

ACOUSTIC IMPEDANCE CHARACTERISTICS OF ARTIFICIAL EARS FOR TELEPHONOMETRIC USE

Gaëtan Lorho, David Isherwood

Nokia Corporation, Tampere, Finland

gaetan.lorho@nokia.com, david.isherwood@nokia.com

Presented at the ITU-T workshop on "From Speech to Audio: bandwidth extension, binaural perception", 10 – 12 September 2008, Lannion, France

ABSTRACT

Artificial ears are an integral part of the audio design process for telephony devices such as mobile phones. The mechanical and electro-acoustical characteristics of these artificial ears should primarily provide an overall acoustic impedance similar to that of the average human ear over a given frequency range. This paper presents work conducted within the ITU-T Study Group 12 to quantify the degree of similarity between human ears and a subset of ITU-T Rec. P.57 Type 3 artificial ears with respect to their acoustic impedance when measured using a mobile phone-like device.

1. INTRODUCTION

An artificial ear intended for telephonometric use is designed to simulate the acoustic properties of the human outer ear when coupled with the acoustic interface of a telephony device, e.g., the earpiece of a mobile phone. Measurement with such systems can then be used to predict the audio characteristics of a device in its final usage case.

Specifications for the design and usage of such artificial ears, along with those of the head and torso simulator (HATS) to which some artificial ears can be attached, have been agreed and defined by the standardization sector of the International Telecommunications Union (ITU-T) in their Series P recommendations. Members of the ITU-T Study Group 12 (SG12 - Performance and quality of service) have been responsible for the development of recommendation P.57 [1] defining artificial ears.

Different types of artificial ears, having different design criteria and scope of application, are included in this recommendation. While the criteria for each type of artificial ear may include the device to be measured, ease of manufacture, durability, consistency of measurements and other important factors, the primary objective is to have “[...] an overall acoustic impedance similar to that of the average human ear over a given frequency band.”

For the purpose of testing the complex acoustic coupling of a modern mobile phone in the so-called ‘hand-portable’ usage mode, the Type 3.3 anatomically shaped pinna simulator and Type 3.4 simplified pinna simulator, both defined in ITU-T Rec. P.57, are both in common use within the telecommunications industry for many different measurements and metrics. The addition of an ear canal extension terminated with a pinna simulator to an IEC

60711 [2] occluded-ear simulator allows the Type 3.3 and 3.4 artificial ears to simulate the acoustic impedance characteristics of a real ear when coupled to a mobile handset dependant on its position, orientation and acoustic leakage.

Recent efforts within ITU-T Study Group 12 [3] have focused on gaining a better understanding of the performance of Type 3.3 and 3.4 artificial ears as simulators of the acoustic impedance of the average human ear, especially for frequencies above the typical narrowband range ($>3.4\text{kHz}$). These efforts are intended to steer current recommendations for narrowband and wideband artificial ear measurements as well as guiding the development of future artificial ears applicable to super-wideband or full band signals.

For this purpose a round-robin¹ measurement campaign was initiated by members of SG12 Q.5/12 (“Telephonometric methodologies for handset and headset terminals”). Measurements of acoustic impedance using a phone-like probe were made on a large set of human subjects as well as commercially available implementations of the Type 3.3 and 3.4 artificial ears integrated into a HATS. The Type 3.3 ear was provided by Brüel & Kjær and the Type 3.4 ear provided by HEAD acoustics, each manufacturer being responsible for measurements made on their respective ears.

The round-robin experiment was designed and administered by the ITU-T SG12 Q.5/12 including representatives from Nokia, Brüel & Kjær, HEAD acoustics, Motorola and Uniden. The analysis and reporting of test results was performed by Nokia, who present the results here on behalf of the Q.5/12 rapporteur and by invitation of the SG12 chairman.

This paper continues with an overview of the test design in Section 2 followed by a presentation of the measurements made on the Type 3.3 and 3.4 artificial ears in Section 3. Section 4 covers the analysis of the human ear measurements. A comparison between the set of human ear impedance measurements and the impedance measurements made on the Type 3.3 and type 3.4 artificial ears is presented in Section 5. Discussion and conclusions from this work are presented in Section 6 and Section 7 respectively.

2. TEST DESIGN

To measure the impedance of the complex acoustic system created when the ear of a user is loaded by a typical mobile phone handset, a phone-like impedance probe was provided by Brüel & Kjær. Pictures of the device to show scale are give in Figure 1. A ½ inch Brüel & Kjær Microphone Type 4134 is fitted flush to the top-front face of the device and can be used, within the frequencies to be measured, as a constant velocity source. The capillary input tube of a Brüel & Kjær Probe Microphone Type 4182 is fitted longitudinally so as to measure the pressure close to the origin of the constant velocity source. The final result of the measurement is the frequency dependent acoustic impedance $|Z_A(f)\cdot\omega|$ at the virtual ear-cap reference point. Impedance measures were recorded at each ISO R40 1/12th octave centre frequencies between 0.2-8kHz for each test case [4].

¹ A “round-robin” is defined here as a test design where measurements are pooled from multiple contributors independently performing the same methodology

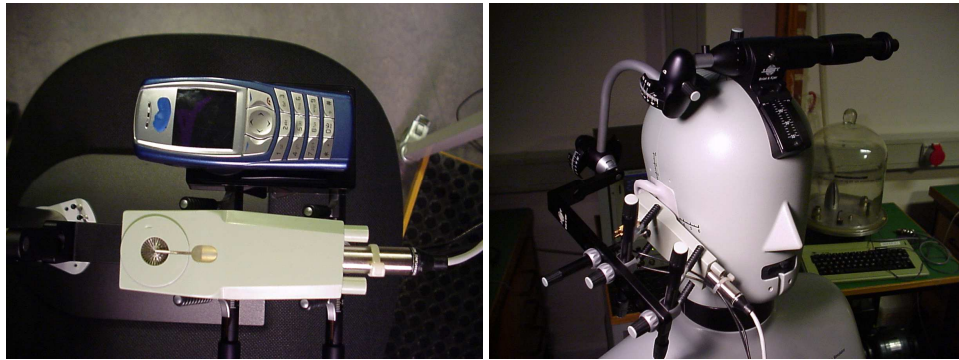


Figure 1. The mobile-phone-like impedance probe device scaled on the left with a Nokia handset and placed on a HATS in the right frame

Measurements on artificial ear types, attached to appropriate head and torso simulators (HATS) as defined in ITU-T Rec. P.58 [5], were made according to ITU-T Rec. P.57 at the standard measurement position according to ITU-T Rec. P.64 [6] for ‘hand-portable’ usage mode. These included:

1. Measurement by Brüel & Kjær on a type 3.3 right artificial ear, attached to a Type 4128D HATS, with separate test cases for application forces between 2 and 18N increasing by 2N steps,
2. Measurement by HEAD acoustics on a type 3.4 right artificial ear, attached to a HMS II HATS, with separate test cases for application forces between 2 and 18N increasing by 2N steps.

This resulted in a total of 18 individual test cases from the 2 ear types (3.3, 3.4) x 9 application forces (2N, 4N, 6N, 8N, 10N, 12N, 14N, 16N, 18N).

Measurements were also made on the ears of 60 male and 46 female human adult subjects, split between the organizations contributing to the tests. The organizational, geographical and age distribution of the human subjects are:

Contributor (Country):

- Lab #1 – Nokia (Finland) : 24 subjects
- Lab #2 – Brüel & Kjær (Denmark) : 30 subjects
- Lab #3 – HEAD acoustics (Germany) : 16 subjects
- Lab #4 – Motorola (USA) : 16 subjects
- Lab #5 – Uniden (USA) : 20 subjects

Age:

- 20-34yrs : 38 subjects
- 35-49yrs : 51 subjects
- ≥ 50 yrs : 17 subjects

Two separate measurements were made for each of the 106 human test subjects.

- ‘Normal’ application force of the handset against the users ear, inferred from placement in a quiet environment (<30dBA background noise).
- ‘Firm’ application force of the handset against the user’s ear, inferred from placement in a noisy environment (~70dBA “Hoth noise” [7] present).

In both measurement cases the user defined what application force was required. The constant velocity source of the probe produced a 1kHz sine signal immediately preceding the measurement to aid the subjects in the placement of the device on their ear, i.e., to confirm that the virtual earpiece was not occluded. For the ‘Firm’ application force test case loudspeakers within the test environment reproduced Hoth noise prior to measurement, which was turned off during measurement.

This resulted in a total of 212 individual test cases from the 106 subjects (60 male, 46 female) × 2 inferred application forces (‘normal’, ‘firm’). No repetitions of test cases for the artificial or human ears were included.

3. ARTIFICIAL EAR MEASUREMENTS

Presented in this section are the results of measurement on type 3.3 and type 3.4 artificial ears.

(Note: Although not part of the planned test comparisons, results of measurement on type 3.2LL and 3.2HL ears are supplied as references in Appendix 1)

3.1 Type 3.3 ear

The set of measurements made by Brüel & Kjær on a type 3.3 right artificial ear shown in Figure 2 includes separate test cases for application forces between 2 and 18N increasing by 2N steps.

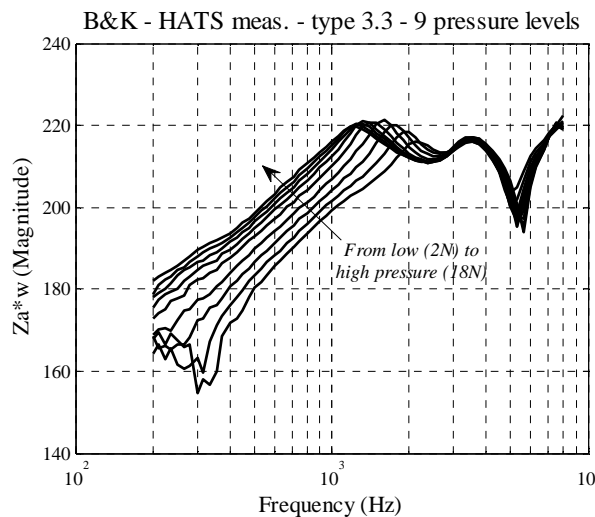


Figure 2. Measurement by Brüel & Kjær on a type 3.3 artificial ear with separate test cases for application forces between 2 and 18N increasing by 2N steps.

3.2 Type 3.4 ear

The set of measurement made by HEAD acoustics on a type 3.4 artificial ear shown in Figure 3 includes separate test cases for application forces between 2 and 18N increasing by 2N steps.

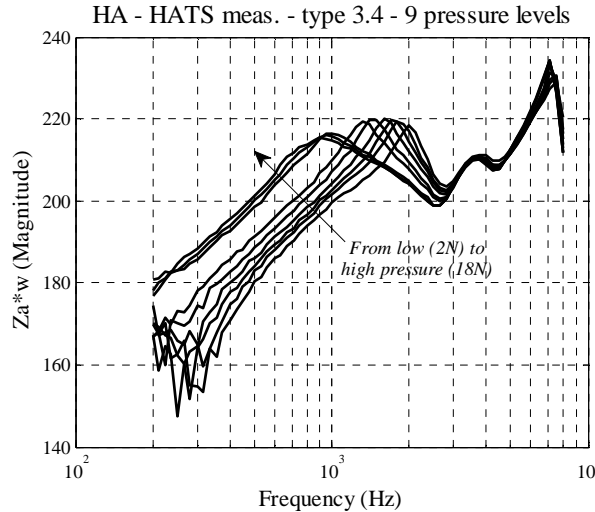


Figure 3. Measurement by HEAD acoustics on a type 3.4 artificial ear with separate test cases for application forces between 2 and 18N increasing by 2N steps

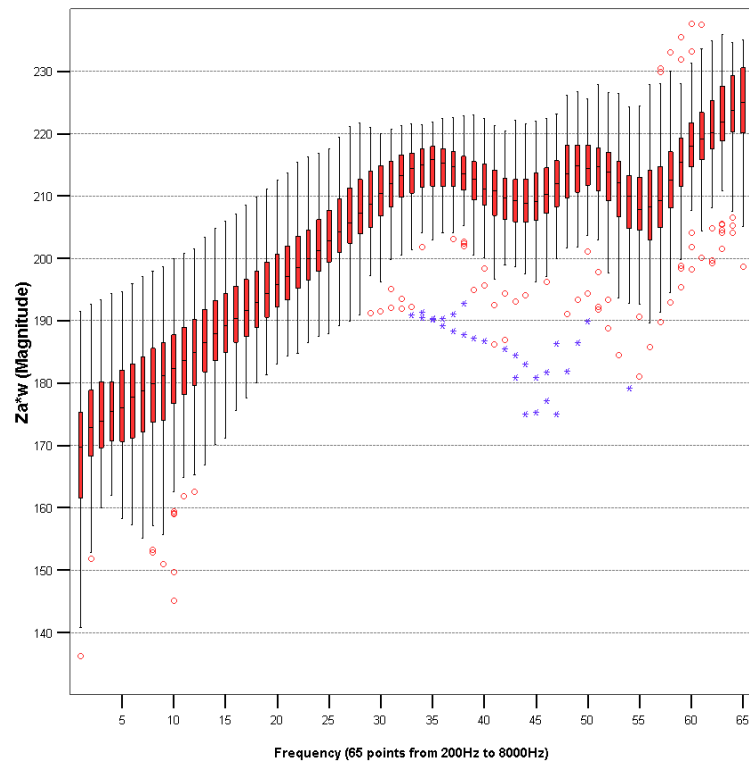
4. HUMAN EAR MEASUREMENTS

4.1 Univariate analysis of the human ear measurements

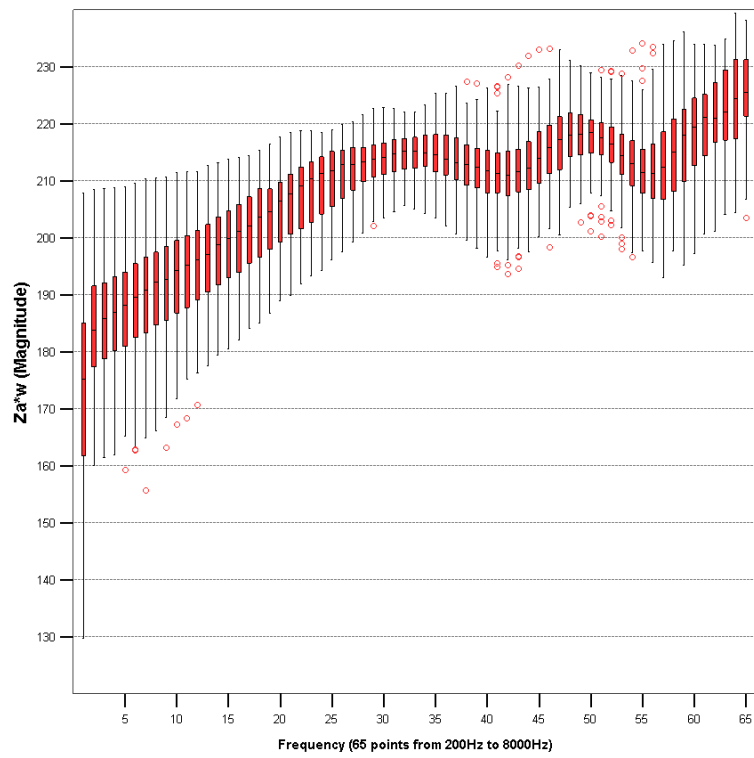
4.1.1 Descriptive statistics

Figure 4 presents a statistical summary plot of the impedance versus 1/12th octave band for the normal (a) and firm (b) application force cases separately. Each of these graphs includes a boxplot with potential outliers in red and extreme outliers in blue for each of the frequency bins individually. From Figure 4(a), we see that the normal application force case contains a set of extreme outlying points, which have been identified as originating from two individual measurements (subjects #12 and #50). The firm application force case does not include any extreme outlying point. An identification of the outlier points from both cases (see graphs in Appendix 2) highlighted that the outlier points in the two measurement sets are not due to one or several isolated measurements that would be clearly inconsistent with the general shape of this set of impedance measurements.

Figure 5 and Figure 6 present an impedance versus frequency bin line chart of the raw data (left plot) and the mean and sample standard deviation (right plot) for the normal and firm application force cases separately. The graph of the raw data for the normal application force clearly shows the two extreme outlying cases highlighted above. Note that these two subjects (#12 and #50) were removed from the analysis presented in the later sections of this document. The standard deviation of the human ear measurements represented by the grey area in the right plots of Figure 5 and Figure 6 illustrates the large variability in individual impedance at the different frequency bands.



(a)



(b)

Figure 4. Boxplot with potential outliers (red circles) and extremes (blue stars) of the impedance versus 1/12th octave band for the normal (a) and firm (b) application force cases.

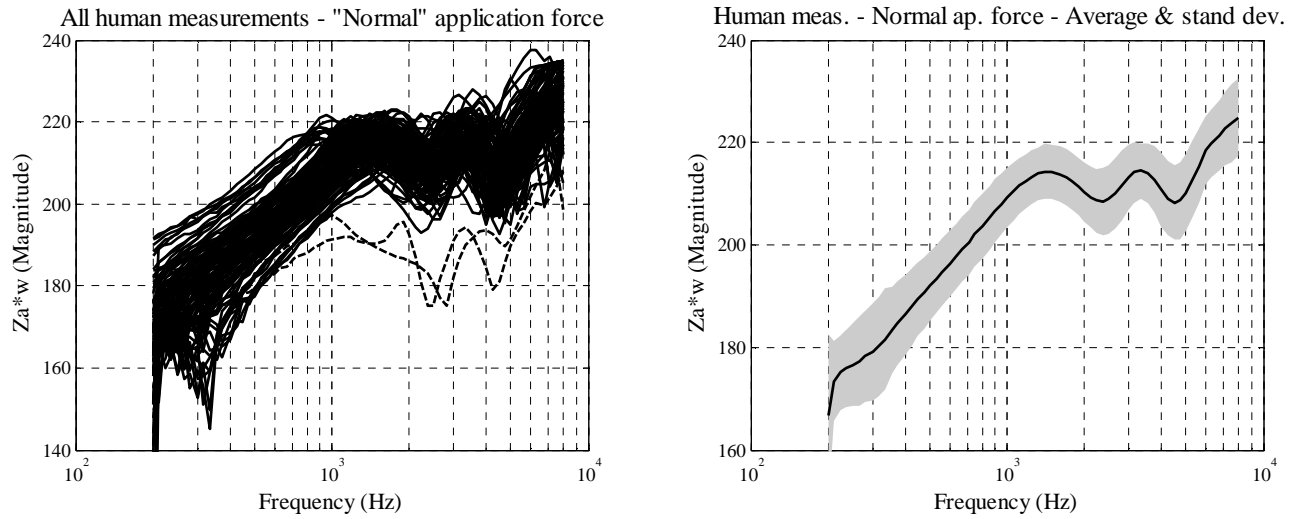


Figure 5. Impedance vs. frequency bin line chart of the raw data (left plot) and the mean and sample standard deviation (right plot) for the normal application force case.

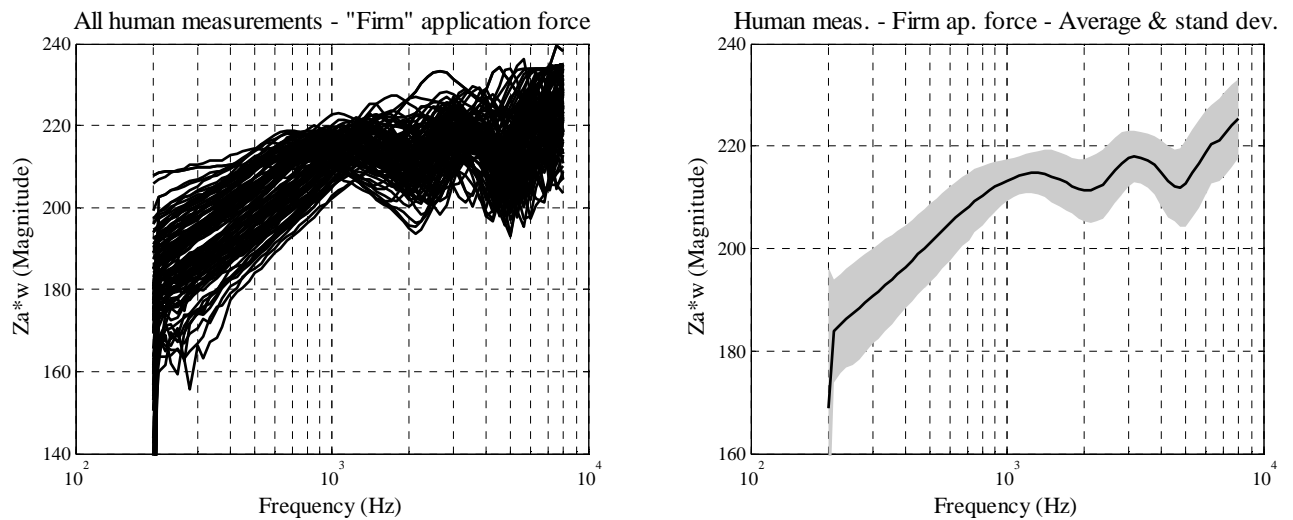


Figure 6. Impedance vs. frequency bin line chart of the raw data (left plot) and the mean and sample standard deviation (right plot) for the firm application force case.

4.1.2 Significance testing of experimental factors

An analysis of variance (ANOVA) was applied to each of 65 frequency bins separately considering the four following factors: *Lab* (five contributing organizations); *Force* (normal and firm application forces); *Subject* (104 measured individuals); *Gender* (male and female). The fact that a given individual was only measured in one laboratory and has one of the two genders has to be accounted for in the ANOVA by considering a nesting of factors. Two separate ANOVA models were considered to handle the nesting of the factor *Subject* in the factor *Lab* on one hand and the nesting of the factor *Subject* in the factor *Gender* on the other side. The first set of ANOVA models includes the factors *Lab*, *Force*, *Subject*(*Lab*) and the interaction *Lab* × *Force*. The second ANOVA model includes the factors *Gender*, *Force*, *Subject* (*Gender*) and the interaction *Gender* × *Force*. For each of these two ANOVA models, a summary table of the F-ratios and associated levels of significance for the different factors and interactions is presented for each frequency point in appendix 3. In this section, impedance verses frequency bin line charts are used to display the impedance means and associated 95% confidence intervals about these means for the different factors. Also, we will refer to the column of significant levels for each ANOVA table in appendix 3 to assess the confidence of the differences observed in these graphs.

Figure 7 illustrates the effect of the factor *Force*, which is by far the largest for most frequency bins as can be seen from the ANOVA results shown in Table 2 and Table 3 (appendix 2). The difference is clearly visible in Figure 7 for the frequency range below 1.3 kHz and the range 2.1 kHz to 5 kHz.

The factor *Lab* (the five contributing organizations) has a much smaller effect than the factor *Force* as illustrated in the left plot of Figure 8. The ANOVA results of Table 2 also highlight some significant differences for this factor in the frequency range 1.9-2.3 kHz and around the frequency 4.2 kHz. Note, however, that the F-ratios are smaller than for the factor *Force* overall. The right plot of Figure 8 illustrates the most salient example of difference observed for this factor when comparing the laboratories *Lab* #2 and *Lab* #5. An additional plot is presented in Figure 9 to compare human ear measurement means per region, i.e. between the three European laboratories (blue line) and the two North American laboratories (red line). This plot illustrates that differences between the two regions are not significant based on the overlapping 95% confidence intervals. Considering now the interaction *Lab* × *Force* in the ANOVA results shown in Table 2 (appendix 2), we can note a significant effect for the same frequency regions as those found for the factor *Force*. This effect is less important though as can be seen from the relatively small F-ratios, but it indicates that the application force used by subjects for the two cases might have differed from one laboratory to another.

Considering finally the factor *Gender*, it appears that human ear measurements made on male and female subjects do not follow the exact same pattern. This difference is illustrated in the left plot of Figure 10 and is also visible from the ANOVA results of Table 3 in appendix 2. The factor *Gender* is significant in the region 1.5-2.3 kHz and 3.5-4.5 kHz with relatively high F-ratios. We can note however from Table 3 that the interaction *Gender* × *Force* is not significant, except for a few isolated frequency bins, which indicates that the gender difference is not clearly related to a difference in application force. The right plot of Figure 10 compares human ear measurement means per gender and per application force. This graph illustrates that differences seen between the two genders follow roughly the same pattern for the normal and firm application force cases.

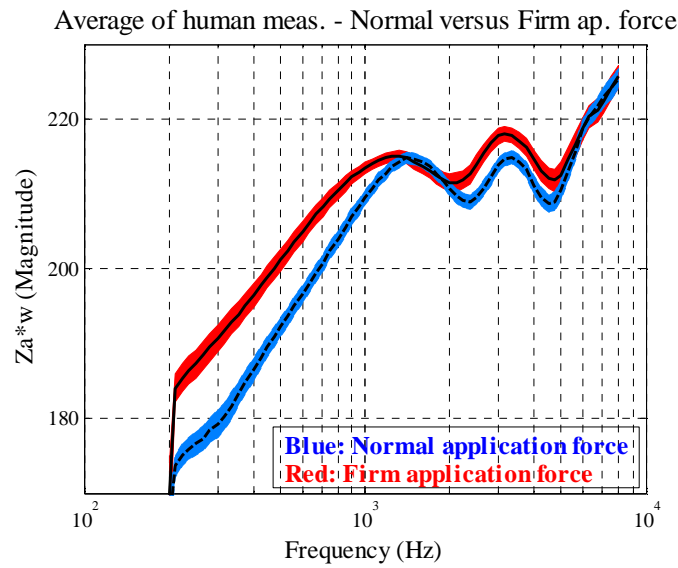


Figure 7. Comparison of human ear measurement means for the normal (blue curve) and firm (red curve) application force cases.

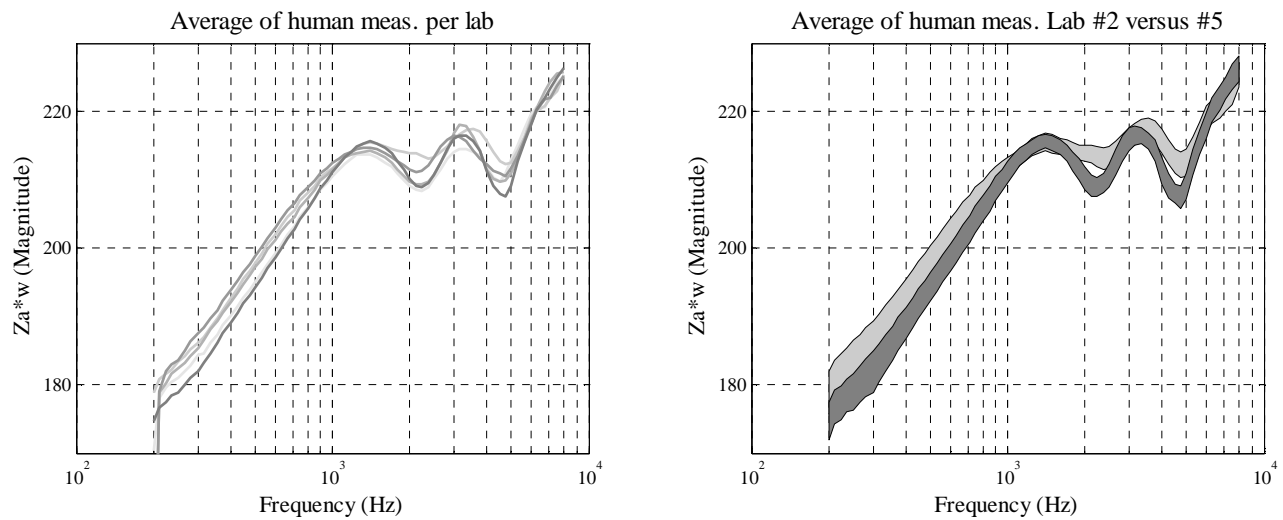


Figure 8. Comparison of human ear measurement means for the five contributing organizations (Factor Lab). The left plot illustrates the level of differences between the five laboratories. The ANOVA table shown in appendix 2 (Table 3) indicates that the Factor Lab is significant for few frequency bins, which is visible when comparing the means and 95% confidence intervals of e.g. the Labs #2 and #5, as illustrated in the right plot.

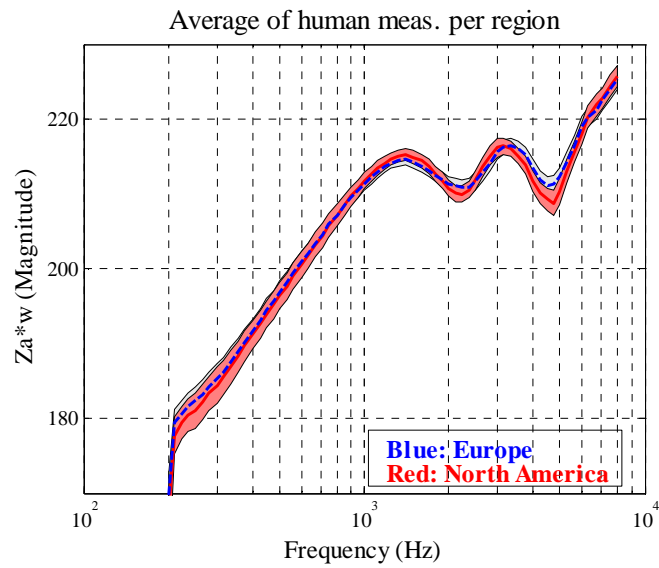


Figure 9. Comparison of human ear measurement means per region, i.e. between the three European laboratories (blue line) and the two North American laboratories (red line). This plot illustrates that differences between the two regions are not significant based on the overlapping 95% confidence intervals.

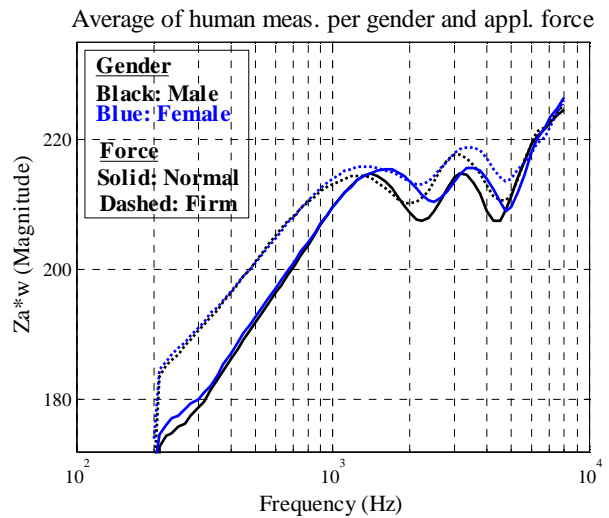
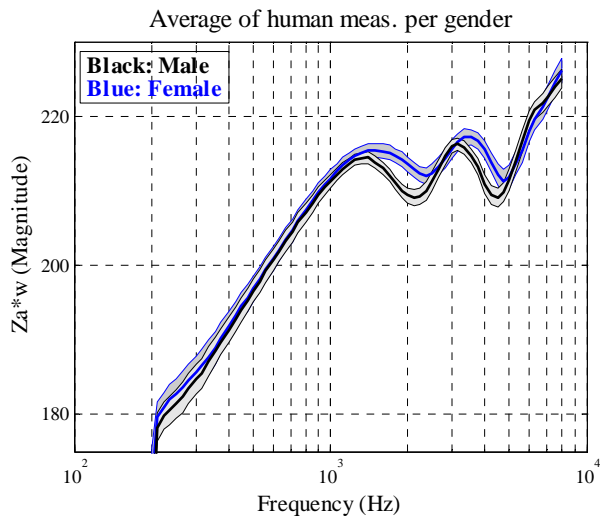


Figure 10. Comparison of human ear measurement means per gender (shown on the left plot) illustrating a significant difference between male and female measurements based on the non-overlapping 95% confidence intervals. Human ear measurement means per gender and application force (shown in the right plot) follow roughly the same pattern for the normal and firm application forces.

4.2 Bivariate parametric analysis of the human ear measurements

4.2.1 Presentation of the analysis method

An inspection of the large set of human ear measurements made in this Round Robin test reveals a common structure in the shape of the impedance response as a function of frequency. The curve formed by most of the individual impedance measurements show a series of extrema which can be used as a basis for applying a structural analysis on this dataset. For this purpose, a routine to detect curve extrema was applied to all human ear measurements and a bivariate parametric analysis was then considered to describe the variability of the impedance and frequency variables for this set of frequency response extrema.

The routine used for the detection of curve extrema consists of an identification of maxima (resp. minima) in the curve, i.e. points that are preceded and followed by lower (resp. higher) values. The number of extrema detected from the set of 104 individual measurements in each of the two cases is reported in Table 1. The automatic peak detection routine did not work 100% of the time because some curves did not follow the general shape of the dataset. Such curves lead to detected points that could be identified visually as clear outliers and were therefore removed from the dataset of extremum points. We can see from Table 1 that the first two minima and maxima cover more than 90% of the individual measurements, except for the second minimum of the firm application force case, which includes only 70% of the individual measurements. The low values seen for the third maximum relates to the fact that this maximum lies around the 6-8 kHz region. As the impedance measurement was limited to 8 kHz in the present study, any maximum occurring above 8 kHz cannot be detected in this set of measurements. Therefore the information presented here for the third maximum should be interpreted with caution.

Table 1. Number of extrema detected from the set of 104 individual measurements for the normal and firm application force cases.

Extremum index	Application force case			
	Normal		Firm	
	Maximum	Minimum	Maximum	Minimum
1	102	98	94	87
2	99	96	91	72
3	44		54	

4.2.2 Results of the parametric analysis

The resulting set of data points comprises the impedance and frequency values of each detected point and the bivariate distribution of this dataset was studied for each extrema and application force case separately. A scatter plot of the extremum points is presented in Figure 11 for the normal application force case (left plot) and the firm application force case (right plot). Three individual impedance responses are also included in this graph to illustrate how these clouds of points relate to the structural shape of the human ear impedance. To describe statistically each cloud of points, a bivariate mean was computed and an ellipse covering 95% of the data points was derived based on the Hotelling T^2 statistic. Figure 12 illustrates the resulting structural representation of the individual human ear measurements for the normal application force case (left plot) and the firm application force case (right plot). We see from this graph that the different clouds are relatively well discriminated. In Figure 13, a comparison of the structural mean and the arithmetic mean is shown for the normal application force case (left plot) and the firm application force case

(right plot). In these two plots, the size of the ellipses represents now the 95% confidence level for the mean value of the extrema, which can be compared to the 95% confidence interval of the arithmetic mean represented by the width of the blue and red curves in this figure. These two plots illustrate some differences in the characteristics of the structural and arithmetic means for both the normal and firm application force cases. The frequency of the extrema relate relatively well with the two methods, except maybe for the first maximum of the firm application force which shows a slight shift in frequency. However, the amplitude between two successive extrema (i.e. maximum impedance to minimum impedance) is about twice larger for the structural mean.

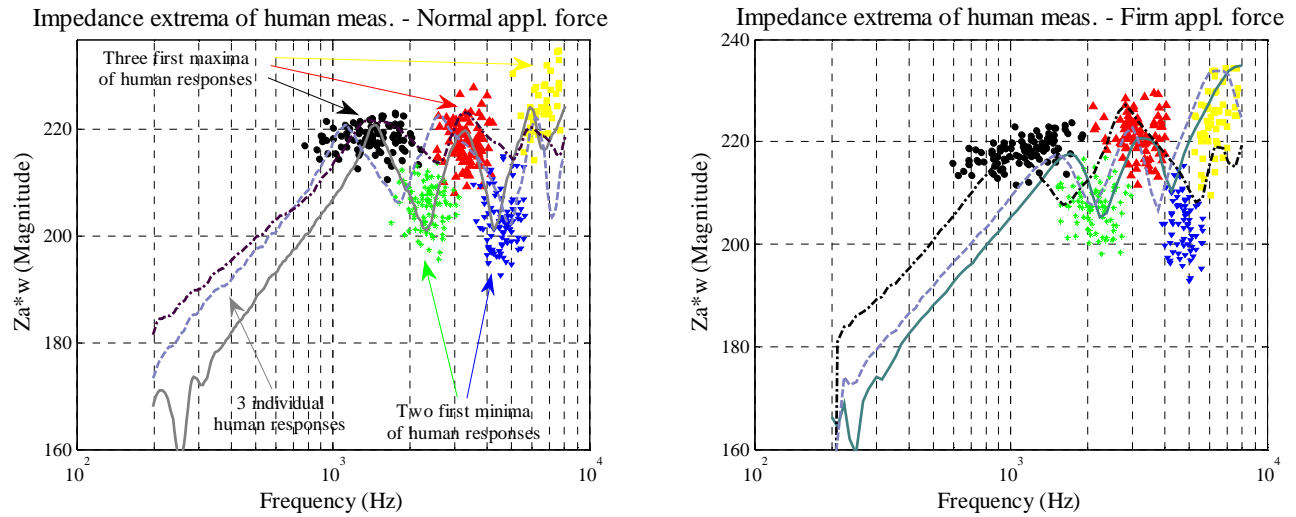


Figure 11. Scatter plot of the extremum points derived from the individual human ear impedance measurements for the normal application force case (left plot) and the firm application force case (right plot). The three individual impedance responses shown in these graphs illustrate how the clouds of points relate to the structural shape of the human ear impedance.

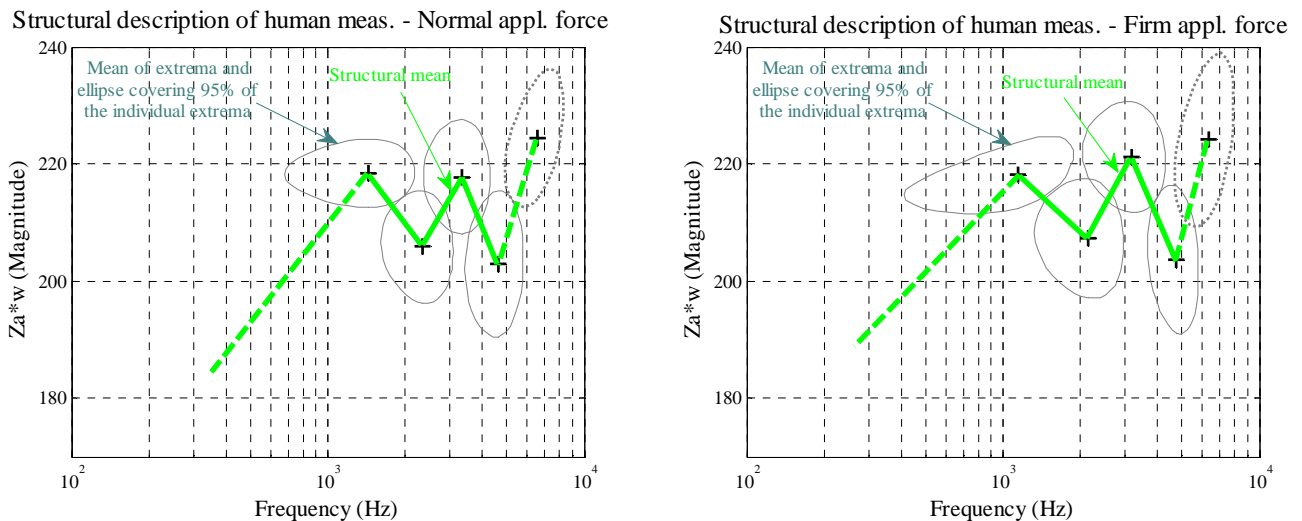


Figure 12. Structural representation of the individual human ear measurements for the normal application force case (left plot) and the firm application force case (right plot). The clouds of points shown in Figure 9 are now represented by a bivariate mean and an ellipse covering 95% of the data points.

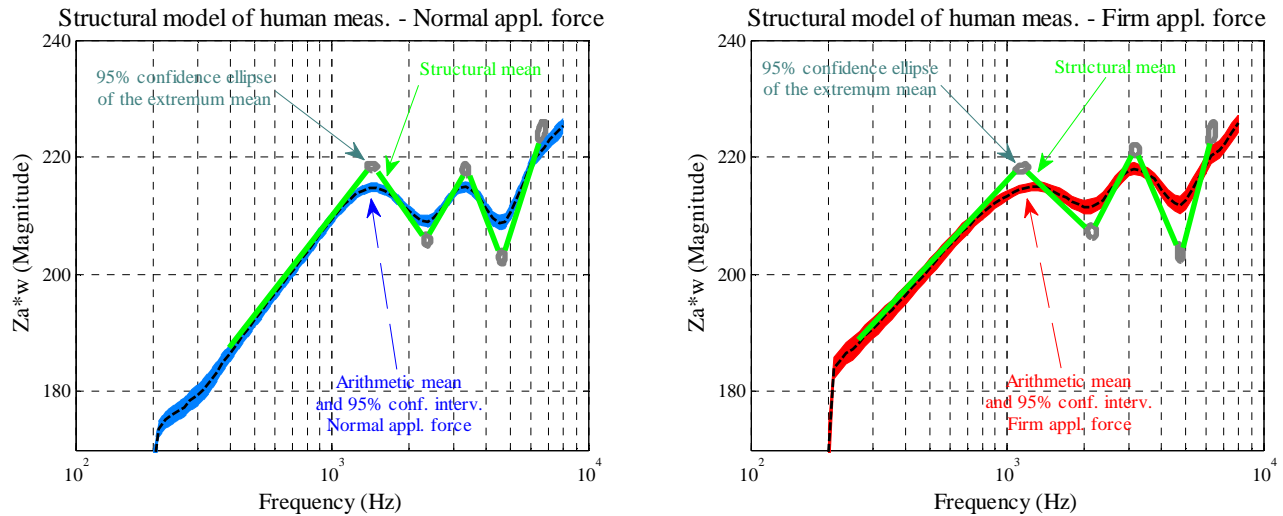


Figure 13. Comparison of the structural and arithmetic means of the individual human ear measurements for the normal application force case (left plot) and the firm application force case (right plot). The size of the ellipses represents now the 95% confidence level for the mean value of the extrema, which can be compared to the 95% confidence interval of the arithmetic mean represented by the width of the blue and red curves.

5. COMPARISON BETWEEN HUMAN AND ARTIFICIAL EAR MEASUREMENTS

5.1 Univariate comparison of human and artificial ear measurements

The series of graphs presented in this paragraph summarizes the results of the set of round-robin test measurements made in this study from a univariate point of view. The graphs presented in Figure 14 and Figure 15 compare the arithmetic means of the human ear measurements with the set of measurements made on the two artificial ear types at nine different application forces (from 2 and 18N increasing by 2N steps). Figure 14 focuses on the measurements made on the artificial ear type 3.3 while Figure 15 focuses on the type 3.4 ear. The curve shown on the left plot of each figure (blue curve) corresponds to the mean of the human ear measurements made with a normal application force and the curve shown on the right plot (red curve) corresponds to the mean of the human ear measurements made with a firm application force. The width of the red and blue curves represents the 95% confidence interval about the human ear measurement mean per frequency bin.

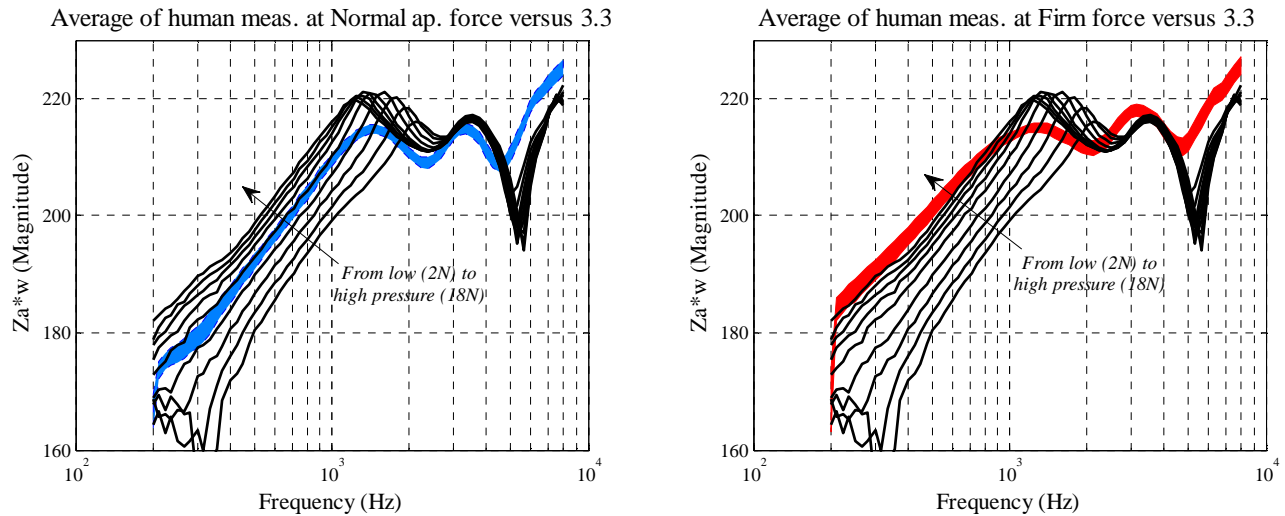


Figure 14. Comparison between the human ear measurements made with normal application force (left plot) and with firm application force (right plot) and the measurements made on the artificial ear type 3.3 at nine different application forces. The width of the red and blue curves represents the 95% confidence interval about the human ear measurement mean per frequency bin.

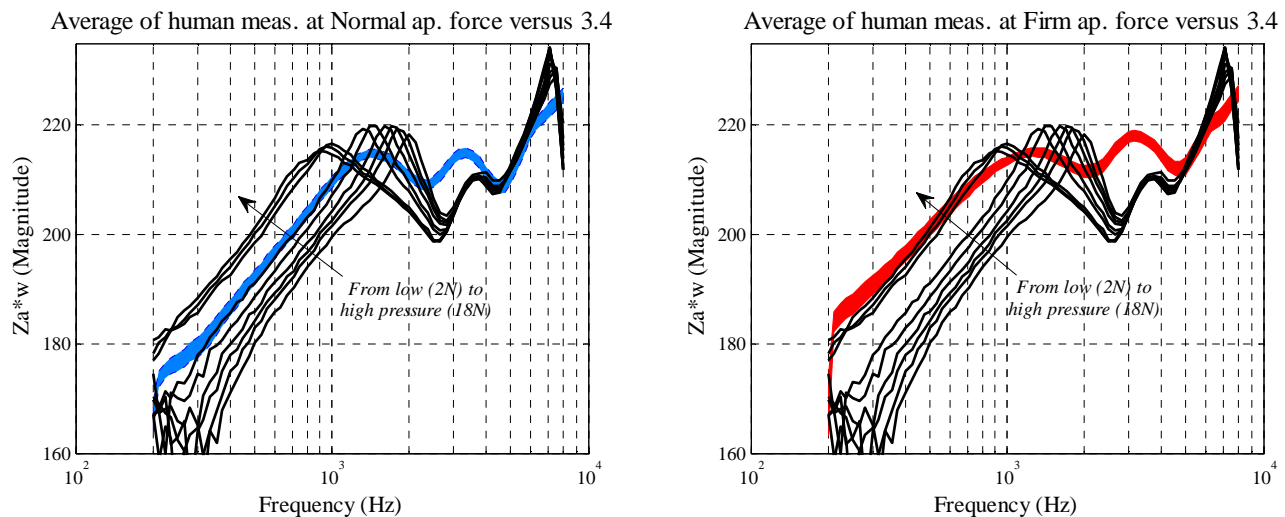


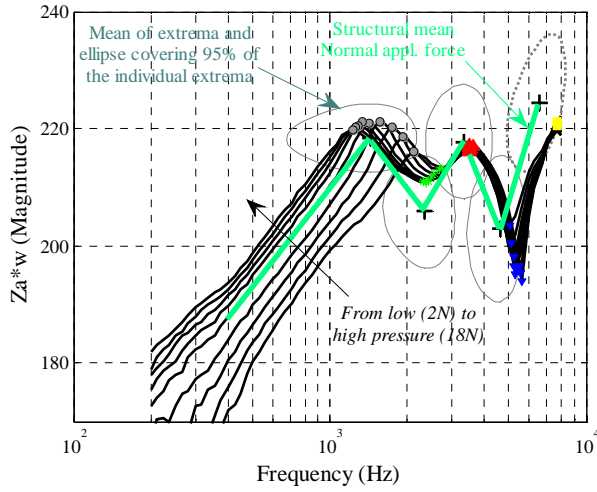
Figure 15. Comparison between the human ear measurements made with normal application force (left plot) and with firm application force (right plot) and the measurements made on the artificial ear type 3.4 at nine different application forces. The width of the red and blue curves represents the 95% confidence interval about the human ear measurement mean per frequency bin.

5.2 Bivariate parametric comparison of human and artificial ear measurements

The series of graphs presented in this paragraph summarizes the results of the set of round-robin test measurements made in this study from a bivariate structural analysis point of view. The graphs presented in Figure 16 and Figure 17 compare the structural model of the human ear measurements with the amplitude extrema of the two artificial ear types at nine different application forces (from 2 and 18N increasing by 2N steps). Figure 16 focuses on the measurements made on the artificial ear type 3.3 while Figure 17 focuses on the ear type 3.4. The structural means shown in these graphs have been described in section 4.2 but it should be noted that the ellipses presented here describe the distribution of the detected extrema, as in Figure 12, and not the 95% confidence ellipse of the mean as in Figure 13. These ellipses are better suited to visually check how well the amplitude extrema of a given

artificial ear type and application force relates to the associated distribution of individual ear impedance extrema.

Structural model of hum. meas. at Normal force versus 3.3



Structural model of hum. meas. at Firm force versus 3.3

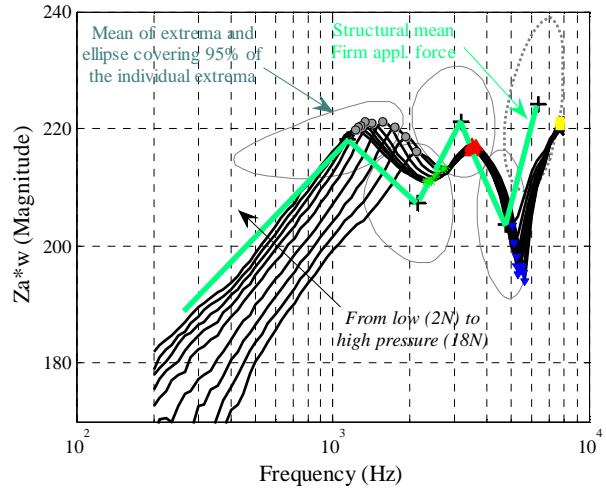
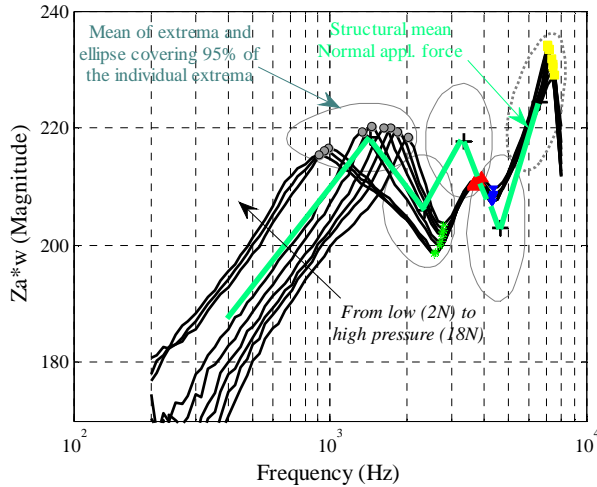


Figure 16. Comparison between the human ear measurements made with normal application force (left plot) and with firm application force (right plot) and the measurements made on the artificial ear type 3.3 at nine different application forces.

Structural model of hum. meas. at Normal force versus 3.4



Structural model of hum. meas. at Firm force versus 3.4

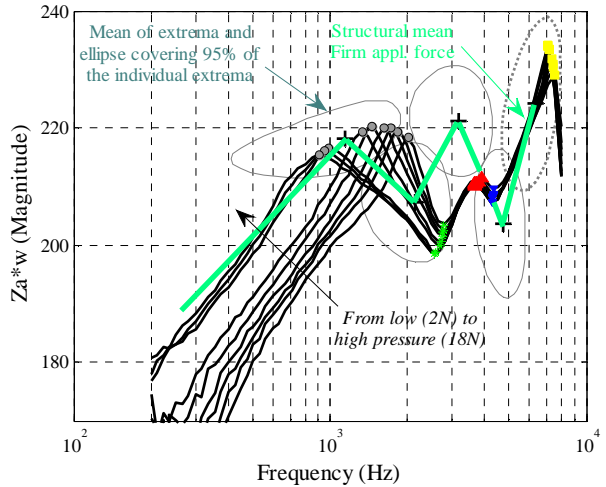


Figure 17. Comparison between the human ear measurements made with normal application force (left plot) and with firm application force (right plot) and the measurements made on the artificial ear type 3.4 at nine different application forces.

6. DISCUSSION

These results were discussed during an ad hoc meeting of the ITU-T Q.5/12, 26th May 2008, and included the contributors to the round-robin test as well as other interested members. The outcomes of these discussions, as reported by the Q.5/12 rapporteur, are presented here.

Generally very similar responses are seen on the two artificial ears up to around 2kHz. Above 2kHz pronounced deviations are observed qualitatively and quantitatively. Between 2kHz and 4kHz the first minimum of the 3.4 ear is lower than the one observed at human subjects and the 3.3 ear responses. Here the 3.3 ear correlates better the human's ear impedance . At 5kHz the second minima of the 3.3 ear is too dominant.

Between 4 and 6kHz type 3.4 seems to correlate better the human's ear impedance, and no other conclusion can be drawn for the frequency range 6kHz – 8kHz .

From the results a basic agreement was reached that for wideband (100-8kHz) it was not possible to conclude whether one of the two artificial ear impedances correlated better with humans. For narrow band (100-4kHz) a consensus was reached that the 3.3 artificial ear had a better match to human impedance responses and therefore should be the preferred type 3 in this range for an impedance point of view.

These findings are currently (May 2008) proposed to be described in an update to Rec. P.57.

7. CONCLUSIONS

The work presented in this paper highlights the challenges in accurately simulating the acoustic impedance of an average human ear for measurement of mobile phones in the 'hand-portable' usage mode at the standard position. Conclusions have been proposed within the ITU-T SG12 for describing the Type 3.3 ear as having a more similar response in narrow-band measurements. It is hoped that future worked based on these results will aid the development of ear simulation devices having greater similarity to that of the human ear for frequency ranges beyond narrow-band.

ACKNOWLEDGEMENTS

The authors wish to thank all of the companies who contributed to the work presented and to the rapporteur for ITU-T Q.5/12, Mr. Luc Madec, and the chairman of ITU-T SG12, Mr. Jean-Yves Monfort, for proposing its presentation during this workshop.

REFERENCES

- [1] ITU-T, *Recommendation P.57, Artificial ears*, International Telecommunications Union Standardization Sector (11/2005).
- [2] IEC 60711, *Occluded-ear simulator for the measurement of earphones coupled to the ear by ear inserts*, International Electrotechnical Commission (1981).
- [3] URL: <http://www.itu.int/ITU-T/studygroups/com12/index.asp> (2008)
- [4] ISO 3, *Preferred numbers - Series of preferred numbers*, International Organization for Standardization (1973).
- [5] ITU-T, *Recommendation P.58, Head and torso simulator for telephonometry*, International Telecommunications Union Standardization Sector (08/1996).
- [6] ITU-T, *Recommendation P.64, Determination of sensitivity/frequency characteristics of local telephone systems*, International Telecommunications Union Standardization Sector (09/1999).
- [7] Hoth, D.F., "Room noise spectra at subscribers' telephone locations," *Journal of the Acoustical Society of America*, vol. 12, pp. 99-504 (April 1941).

APPENDIX 1: COMPARISON BETWEEN HUMAN MEASUREMENTS AND TYPE 3.2 ARTIFICIAL EAR MEASUREMENTS

Included in this appendix are the results of the human ear analysis described herein presented with the results of equivalent measurement on a type 3.2 Low-Leak (Figure 20) and type 3.2 High-Leak (Figure 21) artificial ear. Artificial ear measurement data is supplied by Brüel & Kjør as a normative reference to the round-robin study results.

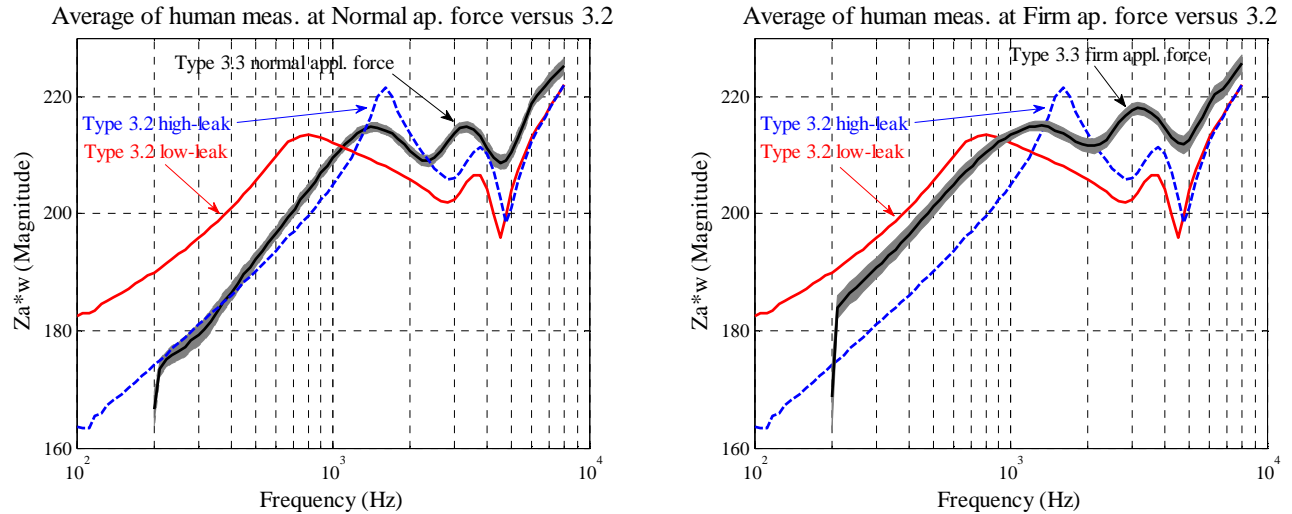


Figure 18. Comparison between the measurements made on the type 3.2 artificial ear and the human ear measurements made at normal application force (left plot) and at firm application force (right plot). The width of the gray curves represents the 95% confidence interval of the human ear measurement mean per frequency bin.

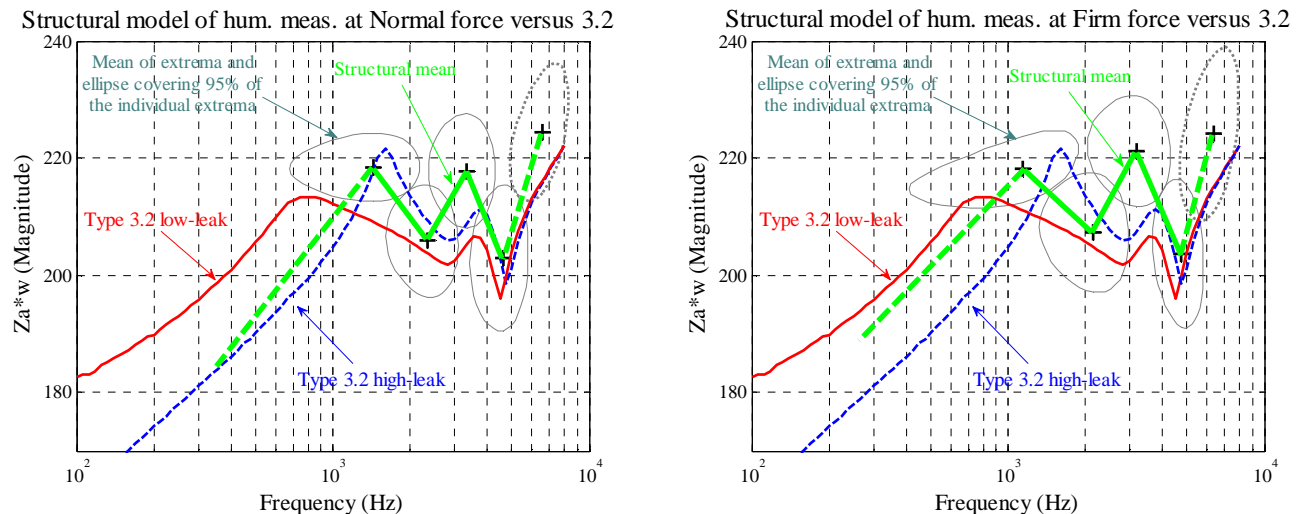


Figure 19. Comparison between the measurements made on on the type 3.2 artificial ear and the structural model derived from the human ear measurements made at normal application force (left plot) and at firm application force (right plot).

APPENDIX 2: IDENTIFICATION OF OUTLIERS FROM THE BOXPLOT ANALYSIS

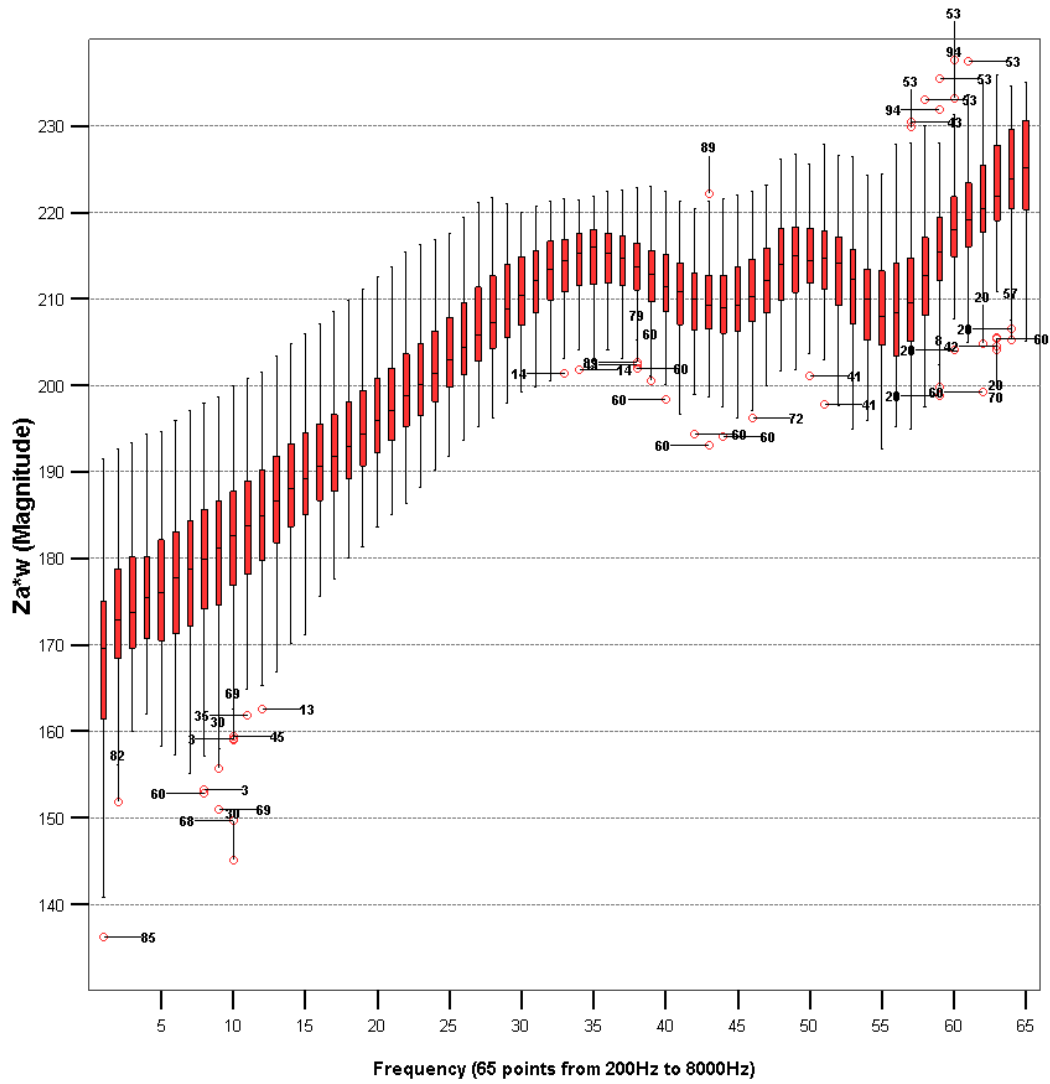


Figure 20. Human ear measurements in the normal application force case: this box plot highlights the subject index of the outlier data points for the different frequency bins.

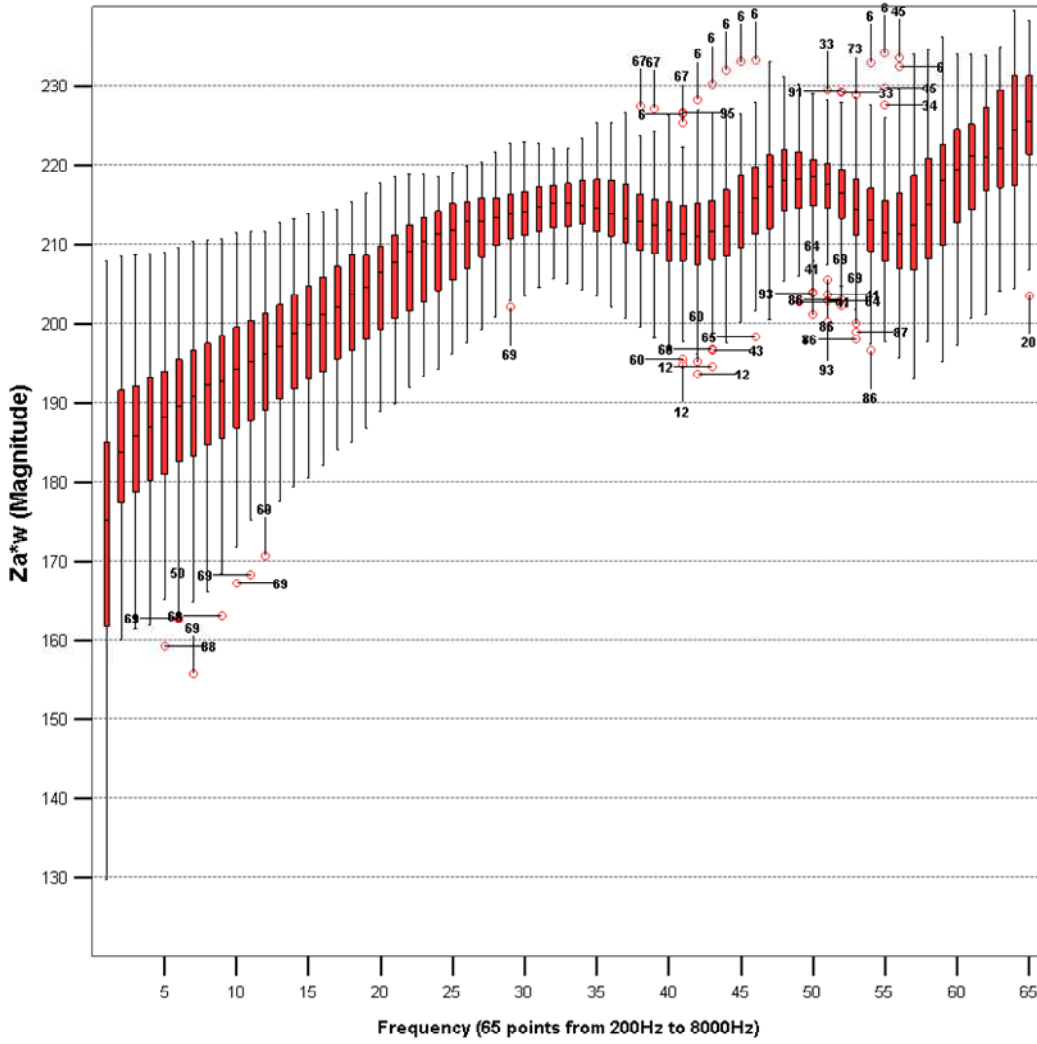


Figure 21. Human ear measurements in the firm application force case: this box plot highlights the subject index of the outlier data points for the different frequency bins.

APPENDIX 3: STUDY OF FACTOR EFFECTS BY UNIVARIATE ANALYSIS OF VARIANCE

Table 2. ANOVA applied separately to each frequency bin with the factors Lab (5 levels), Force (normal or firm), Subject (nested in Lab) and the interaction LabxForce.

Frequency	Lab		Force		Subject(Lab)		Lab*Force	
	F-ratio	Sig. Level	F-ratio	Sig. Level	F-ratio	Sig. Level	F-ratio	Sig. Level
200	14.05	***	0.07		2.42	***	9.52	***
212	0.87		194.66	***	4.14	***	6.09	***
224	1.19		252.54	***	5.48	***	7.37	***
236	1.09		262.33	***	5.72	***	7.04	***
250	1.17		266.93	***	6.13	***	8.42	***
265	1		292.44	***	6.54	***	8.07	***
280	1.16		289.59	***	6.77	***	7.08	***
300	1.15		312.74	***	7.06	***	6.27	***
315	1.03		318.61	***	7.5	***	6.8	***
335	1		271.04	***	6.53	***	5.53	***
355	1.19		303.78	***	6.76	***	6.13	***
375	1.27		306.91	***	6.9	***	6.67	***
400	1.44		314.48	***	6.68	***	6.72	***
425	1.51		307.64	***	6.61	***	6.08	***
450	1.57		313.85	***	6.63	***	6.74	***
475	1.64		311.82	***	6.55	***	6.53	***
500	1.74		313.5	***	6.53	***	6.32	***
530	1.74		309.54	***	6.48	***	6.22	***
560	1.79		295.51	***	6.32	***	5.58	***
600	1.82		280.43	***	6.02	***	5.04	***
630	1.79		276.86	***	5.92	***	4.46	**
670	1.69		253.77	***	5.73	***	3.67	**
710	1.56		230.31	***	5.44	***	3.34	*
750	1.45		204.23	***	5.08	***	2.78	*
800	1.26		168.08	***	4.46	***	2.02	
850	1.11		131.66	***	3.84	***	1.54	
900	0.98		98.74	***	3.24	***	1.17	
950	0.77		75.06	***	2.79	***	1.05	
1000	0.67		53.26	***	2.3	***	0.97	
1060	0.53		33.56	***	1.75	**	0.86	
1120	0.5		20.49	***	1.43	*	1.04	
1180	0.46		11.22	**	1.22		1.03	
1250	0.72		4.26	*	1.1		1.12	
1320	1.14		0.85		1.14		1.03	
1400	1.37		0.06		1.36		1.02	
1500	1.33		1.4		2.02	***	0.95	
1600	1.35		3.24		3.1	***	1.04	
1700	1.5		3.32		4.06	***	1.1	
1800	1.94		1.62		4.56	***	1.06	
1900	2.92	*	0.01		4.57	***	0.91	
2000	4.19	**	2.72		4.27	***	0.6	
2120	5.13	***	15.03	***	4.14	***	0.66	
2240	4.52	**	37.25	***	4.19	***	1.71	
2360	3.03	*	61.49	***	4.4	***	2.47	*
2500	1.9		90.8	***	4.99	***	3.08	*
2650	1.16		112.9	***	5.69	***	3.46	*
2800	0.81		110.26	***	5.61	***	3.92	**
3000	1		93.99	***	5.3	***	3.97	**
3150	1.74		73.31	***	4.75	***	3.08	*
3350	2.16		54.39	***	4.87	***	2.87	*
3550	2.16		52.21	***	6.13	***	3	*
3750	2.09		59.8	***	7.4	***	3.33	*
4000	2.45		59.15	***	6.77	***	4.02	**
4250	2.67	*	40.23	***	4.37	***	3.25	*
4500	2.46		24.17	***	2.62	***	2.6	*
4750	1.92		12.35	***	2.39	***	2.91	*
5000	1.15		5.11	*	2.24	***	1.83	
5300	0.44		2.61		2.51	***	0.66	
5600	0.26		0.97		2.65	***	0.4	
6000	0.24		0.01		2.86	***	0.47	
6300	0.05		0.28		2.69	***	0.71	
6700	0.41		0.1		3.05	***	1.07	
7100	0.42		0.29		3.5	***	1.01	
7500	0.49		0.04		3.55	***	1.02	
8000	0.11		0		3.64	***	1.36	

Significant levels are represented as follows: $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$

Table 3. ANOVA applied separately to each frequency bin with the factors Gender (male or female), Force (normal or firm), Subject (nested in Gender) and the interaction Gender×Force.

Frequency	Gender		Force		Subject(Gender)		Gender*Force	
	F-ratio	Sig. Level	F-ratio	Sig. Level	F-ratio	Sig. Level	F-ratio	Sig. Level
200	3.11		1.12		2.7	***	1.39	
212	0.83		154.63	***	3.43	***	0.37	
224	0.5		197.02	***	4.42	***	0.4	
236	0.88		212.18	***	4.66	***	1.02	
250	0.62		203.99	***	4.77	***	0.25	
265	0.71		227.42	***	5.17	***	1.54	
280	0.47		233.91	***	5.52	***	0.66	
300	0.28		261.01	***	5.9	***	0.31	
315	0.33		264.12	***	6.13	***	0.31	
335	0.12		237.12	***	5.55	***	0.01	
355	0.1		258.66	***	5.68	***	0.01	
375	0.13		256.6	***	5.72	***	0.23	
400	0.15		259.98	***	5.58	***	0.45	
425	0.12		259.59	***	5.64	***	0.39	
450	0.13		256.59	***	5.56	***	0.47	
475	0.07		256.78	***	5.56	***	0.62	
500	0.05		260.17	***	5.61	***	0.81	
530	0.05		256.53	***	5.61	***	1.2	
560	0.04		250.63	***	5.58	***	1.01	
600	0.01		242.4	***	5.43	***	1.13	
630	0.02		244.32	***	5.43	***	0.93	
670	0.04		231.23	***	5.37	***	0.85	
710	0.04		211.97	***	5.13	***	0.68	
750	0.06		191.94	***	4.84	***	0.31	
800	0.09		163.45	***	4.33	***	0.07	
850	0.12		131.71	***	3.77	***	0	
900	0.11		102.02	***	3.21	***	0.13	
950	0.15		80.16	***	2.76	***	0.43	
1000	0.22		58.52	***	2.29	***	0.78	
1060	0.42		38.46	***	1.73	**	1.13	
1120	0.56		24.14	***	1.41	*	1.33	
1180	0.78		14.14	***	1.19		1.21	
1250	1.07		6.24	*	1.08		1.57	
1320	1.19		1.94		1.16		1.97	
1400	1.93		0.08		1.38		2.41	
1500	4.29	*	0.42		2.01	***	2.1	
1600	8.97	**	1.68		2.91	***	0.95	
1700	14.81	***	1.93		3.61	***	0.15	
1800	20.78	***	0.94		3.92	***	0.07	
1900	23.28	***	0.01		4.05	***	1.01	
2000	21.04	***	2.67		4.16	***	2.96	
2120	15.19	***	13.44	***	4.43	***	4.69	*
2240	8.93	**	31.54	***	4.43	***	4.09	*
2360	3.93	*	52.23	***	4.41	***	2.08	
2500	1.06		79.06	***	4.75	***	0.53	
2650	0.11		99.94	***	5.22	***	0.01	
2800	0.06		96.8	***	5.01	***	0.1	
3000	0.13		84.83	***	4.78	***	0.89	
3150	0.05		71.82	***	4.58	***	1.58	
3350	2.7		55.36	***	4.65	***	0.57	
3550	7.42	**	52.43	***	5.55	***	0.13	
3750	11.98	***	55.3	***	6.33	***	0.02	
4000	16.43	***	48.86	***	5.54	***	0.64	
4250	15.9	***	34.94	***	3.71	***	0.26	
4500	8.05	**	23.84	***	2.42	***	0.26	
4750	1.46		14.56	***	2.35	***	3.54	
5000	0.09		7.59	**	2.32	***	6.23	*
5300	0.08		4.64	*	2.61	***	5.2	*
5600	0.91		1.48		2.65	***	1.42	
6000	1.14		0.01		2.81	***	0.08	
6300	1		0.03		2.62	***	0.98	
6700	0.22		0.59		3.02	***	1.61	
7100	0		0.72		3.46	***	1.2	
7500	0.19		0		3.49	***	0.67	
8000	0.83		0.18		3.46	***	0.75	

Significant levels are represented as follows: $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$