



HAL
open science

Effect of headphone position on absolute threshold measurements

Mathieu Paquier, Vincent Koehl, Brice Jantzem

► **To cite this version:**

Mathieu Paquier, Vincent Koehl, Brice Jantzem. Effect of headphone position on absolute threshold measurements. *Applied Acoustics*, 2016, 105, pp.179-185. 10.1016/j.apacoust.2015.12.003 . hal-01249905

HAL Id: hal-01249905

<https://hal.science/hal-01249905>

Submitted on 4 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EFFECT OF HEADPHONE POSITION ON ABSOLUTE THRESHOLD MEASUREMENTS

Mathieu Paquier¹, Vincent Koehl¹, Brice Jantzen²

¹ University of Brest, Lab-STICC CNRS UMR 6285, 6 avenue Le Gorgeu - 29238 Brest, France.

Mathieu.Paquier@univ-brest.fr

² Centre Auditif Entendre, 4 square Commandant L'Herminier, 29200 Brest, France

Pure-tone audiometry (measurement of absolute thresholds using pure tones) is the main test for the diagnosis of hearing loss. The aim of the present study is to determine whether the headphone placement over a listener's ears has an influence on pure-tone audiometric tests, for a large frequency range, for Sennheiser HD600 and Telephonics TDH39 headphones. Audiograms (with 1 dB step, and including 10 frequencies up to 14kHz) were performed several times on normal-hearing subjects, for different –or not different- headphone positions (allowing to dissociate between effects of headphone position and cognitive factors). Globally, the results seem to indicate that the reliability without headphone removing was quite close to the one observed with removing. The influence of removing did not appear more crucial for high frequencies. The rare frequencies for which a removing effect was seen seem to be function of the headphone model. Finally the results were quite different among the subjects.

Keywords : audiometry, absolute thresholds, headphone position

1. INTRODUCTION

1.1. Reliability of hearing threshold measurements

Pure-tone audiometry consists in measuring absolute hearing thresholds by using pure tones and is used as a primary diagnosis of hearing loss. It can be accomplished by measuring the Minimum Audible Field (MAF) for stimuli presented with loudspeakers or the Minimum Audible Pressure (MAP) for stimuli presented with headphones. In the latter case, the headphone model has to be chosen with care.

Several studies have compared the reliability of hearing thresholds as a function of the headphone type [1-7]. In these studies (which used different step precisions, from 1 to 5 dB), reliability was observed with a test-retest method: subjects' thresholds were measured, then the headphone was removed, then replaced, and new thresholds were measured. It has been shown that supra-aural headphones were less reliable than circumaural headphones [4]. Several studies reported supra-aural headphones to be less reliable than insert earphones [6, 7], while other studies showed the same variability with supra-aural and insert earphone [2, 3]. In addition, some studies showed that the cushions themselves could have an influence, independently of the headphone model and type [8, 9, 10].

1.2. A high frequency specific issue?

Some studies showed that the auditory thresholds were less reliable in high frequencies, and measurements of acoustic level in subjects' ears even showed that above 8 kHz, standing wave patterns create large variations at different points within the ear canal, and that a specific high-frequency audiometer was required, including a feedback in-ear microphone [11-14].

On the contrary, [15] indicated that the intra-subject reliability of threshold estimates should be nearly the same at both the low (0–8 kHz) and the high (8–16 kHz) frequency regions. Authors explained that conventional headphones reduce the size of the standing wave ratios compared with a hard wall termination. An additional factor was the steeper slope of the psychometric function at the higher frequencies. The combination of these factors produced a standard deviation for threshold estimates that was only about 1 dB larger at the high than at the low frequency region.

1.3. Two factors explaining the global reliability

Actually, there are at least two explanations for limitation of threshold reliability [11]:
i) the “threshold variance”, determined by the number of trials, the slope of the underlying psychometric function (possibly different in low and high frequencies [15]), the psychophysical procedure, and the consistency of the subject's attention and criterion. This variance is between 1 and 3 dB according to previous studies [16, 17] using 1-dB step-size audiometry.

ii) the “fitting variance” concerns differences in earphone placement. At high frequencies, earphone placement or fitting may alter the effective geometry of the ear canal for a given

subject, resulting in large changes in effective SPL. This fitting variance is the consequence of the scattering of HPTF with headphone position: the HeadPhone Transfer Function (HPTF) describes both the headphone response and the coupling to a listener's ear. It has been shown that slight modifications in the headphone placement can result in large spectral (>9dB) differences, especially in high frequency [18-22]. Moreover the modification of timbre introduced by these differences were audible with pink noise and music, for several different headphone models [23].

1.4. Variability specifically due to headphone positioning

Most of test-retest (with systematic removing of headphone) experiments cited above have mixed the two contributions "threshold variance" and "fitting variance". Indeed, when only two audiograms are realized, the effect of headphone repositioning is merged with cognitive factors, procedure validity, etc.

In order to separate between "threshold variance" and "fitting variance", three audiograms have to be measured: two with the same headphone position (the difference between these measurements involves only "threshold variance"), and a third with an other headphone position (the difference between this third measurement and the two other ones involves both "threshold variance" and "fitting variance").

Few studies have separated the two contributions by repeating the threshold measurement both with and without headphone replacement. Two of them, using insert earphones, found an effect of earphone replacement [11, 24]. Hickling [5] showed that the removal and replacement of TDH39 supra-aural headphone significantly reduced the reliability of 6 and 8 kHz thresholds in comparison to repeated threshold measurements for the same headphone position (but no effect at 1 and 2 kHz). However, only these four frequencies were under test. Erlandsson et al. [7] and Gauz et al. [25] both compared auditory thresholds obtained when a circumaural headphone was repeatedly replaced to thresholds obtained when the headphone position was fixed. Erlandsson et al. [7] reported a significant effect of circum-aural headphone replacement on threshold measurements but Gauz et al. [25] reported no such significant effect.

1.5. Summary

From all these previous works it can be seen that :

- Studies exhibited large differences in their measurement methods and most of them consider a 5 or 10dB precision. As a consequence, results are often contradictory and do not enable to really know if circumaural headphones are more or less reliable than supra-aural headphones.
- Most of the studies used only one retest (with headphone repositioning), and did not enable to draw a global conclusion about the effect of the headphone placement independently of the “threshold variance”.
- Most of the studies tested frequencies up to 8kHz (however few papers tend to indicate that positioning reliability would be more problematic in higher frequencies, but counter-arguments might be found among this literature [15]).
- In addition, the circumaural headphone HD600 is nowadays often used in psychoacoustics experiments and a recent study [23] has shown that modifications of the position of this headphone led to audible changes. However no study has specifically studied the threshold reliability with this headphone.

Taking all these considerations into account, the present study aimed at measuring auditory thresholds with high precision (1 dB step) on a large frequency range (from 125 Hz up to 14 kHz) using both Sennheiser HD600 (circum-aural) and Telephonics TDH39 (supra-aural) headphones, with and without headphone repositioning. Threshold variability was then investigated by comparing the audiograms obtained on both identical and different headphone positions, and enabled to separate between “threshold variance” and “fitting variance”.

2. MATERIAL AND METHOD

2.1. Audiometry procedure

A modified ANSI procedure [26] was used. The audiometry included measurements at octave intervals from 250 Hz through 8 kHz, and additional thresholds were tested at 125 Hz, 6 kHz, 11 kHz, and 14 kHz. We used the octave intervals from 250 Hz through 8 kHz because they are used in clinical audiometry (according to ANSI). Several frequencies were added to expand the measurement range towards the low and high frequencies (125 Hz, 11 and 14 kHz respectively). Threshold exploration was carried out by presenting pulsed tones lasting 2.5 s.

The level of successive presentations was determined by the preceding response: in the original ANSI recommendation [26], each failure to respond to a signal leads to successive increases of 5 dB steps until the first response occurs. The intensity is then decreased by 10 dB and another ascending series is begun. In this study, a higher precision was sought. The algorithm consisted then of a series of three bracketing procedures, each providing progressively smaller step sizes to finally result in threshold responses with 1 dB resolution. The initial bracketing series (Series 1) used step increments of up 5 dB, down 10 dB to quickly bracket the threshold level to within 10dB. Subsequent bracketing series used step sizes of up 2 dB, down 5 dB (Series 2) and up 1 dB, and down 2 dB (Series 3). Final threshold was defined in Series 3, as the lowest hearing level at which responses occur in at least one-half of a series of ascending trials, with a minimum of two responses out of three required at a single level.

2.2. Participants

Twenty normal-hearing subjects (aged 20-40 years; 7 females and 13 males; mean age = 25.4 yr; SD =6.3 yr), who were all unpracticed in hearing experiments, participated in the study. Listeners were required to have hearing thresholds ≤ 20 dBHL in the 125 Hz to 8 kHz frequency range.

2.3. Auditory Listening test

Threshold measurements were made only on the subject's left ear. According to ANSI recommendation [26], the headphone was centered over the ears and its position was adjusted by test subject for most comfortable listening. It is worth noting that some other standards about audiometry [27, 28] recommend the tester to adjust himself the headphone placement. However, studies about HPTF measurements showed that a better reproducibility can be obtained when the headphone was placed by the subject himself [29, 30].

Subject sat in front of a computer screen in an audiometric booth. The automatic procedure was run using a MATLAB graphical user interface controlled by a PC located out of the booth. Stimuli were played back over the test headphone (Sennheiser HD600 or Telephonics TDH39) by using a RME Fireface 800 soundcard.

During the playback, the instruction "Listen" was displayed on the screen. After the stimulus presentation, the subject was asked whether he had perceived the sound or not.

After a test (measurement of thresholds for each of the ten frequencies), two retests were conducted (with and without headphone repositioning, enabling to separate between

“threshold variance” and “fitting variance”). Between the first and the second test or between the second and the third test (randomized), listener received the instruction (by displaying a message on the screen) to remove and replace the headphone on his head. No pause was allowed. The whole test for one given headphone model lasted between 45 and 60 minutes. The whole test was first carried out with one of the two headphones, then repeated with the other headphone one week later.

3. RESULTS

3.1. Calculation of threshold differences

Let's denote the three consecutive audiograms A, B, and C. Three differences were calculated: B-A, C-B, C-A. If the headphone was repositioned between the first (A) and the second audiogram (B), the two differences $|B-A|$ and $|C-A|$ were obtained from different headphone positions (and respectively with consecutive and non-consecutive measurements), and the difference $|C-B|$ was obtained from identical headphone positions. If the headphone was repositioned between the second (B) and the third audiogram (C), *identical positions* applies to $|B-A|$, and *different positions* to $|C-B|$ and to $|C-A|$ (respectively with consecutive and non-consecutive measurements). Let us remind that the differences from *identical position* involves only “threshold variance”, and the differences from *different position* involves both “threshold variance” and “fitting variance”. As a consequence, if the difference from *different positions* is significantly larger than the difference from *identical positions*, it means that the headphone repositioning affects the audiogram reliability.

Figure 1 (for HD600) and Figure 2 (for TDH39) show the absolute value of the differences between *different* or *identical* positions, for each frequency and for all subjects. Note that for *different positions*, the two differences for consecutive and non-consecutive measurements ($|B-A|$ and $|C-A|$, or $|C-B|$ and $|C-A|$) were merged. As the distribution of differences did not fit a normal distribution according to a Kolmogorov –Smirnov test, means and standard deviations were not suitable, so boxplots were displayed instead: on each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, and the crosses indicate the subjects out of this range.

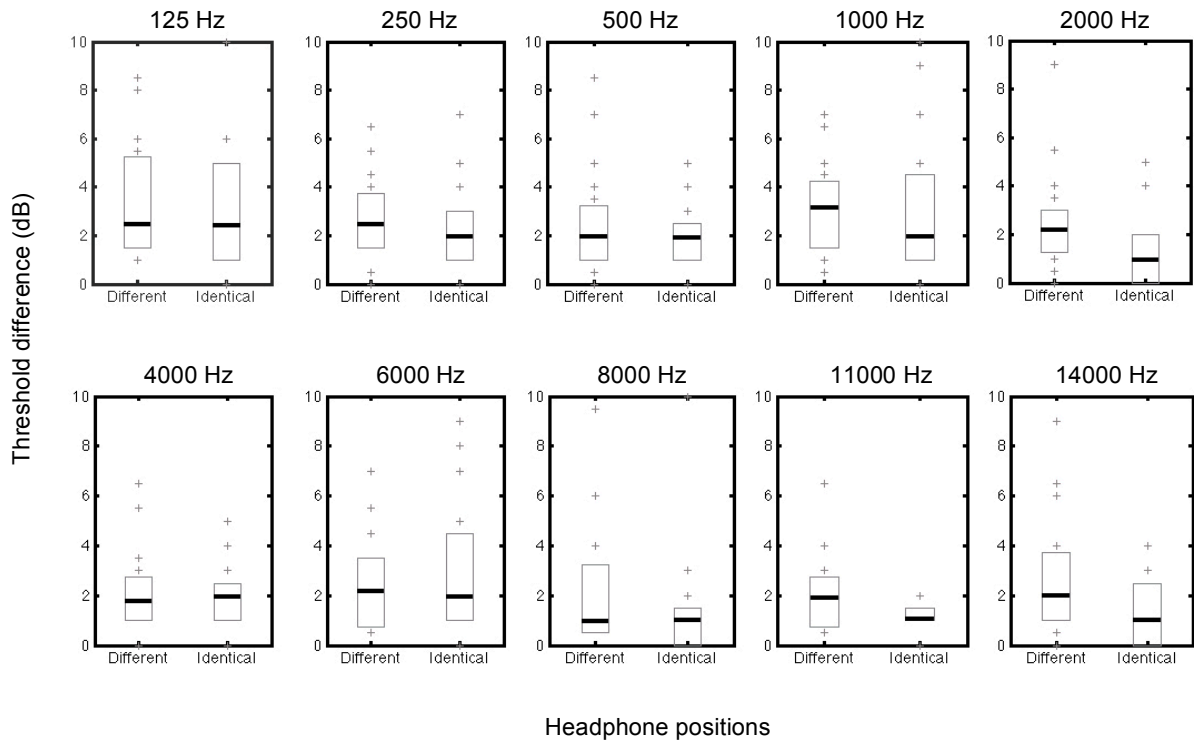


Figure 1. HD600 headphone - Boxplots of the differences between thresholds from different vs identical headphone positions, for the ten frequencies under test.

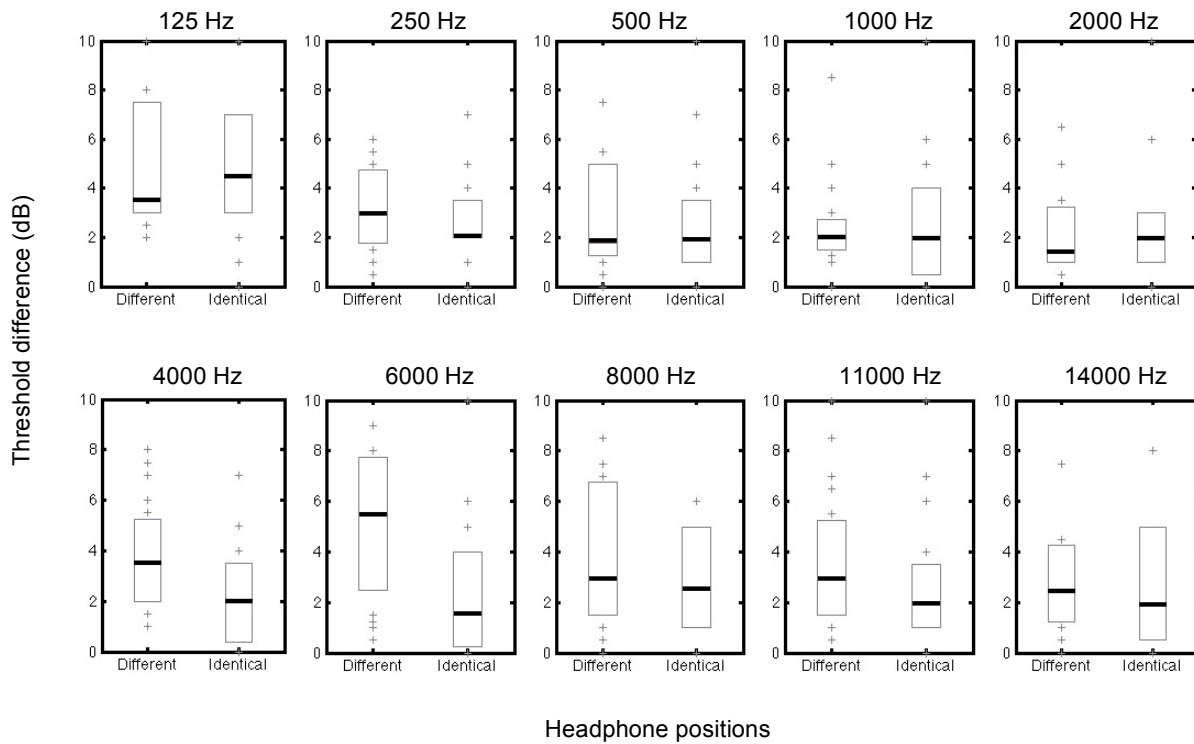


Figure 2. TDH39 headphone - Boxplots of the differences between thresholds from different vs identical headphone positions, for the ten frequencies under test.

3.2. Order effects

During the test, practice could progressively decrease the thresholds of subjects [31], or conversely tiredness could progressively increase the thresholds. As a result, this potential order effect could hide an effect of the headphone removing. A Friedman test carried out on the thresholds (not the differences) indicated no significant effect of the measurement order: no effect of practice nor tiredness was observed.

3.3. Effect of headphone position

Firstly, let's note that the differences between thresholds both with and without headphone repositioning provided results of the same order (median inferior to 6dB, some outliers up to 10 dB) as those of the literature [17, 32, 33, 34].

For HD600, a Wilcoxon test indicates that the reliability of threshold measurements was significantly lower when headphone was removed, only at 2000 Hz ($p=0.008$) and 11000 Hz ($p=0.011$). This means that at these frequencies, the combination of "threshold variance" and "fitting variance" was larger than the only "threshold variance", so the headphone repositioning affected the audiogram reliability. For all other frequencies, the measurement variability was not modified when repositioning the headphone.

For TDH39, a Wilcoxon test indicates that the reliability of threshold measurements was significantly lower when headphone was removed, only at 4000 Hz ($p=0.014$) and 6000 Hz ($p=0.019$). For all other frequencies, the measurement variability was not modified when repositioning the headphone.

This finding is quite surprising: at 2 kHz, an effect of the HD600 position was hardly expectable because of the large wavelength compared to the dimensions of the headphone and ear canal. A theoretical explanation of this effect can be easily proposed for 11 kHz [11-14], but if the HD600 position has an effect at 11 kHz, why not at 14 kHz where it is far from being significant ($p=0.9213$ at 14000Hz)? In addition, except for 2 kHz and 11 kHz, this effect did not prove to be significant at any frequency and no trend could be revealed.

Identically, it could be expected that 4000 and 6000Hz were quite low frequencies to have an effect of TDH39 position, but actually an effect was found at these frequencies, but not higher ($p=0.326$ at 8000Hz, $p=0.247$ at 11000Hz, and $p=0.336$ at 14000).

The comparison of our results with those of previous studies must be taken with care for two reasons. Firstly, most of the previous works used 5 dB steps (whereas we used a 1 dB step). Secondly, as the distributions of our results are not normal, the use of means is

theoretically not advised; we should rather use the median, but in most of previous studies, means are indicated rather than medians (without taking care to the distribution) which makes the comparison with our results not straightforward. As an indication, the means of significant differences are indicated in Table 1 and are relatively close to the median values (see Figures 1 and 2).

Table I. Average differences between thresholds from for identical and different positions, for the frequencies giving significative differences.

	Identical position	Different positions
HD600 – 2 kHz	1.1 dB	2.6 dB
HD600 – 11 kHz	1.1 dB	2.4 dB
TDH39 – 4 kHz	2 dB	3.9 dB
TDH39 – 6 kHz	3 dB	5.5 dB

Results of measurement differences using TDH39 partially agree with those of previous studies: Hickling [5] found an effect of TDH39 removing at 6kHz and 8 kHz, but not at 1 and 2kHz (he did not test the other frequencies). We found an effect at 6kHz, but not at 8kHz. Moreover, Flottorp [35] described an experiment about annual monitoring of workers' auditory thresholds. When the audiograms showed suspicious losses at frequencies above 3KHz (in 32% of the cases), the headphone (TDH39) was removed and thresholds at 6kHz et 8kHz were measured another time. The new thresholds showed improvements superior to 5dB in more than 50% of the cases (but one should note that the audiogram resolution was only 5dB). Lindgren [3], using the TDH-49P supra-aural headphone and sessions including 5 test-retests, indicated differences up to 1.3 dB (from 250 Hz up to 8 kHz). Finally in a single test-retest experiment, Stuart [2] indicated differences up to 3 dB from 250 Hz up to 4 kHz, and 4 dB up to 8 kHz).

Erlandsson et al. [7] and Gauz et al. [25] used circumaural headphones (but not exactly HD600). Erlandsson et al. found an effect of headphone replacing and affirmed that a earphone adapted to an ear speculum could increase the reliability in the frequency range of 2 to 8 kHz. On the contrary, Gauz et al. did not find any significant difference when the headphone was removed. Valente et al. [36] realized single test-retests in high frequencies (between 8 and 18 kHz) with a circumaural Koss HV/1 A+. Differences between the test and

retest thresholds were not significant (but with a 5 dB step). In a test-retest with the circumaural Sennheiser HDA200 [37], 98% of the threshold differences were within a “clinically acceptable range of ± 10 dB from 8 to 14 kHz”. In another study [38], test-retest repeatability with HDA200 from 0.5 to 6 kHz and 8 to 16 kHz was within 10 dB for >99% and >94% of measurements, respectively. These results are in agreement with the distributions visible on Figure 1.

In summary, depending on the procedure, previous studies revealed or did not reveal an effect of headphone repositioning (at low or high frequencies). Our results showed an effect for several frequencies only, but not the highest ones. Hence, this seems to confirm perceptually the theory developed by Zhou and Green [15] (reduction of the size of the standing wave because of the non-infinite acoustic impedance of headphone and steeper slope of the psychometric function at the higher frequencies).

3.4 Effect of the headphone model

Agreeing with the examination of Figures 1 and 2, a Wilcoxon test indicated that the differences were higher for TDH39 than for HD600, when headphone was removed ($p < 0.0001$), and also when the headphone was not removed ($p = 0.001$). About “fitting variance”, this result agrees with those of Atherley et al. [4], who indicated that measurements with supra-aural headphones were less reliable than with circumaural headphones. Moreover, the “closed” headphones (like TDH39) generally do not account for a Free-air Equivalent Coupling (FEC) to the ear [29, 39]. That means that the acoustical loading applied by the TDH39 on the ear canal is not negligible, and could modify the resonances. However the difference between headphones when they were not removed is more surprising and seems indicate that even the “threshold variance” would depend on the headphone. A possible explanation is that the HD600, as it is circumaural, is more comfortable than the TDH39, which particularly press on ears. As a consequence, the subject, in a more comfortable position, could better focus on the task (especially in a test lasting around 45 minutes).

3.5 Effect of headphone position for each subject

Important differences between subjects could mean that the headphone position would be important for several subjects, whereas it would have a negligible effect for other ones. This subject effect could hide a potential smaller effect of the type of difference. Figure 3 and 4 shows the absolute value of the three difference types :

- *different positions, consecutive measurements* (dashed lines),
- *different positions, non-consecutive measurements* (dotted lines),
- *identical positions* (solid lines),

for each subject and for each frequency. No statistical analysis could be carried out because of the unique measurement for a given subject, a given frequency, and a given type of difference. However, several observations can be made.

Firstly, for several subjects, the variability between threshold measurements was large, whatever the difference type (when the headphone was removed or not): S7 for both HD600 and TDH39 ; S5, S8, S13, S19 for TDH39 exclusively. On the contrary, several other subjects always showed very low variability (whatever the difference type): S12 for both HD600 and TDH39 ; S6, S14, S15, S16, S19, S20 for HD600 exclusively; S4, S10 for THD 39 exclusively. The number of subjects having a poor reliability, whatever the difference type, appears to be higher for TDH39. This finding agrees with the comparison between Figures 1 and 2: the median was globally higher for TDH39 than for HD600 (confirmed by a Wilcoxon test: $p < 0.0001$, different and identical headphone positions together).

Secondly, several subjects seemed to be particularly sensitive to the headphone position (the difference “same position” is lower than the differences “different positions”): S6, S11, S14 for both HD600 and TDH39; S3, S8, S17, S19 for HD600 exclusively ; S1, S2, S9, S16, S17, S20 for TDH39 exclusively. On the contrary for the other subjects the three difference types were equivalent and the “threshold variance” seemed to be as large as the “fitting variance”.

The results for a given subject and a given headphone could be due to the coupling between the headphone and the external ear. Indeed, Pralong and Carlile [20] showed that the HPTF were highly individual, and these inter-individual differences could be enhanced by a supra-aural headphone (like TDH39), which cushions are more likely to deform the pinna than the circumaural headphones (like HD600). Flottorp [35] indicated that some persons with a specially shaped outer ear, particularly of the tragus and the ear canal opening, may require special care when the phone is placed at the ear. He added that for the determination of the standardized hearing threshold, the test subjects have to be carefully selected, using preliminary threshold measurement, and that persons who are especially sensitive to the placement of the ear cushion have to be excluded.

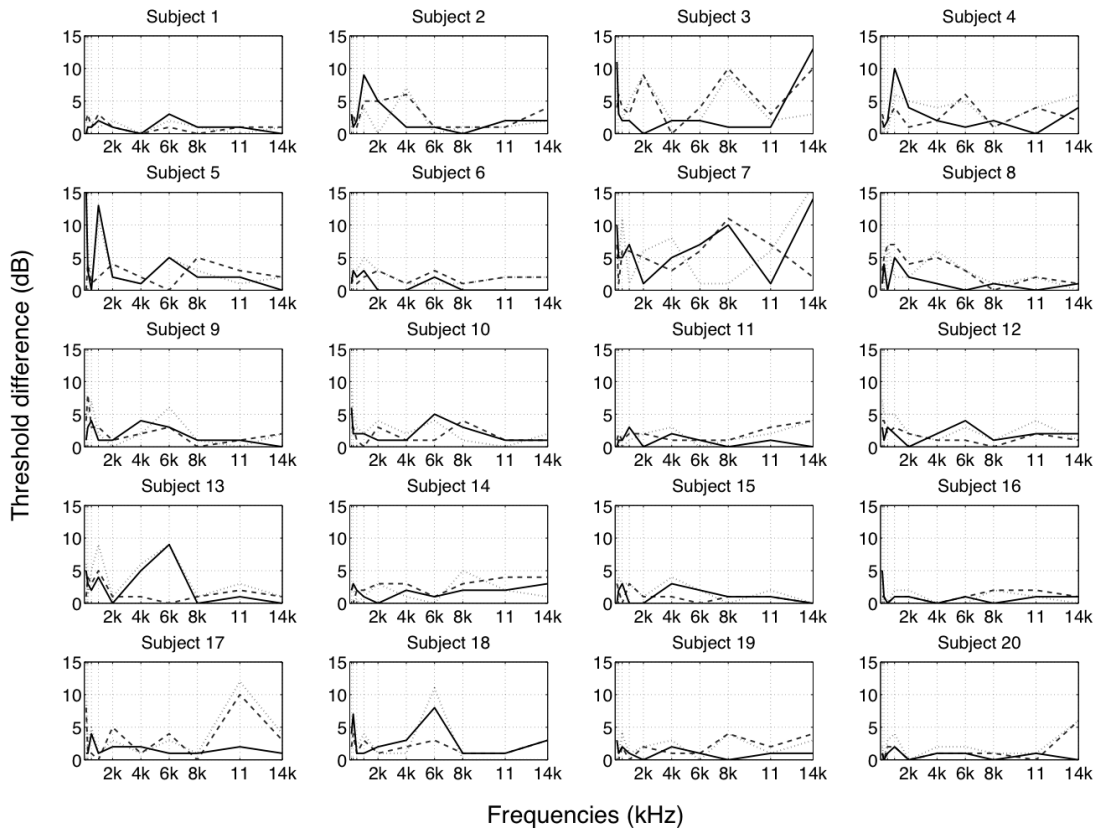


Figure 3. HD600 headphone - Differences between thresholds from different (dashed and dotted lines) vs identical (solid line) headphone positions, for the ten frequencies under test and each subject.

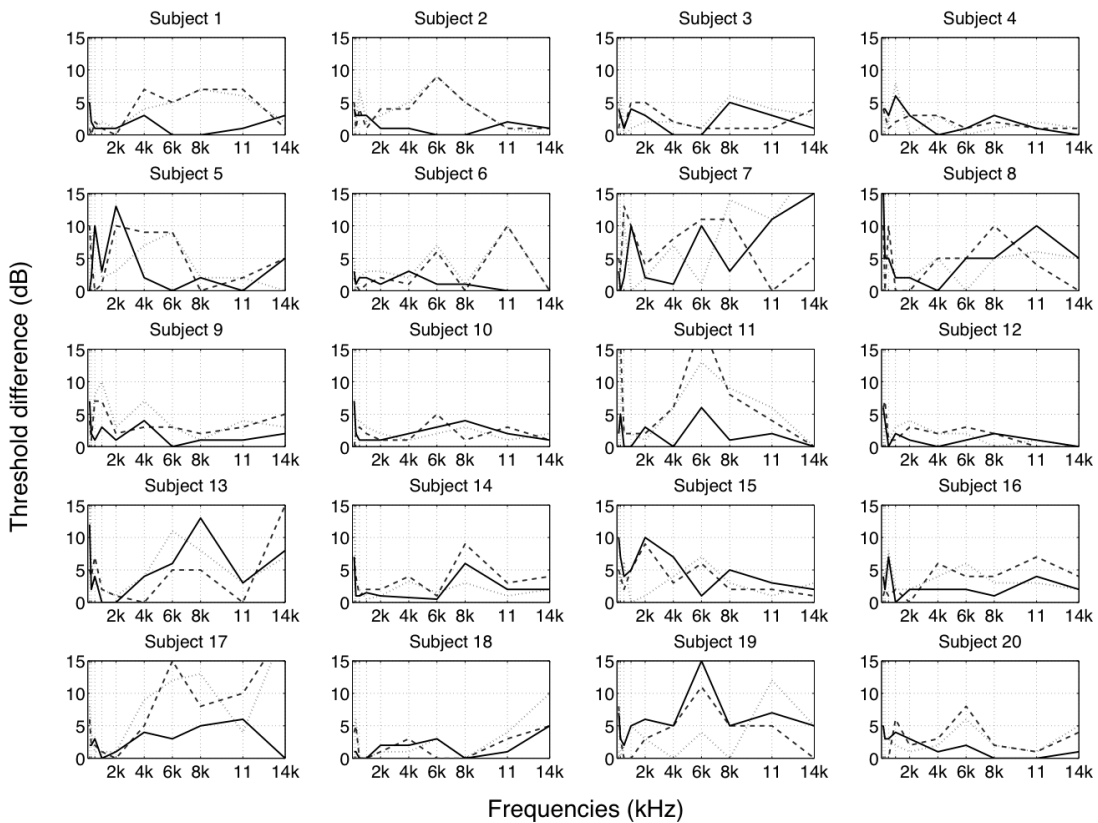


Figure 4. TDH39 headphone - Differences between thresholds from different (dashed and dotted lines) vs identical (solid line) headphone positions, for the ten frequencies under test and each subject.

4. CONCLUSION

In this study, audiograms (with 1 dB step, and including 10 frequencies ranging from 125 Hz to 14 kHz) were performed several times on normal-hearing subjects, for different –or not different- headphone positions (allowing to separate between effects of headphone position and cognitive factors).

For HD600, the reliability of threshold measurements was significantly lower when the headphone was removed, only at 2000 Hz and 11000 Hz. For all other frequencies, the measurement variability was not modified when repositioning the headphone. For TDH39, the reliability of thresholds was significantly lower when headphone was removed, only at 4000 Hz and 6000 Hz. As a result, measurements did not simply show a good reliability for low frequencies and a low reliability in high frequencies. So this experiment confirmed perceptually the alternative theory explaining that the reliability was not more crucial for high frequencies because of both the reduction of the size of the standing wave due to the non-infinite acoustic impedance of headphone and the steeper slope of the psychometric function at the higher frequencies.

Reliability between thresholds was better for HD600 than for TDH39, when the headphone was removed, and also when the headphone was not removed.

Finally large reliability differences were observable between subjects (with and without headphone removing). As a result, headphone position could be a crucial issue when measuring hearing thresholds for some subjects at certain frequencies that depends on the headphone model.

REFERENCES

- [1] Arlinger, S. D. (1991). Normal hearing threshold levels in the low-frequency range determined by an insert earphone. *The Journal of the Acoustical Society of America*, 90(5), 2411-2414.
- [2] Stuart, A., Stenstromb, R., Tompkins, C., & Vandenhoff, S. (1991). Test-retest variability in audiometric threshold with supraaural and insert earphones among children and adults. *International Journal of Audiology*, 30(2), 82-90.
- [3] Lindgren, F. (1990). A comparison of the variability in thresholds measured with insert and conventional supra-aural earphones. *Scandinavian audiology*, 19(1), 19-23.
- [4] Atherley, G. R. C., Hempstock, T. I., Lord, P., & Walker, J. G. (1967). Reliability of Auditory Threshold Determinations Using a Circumaural-Earphone Assembly. *The Journal of the Acoustical Society of America*, 42(1), 199-203.
- [5] Hickling, S. (1966). Studies on the reliability of auditory threshold values. *Journal of Auditory Research*.
- [6] Ivarsson, A., Erlandsson, B., Håkanson, H., Nilsson, P., Bennrup, S., & Sheppard, H. (1983). Comparison between hearing threshold measurements made with and without the aid of an ear speculum. *International Journal of Audiology*, 22(5), 421-429.
- [7] Erlandsson, B., Håkanson, H., Ivarsson, A., & Nilsson, P. (1980). The reliability of Bekesy sweep audiometry recording and effects of the earphone position. *Acta Oto-Laryngologica*, 89(S366), 99-112.
- [8] Stein, L., & Zerlin, S. (1963). Effect of circumaural earphones and earphone cushions on auditory threshold. *The Journal of the Acoustical Society of America*, 35(11), 1744-1745.
- [9] Jerger, J. F., & Tillman, T. W. (1959). Effect of earphone cushion on auditory threshold. *The Journal of the Acoustical Society of America*, 31(9), 1264-1264.
- [10] Speaks, C. (1969). Effect of Circumaural Earphone Cushions on Threshold Sensitivity for Speech. *The Journal of the Acoustical Society of America*, 46(6B), 1515-1518.
- [11] Stelmachowicz, P. G., Beauchaine, K. A., Kalberer, A., Kelly, W. J., & Jesteadt, W. (1989). High-frequency audiometry: Test reliability and procedural considerations. *The Journal of the Acoustical Society of America*, 85(2), 879-887.
- [12] Vinay, U, Svensson, P., Kvaløy, O., & Berg, T. (2015). A comparison of test-retest variability and time efficiency of auditory thresholds measured with pure tone audiometry and new early warning test. *Applied Acoustics*, 90, 153-159.
- [13] Stelmachowicz, P. G., Beauchaine, K. A., Kalberer, A., Langer, T., & Jesteadt, W. (1988). The reliability of auditory thresholds in the 8-to 20-kHz range using a prototype audiometer. *The Journal of the Acoustical Society of America*, 83(4), 1528-1535.

- [14] Stinson, M. R., & Shaw, E. A. G. (1982). Wave effects and pressure distribution in the ear canal near the tympanic membrane. *The Journal of the Acoustical Society of America*, 71(S1), S88-S88.
- [15] Zhou, B., & Green, D. M. (1995). Reliability of pure-tone thresholds at high frequencies. *The Journal of the Acoustical Society of America*, 98(2), 828-836.
- [16] Robinson, D. W., & Whittle, L. S. (1973). A comparison of self-recording and manual audiometry: Some systematic effects shown by unpractised subjects. *Journal of Sound and Vibration*, 26(1), 41-62.
- [17] Lutman, M. E., Cane, M. A., & Smith, P. A. (1989). Comparison of manual and computer-controlled self-recorded audiometric methods for serial monitoring of hearing. *British journal of audiology*, 23(4), 305-315.
- [18] Toole, F. E. (1984, May). The acoustics and psychoacoustics of headphones. In *Audio Engineering Society Conference: 2nd International Conference: The Art and Technology of Recording*. Audio Engineering Society.
- [19] Wightman, F. L., & Kistler, D. J. (1989). Headphone simulation of free-field listening. I: Stimulus synthesis. *The Journal of the Acoustical Society of America*, 85(2), 858-867.
- [20] Pralong, D., & Carlile, S. (1996). The role of individualized headphone calibration for the generation of high fidelity virtual auditory space. *The Journal of the Acoustical Society of America*, 100(6), 3785-3793.
- [21] McAnally, K.I., & Martin, R. L. (2002). Variability in the headphone-to-ear-canal transfer function. *Journal of the Audio Engineering Society*, 50(4), 263-266.
- [22] Kulkarni, A., & Colburn, H. S. (2000). Variability in the characterization of the headphone transfer-function. *The Journal of the Acoustical Society of America*, 107(2), 1071-1074.
- [23] Paquier, M., & Koehl, V. (2015). Discriminability of the placement of supra-aural and circumaural headphones. *Applied Acoustics*, 93, 130-139.
- [24] Henry, J. A., Flick, C. L., Gilbert, A., Ellingson, R. M., & Fausti, S. A. (2001). Reliability of hearing thresholds: computer-automated testing with ER-4B Canal Phone™ earphones. *Journal of rehabilitation research and development*, 38(5), 567-581.
- [25] Gauz, M. T., Robinson, D. O., & Peters, G. M. (1981). High-frequency Bekesy audiometry: III. Reliability and validity revisited. *The Journal of auditory research*, 21(3), 167-180.
- [26] American National Standards Institute. (1978). Methods for manual pure-tone threshold audiometry. *ANSI*, 3, 21.
- [27] British Society of Audiology, Recommended Procedure. (2011). Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking », 9th September 2011.

- [28] NF EN 26189 ISO 6189. (1992). Audiométrie liminaire tonale en conduction aérienne pour les besoins de la préservation de l'ouïe.
- [29] Møller, H., Hammershøi, D., Jensen, C. B., & Sørensen, M. F. (1995). Transfer characteristics of headphones measured on human ears. *Journal of the Audio Engineering Society*, 43(4), 203-217.
- [30] Møller, H., Sørensen, M. F., Jensen, C. B., & Hammershøi, D. (1996). Binaural technique: Do we need individual recordings?. *Journal of the Audio Engineering Society*, 44(6), 451-469.
- [31] Zwislocki, J., Maire, F., Feldman, A. S., & Rubin, H. (1958). On the effect of practice and motivation on the threshold of audibility. *The Journal of the Acoustical Society of America*, 30(4), 254-262.
- [32] S. Hickling, "The validity and reliability of pure tone clinical audiometry", *The New Zealand Medical Journal* 63, 379-82 (1964).
- [33] R.S. Tyler, E.J. Wood, "A comparison of manual methods for measuring hearing levels", *Audiology* 19, 316-29 (1980)
- [34] H.A. Newby, "Audiology", New York: Appleton-Century-Crofts (1972)
- [35] Flottorp, G. (1995). Improving audiometric thresholds by changing the headphone position at the ear. *International Journal of Audiology*, 34(5), 221-231.
- [36] Valente, M., Valente, L. M., & Goebel, J. (1992). High-frequency thresholds: circumaural earphone versus insert earphone. *Journal of the American Academy of Audiology*, 3(6).
- [37] Frank, T. (2001). High-frequency (8 to 16 kHz) reference thresholds and intrasubject threshold variability relative to ototoxicity criteria using a Sennheiser HDA 200 earphone. *Ear and hearing*, 22(2), 161-168.
- [38] Schmuziger, N., Probst, R., & Smurzynski, J. (2004). Test-retest reliability of pure-tone thresholds from 0.5 to 16 kHz using Sennheiser HDA 200 and Etymotic Research ER-2 earphones. *Ear and hearing*, 25(2), 127-132.
- [39] Møller, H. (1992). Fundamentals of binaural technology. *Applied acoustics*, 36(3), 171-218.