subjects. In particular statistical analysis showed that these indices are able to significantly discriminate low from high stress conditions.

Further research is needed to investigate these indexes also in hearing impaired subjects and during a wider range of listening conditions.

References

[1] Pichora-Fuller MK, Kramer SE, Eckert MA, Edwards B, Hornsby BW, Humes LE, Lemke U, Lunner T, Matthen M, Mackersie CL, et al. Hearing impairment and cognitive energy: The framework for understanding effortful listening (fuel). Ear and Hearing 2016;37:5S–27S.

[2] Peelle JE. Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. Ear and Hearing 2018;39(2):204.

[3] Zekveld AA, Kramer SE. Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. Psychophysiology 2014;51(3):277–284.

[4] Vaden KI, Kuchinsky SE, Cute SL, Ahlstrom JB, Dubno JR, Eckert MA. The cingulo-opercular network provides word-recognition benefit. Journal of Neuroscience 2013; 33(48):18979–18986.

[5] Peelle JE. Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. Ear and Hearing 2018;39(2):204.

[6] Mackersie CL, Calderon-Moultrie N. Autonomic nervous system reactivity during speech repetition tasks: Heart rate variability and skin conductance. Ear and Hearing 2016; 37:118S–125S.

[7] Francis AL, MacPherson MK, Chandrasekaran B, Alvar AM. Autonomic nervous system responses during perception of masked speech may reflect constructs other than subjective listening effort. Frontiers in Psychology 2016;7:263.

[8] Bourdillon N, Schmitt L, Yazdani S, Vesin JM, Millet GP. Minimal window duration for accurate hrv recording in athletes. Frontiers in Neuroscience 2017;11:456.

[9] Paglialonga A, Polo EM, Zanet M, Rocco G, van Waterschoot T, Barbieri R. An automated speech-in-noise test for remote testing: Development and preliminary evaluation. American Journal of Audiology 2020;29(3S):564–576.

[10] Zanet M, Polo EM, Lenatti M, van Waterschoot T, Mongelli M, Barbieri R, Paglialonga A. Evaluation of a novel speech-in-noise test for hearing screening: Classification performance and transducers' characteristics. IEEE Journal of Biomedical and Health Informatics 2021;25(12):4300–4307.

[11] McShefferty D, Whitmer WM, Akeroyd MA. The just-noticeable difference in speech-to-noise ratio. Trends in Hearing 2015;19:2331216515572316.

[12] Barbieri R, Matten EC, Alabi AA, Brown EN. A point-process model of human heartbeat intervals: new definitions of heart rate and heart rate variability. American Journal of Physiology Heart and Circulatory Physiology 2005; 288(1):H424–H435.

[13] Camm AJ, Malik M, Bigger JT, Breithardt G, Cerutti S, Cohen RJ, Coumel P, Fallen EL, Kennedy HL, Kleiger RE, et al. Heart rate variability: standards of measurement, physiological interpretation and clinical use. task force of the european society of cardiology and the north american society of pacing and electrophysiology 1996;.

[14] Taylor JA, Carr DL, Myers CW, Eckberg DL. Mechanisms underlying very-low-frequency rr-interval oscillations in humans. Circulation 1998;98(6):547–555.

Address for correspondence:

Edoardo Maria Polo
Via Ponzio 34/5, 20133 Milano, Italia
polo@diag.uniroma1.it