

SENSITIVITY OF AUDITORY PERCEPTION TO CHANGES IN PHASE SPECTRUM

Ivana Štěpánková

Czech Technical University in Prague, Faculty of Electrical Engineering, Prague, Czech Republic

Abstract

This paper deals with sensitivity of human auditory to changes in sound phase spectrum. Although the human auditory system was considered “phase deaf” for a long time, nowadays a few recent studies proved that the change of sound phase spectrum has a significant impact on auditory perception. The aim of this paper is to verify these mentioned claims. To achieve this aim it is fundamental to perform listening tests with phase changed audio signals. The phase changes were applied to two groups of audio signals – synthetic signals and real signals. Thus, the listening test consisted of six different experiments and fourteen subjects participated in the listening test. The results of the test were statistically analyzed by ANOVA and the effect of phase changes to human auditory perception was determined based on these results.

Keywords

Phase change, phase spectrum, psychoacoustic test, signal processing, analysis of variance

Introduction

The sound phase spectrum was considered not very significant for human auditory perception for many years. The human ear was already designated phase deaf in 1845 by Georg Simon Ohm [1]. Another experiments based on Ohm’s assumption were made by Hermann Helmholtz, who confirmed Ohm’s theory of zero effect of phase changes to auditory perception [2]. However, these claims were later disproved by several scientists and recently has been noticed that humans are sensitive to change in the phase spectrum, especially with anechoic speech signals [3] and applause-type signals [4, 5]. The phase spectrum of an audio signal is often modified due to decorrelation, quantization and downmixing [6].

But why was the theory of phase deaf human ears thought to be correct for such a long time? The fact that this assumption holds well for most of the audio signals is believed to be the reason [7]. Several newer studies proved that the phase change has the influence especially on the timbre [6]. The effective method to verify these claims is to perform psychoacoustic test which consists of audio signals having various changes in phase. Two groups of signals (synthetic and real) are chosen for the experiment in this study. The parameters of synthetic signals are taken from the study [6] which used these signals to create auditory model for analysing phase perception. In the second part of this study the real signals are phase modified. The modification in this case

consists in replacement of the exact part of phase spectrum by the exponential function. The aim of this part is to discover the impact of these phase modifications and to find out if the type of the real signal has a significant effect on the results. Different parts of phase spectrum were replaced by exponential function to determine the frequency dependence of these phase changes.

Human auditory system

The function of the human ear is to transform the sound waves to auditory perception. The outer ear works as pressure receiver of the sound. Due to this, the outer ear function consists of transfer of the vibrations by the ear canal and creation of eardrum movement. Next, these vibrations are transmitted through the three ossicles in the middle ear to the cochlea. [8] The cochlea transforms mechanical vibrations into neural pulses of hair cells located on the basilar membrane. The hair cells sensitive to high frequencies are located closest to the inner ear entrance. Conversely, the hair cells tuned on lower frequencies are located farther from the upcoming vibration. The frequency selectivity was found to follow the equivalent rectangular bandwidth (ERB) [9]. Thus, the effect of vibration in one bandwidth is shown as one common neural response. As mentioned, the high frequency cells respond earlier than the low frequency cells. Hence, the frequency-dependent group delay is

created. This delay is partly compensated during the higher processing level. [10] Finally, hair cells transmit the sound information through auditory nerve to the brainstem.

Phase Perception Studies

Until these days several studies have been focused on the effect of phase changes of audio signal spectrum. [1, 2, 6] As mentioned, the earlier studies came up with the opposite results than recent studies.

Ohm's acoustic law

Georg Simon Ohm was the first scientist who came up with the study of sensitivity of human auditory to changes in phase spectrum. In his study called *Über die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen* determined that the phase spectrum isn't important for auditory perception [1, 6, 11].

A few years later, Hermann Helmholtz performed an experiment based on Ohm's result. His experiment consisted in the change of first twelve harmonics. He imagined the cochlea as a spectrum analyzer and claimed that the only factor, which has the effect on auditory perception is the magnitude of particular frequency component. In other words, his result confirmed Ohm's acoustic law [2, 6, 11].

Recent studies

In 1987 M. S. Patterson's study showed that human ears are not phase deaf. He supposed that acoustic properties of a room could eliminate the existing difference of phase modified signal. Following studies also proved that changes of phase spectrum affect auditory perception (Schroeder – 1959, Plomp and Steeneken – 1969, Bilsen – 1973, Patterson – 1987, Moore and Glasberg – 1989, Laitinen, Disch and Pulkki – 2013) [11].

The impact of the phase spectrum on the timbre perception was studied by Plomp and Steeneken. The results of the first part of their experiments proved that the tones with alternating sine and cosine components show a significant difference in the comparison to the signals with only sine or cosine components. In the second part of their study they focused on the quantitative expression of the previous phase changes by changing the slope of the magnitude spectrum of the signals with only sine or cosine components. The effect of the previous phase changes (alternating sine and cosine components) was found to be quantitatively smaller than the effect of changing the slope of the magnitude spectrum by 2 dB/oct.

Other similar psychoacoustic experiments were performed by Patterson. He used the same reference signal (sum of the cosine components) as Plomp and

Steeneken. His phase modifications consisted in the phase shift of every second component by the same amount D . The aim of this experiment was to discover the amount D which represents the just noticeable difference (JND) between reference and modified signal. They found out that this amount of phase shift D is dependent on the fundamental frequency of the signal. In some cases already 15 degrees was determined as JND but in some other cases more than 60 degrees was needed to achieve JND.

Moore's and Glasberg's psychoacoustic experiments were created to verify the ability to detect a phase change of a single component in a harmonic complex tone. These complex tones contained the first 20 harmonics. Only one harmonic was phase shifted in order to find out the minimum amount of phase shift which causes perceived difference. The amount of this minimal phase shift which was determined 2–4 degrees.

Laitinen, Disch and Pulkki studied a human sensitivity to phase changes in order to create an auditory model for analysing phase perception. In the first part of this study they performed many psychoacoustic experiments that consisted of synthetic harmonic complex signals. Based on the results of these tests, an auditory model was developed. The aim of this study was to mimic the firing rate of the neurons in the cochlea. It was found that the crest factor of the neural firing rate (the ratio between the loudest amplitude values and the mean amplitude value) for different frequency bands can be used to explain differences in the perception due to phase modifications. The high crest factor at mid and high frequencies of the tone indicates a perception of a buzzy sound. At the lowest frequencies of the tone the high crest factor indicates a perception of loud bass [6].

Listening tests

At first, the application for the psychoacoustic test was created. This application was developed using The Lazarus Integrated Development Environment. The graphical user interface, which is represented in Fig. 1, was used for all six experiments. The scaling method was chosen for the evaluation of the phase modified signals. According to this, the graphical user interface in Fig. 1 provides the comparison of the reference signal and the phase modified signals and evaluates the difference between them on the scale from 1 to 5 with the 0.1 step. Number 1 on the scale represents absolute match with the reference signal, number 5 absolute difference. Final evaluation of every tested signal was obtain as average rating of all particular evaluations. The testing subjects could arbitrarily and repeatedly play reference or modified signal, as long as they were sure of their evaluation. Hence, the duration of the test depended on every subject's speed of evaluation but the total duration never exceeded ten minutes. The average run of the test took about 5 minutes. The length of every

synthetic signal is 2.5 seconds, applause type signal 4 seconds, and instrumental music and speech 4 seconds. The order of modified signals was randomized for every test. Before beginning of the test, all testing subjects were asked to fill a short form containing questions about age, sex and experience with psychoacoustic tests.

Psychoacoustic test took place at Faculty of Electrical Engineering in Prague, Department of Radioelectronics in the audiology room, which provides the isolation of surrounding noise. The RME Fireface UC sound card and Sennheiser HD650 headphones were used for the test. Fourteen subjects, excluding the author, participated in the listening test. First of all, the evaluation method was explained to the subject. Subsequently the subject was invited to try the trial test. According to the fact, that the test consists of six particular experiments, the subjects were allowed to take a rest between the experiments if necessary.

Two groups of audio signals with phase modification were prepared for the listening test. The first group consists of synthetic signals, the second of real signals.

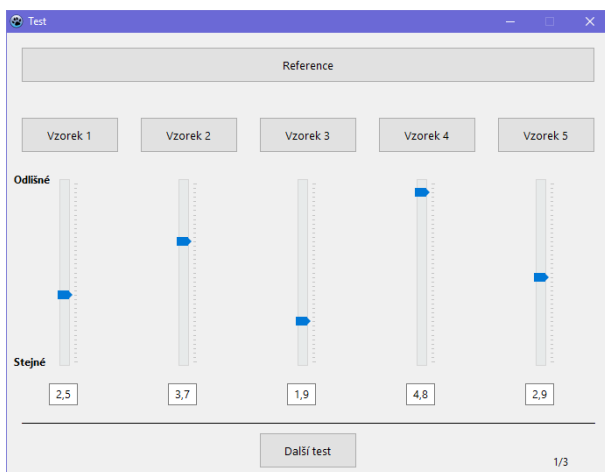


Fig. 1: The graphical user interface.

Synthetic signals

All synthetic signals were created using MATLAB as the sum of cosines with different parameters. This method and parameters of signals were taken from Laitinen's study [6]. All synthetic signals used for listening test have sample frequency 48 kHz and were created by following formula:

$$x(t) = G \cdot \sum_{n=1}^{\infty} \frac{g_n}{n} \cos(2\pi n f_0 (t - \tau_n) + \phi_n), \quad (1)$$

where G is the gain controlling the overall level of the signal, g_n is a frequency-dependent gain for controlling the magnitude spectrum, n is the sequential number of the harmonic, f_0 is the fundamental frequency (100 Hz), τ_n is a frequency-dependent delay, and ϕ_n is

a frequency dependent angle for controlling the phase spectrum [6].

By using this method there were prepared three experiments. Experiment 1 compared the effect of phase and amplitude modification. Experiment 2 was focused on the effect of phase shift of one harmonic and Experiment 3 dealt with the effects of different levels of the phase modification.

Real signals

Three types of audio signals were chosen for the phase modification – speech, applause and instrumental music. Every phase spectrum of signal was changed using MATLAB. The modification in this case consists in replacement of the exact part of phase spectrum by the exponential function with the range from $-\pi$ to π . During the preparation of signals for this experiment other functions were tried (linear, sine, cosine, logarithmic, sawtooth wave). The informal listening test proved that there are not strong differences between signals whose phase spectrums were replaced by different functions. Thus, only one function was chosen for the psychoacoustic experiment. Four modified signals were created for every type of audio signal. Particular signals of one type are different from each other by the bandwidth of modified phase spectrum.

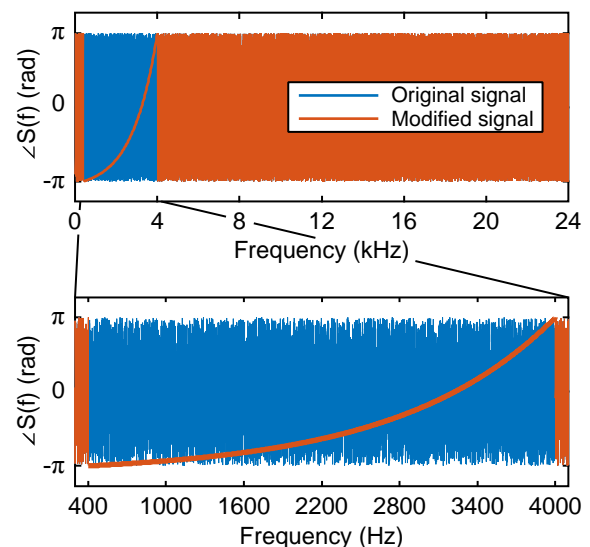


Fig. 2: Example of phase spectrum modification by the exponential function.

Experiments and results

The whole psychoacoustic test consists of six different experiments. First three of them include a synthetic signal, other three experiments include real signals. The exact modification of phase spectrum of every signal we

used is described in this section. The results of evaluation are presented in the graphs. Every experiment contains one of the tested signal which is the same as reference signal. This signal is located at the first place in the tables and graphs.

Statistical analysis

Analysis of variance (ANOVA) was used for the statistical analysis of the results. Applying this analysis, it is possible to verify the results gained from the experiments and to determine if the phase change has the real effect on the evaluation of the signals. The significance level was set at the usual value 0.05.

Having completed Analysis of variance, Tukey's post-hoc multicomparative tests were performed. These tests are applied to obtain if the differences between evaluation of signals in one experiment are statistically significant or not. Almost all 60 multicomparative tests proved that there are statistically significant differences between evaluations of the signals, except 5 pairs of signals with a very low difference of evaluation. The

results of ANOVA are always presented as $F(df_B, df_W)$, where F is the ratio of between-group variability and within-group variability, df_B is the numerator degree of freedom between the groups and df_W is the numerator degree of freedom within the groups. The F ratio is compared with $F_{critical}$ which is table value in the Fisher-Snedecor distribution chosen for the significance level 0.05.

Experiment 1

This experiment compares the effect of phase and amplitude modification. The parameters of the signals we used are presented in Table 1. The function N in the Table 1 means the normal distribution. Applying the ANOVA to the results of Experiment 1 was obtained $F(4, 65) = 128.6; p < 0.05$ and $F > F_{critical}$. It follows from this result, that at least one pair of signals has a statistically significant difference of their evaluations. The multicomparative tests proved that all differences of all signal pairs are statistically significant. The results of this experiment are presented in Fig. 3.

Table 1: Parameters of the signals used in Experiment 1.

Signal	f_0 (Hz)	g_n	Φ_n	τ_n (ms)
sig1_inphase	100	1	$(i - 1) \cdot \pi/2$	0
sig2_randomphase	100	1	$\sim U(0, 2\pi)$	0
sig3_1dBstd	100	$10^{L/20}, L \sim N(0, 1)$	$(i - 1) \cdot \pi/2$	0
sig4_2dBstd	100	$10^{L/20}, L \sim N(0, 2)$	$(i - 1) \cdot \pi/2$	0
sig5_4dBstd	100	$10^{L/20}, L \sim N(0, 4)$	$(i - 1) \cdot \pi/2$	0

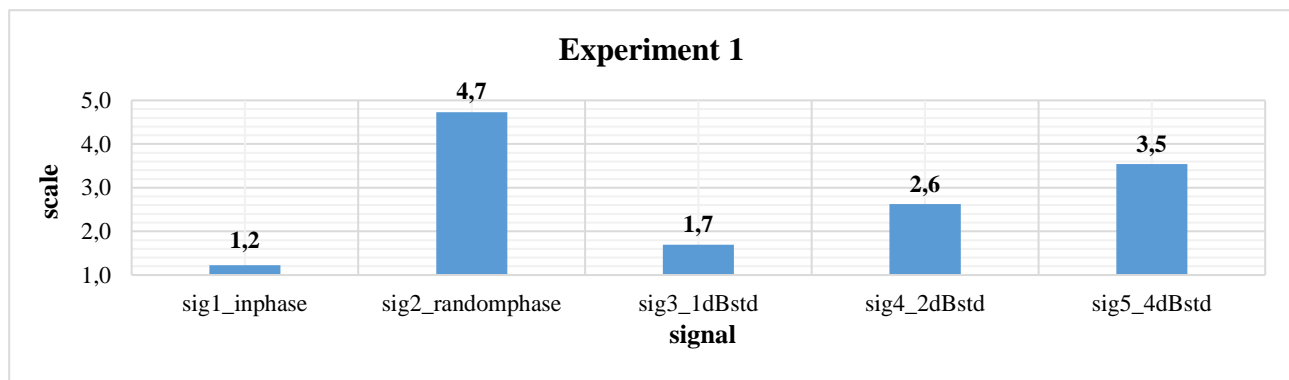


Fig. 3: The results of Experiment 1.

Experiment 2

Experiment 2 was focused on the effect of phase shift of one harmonic. The phase shift was applied to the harmonic at 3 kHz. This harmonic was chosen based on the resonance of human ear which causes the amplification of frequency range 3–4 kHz. The parameters of the signals are presented in Table 2. Applying ANOVA to the results of the Experiment 2

was obtained $F(4, 65) = 29.3; p < 0.05$ and $F > F_{critical}$.

According to the results gained by multicomparative tests, a statistically significant difference was not discovered just between two pairs of signals - sig1_inphase and sig3_3dBboost, sig2_180shift and sig5_9dBboost. The results of the Experiment 2 are presented in Fig. 4.

Table 2: Parameters of the signals used in Experiment 2.

Signal	f_0 (Hz)	g_n	ϕ_n	τ_n (ms)
sig1_inphase	100	1	$(i - 1) \cdot \pi/2$	0
sig2_180shift	100	1	$(i - 1) \cdot \pi/2 + \pi, i = 30$ $(i - 1) \cdot \pi/2, i \neq 30$	0
sig3_3dBboost	100	$10^{\frac{L}{20}}, L = 3, i = 30$ 1, $i \neq 30$	$(i - 1) \cdot \pi/2$	0
sig4_6dBboost	100	$10^{\frac{L}{20}}, L = 6, i = 30$	$(i - 1) \cdot \pi/2$	0
sig5_9dBboost	100	$10^{\frac{L}{20}}, L = 9, i = 30$	$(i - 1) \cdot \pi/2$	0

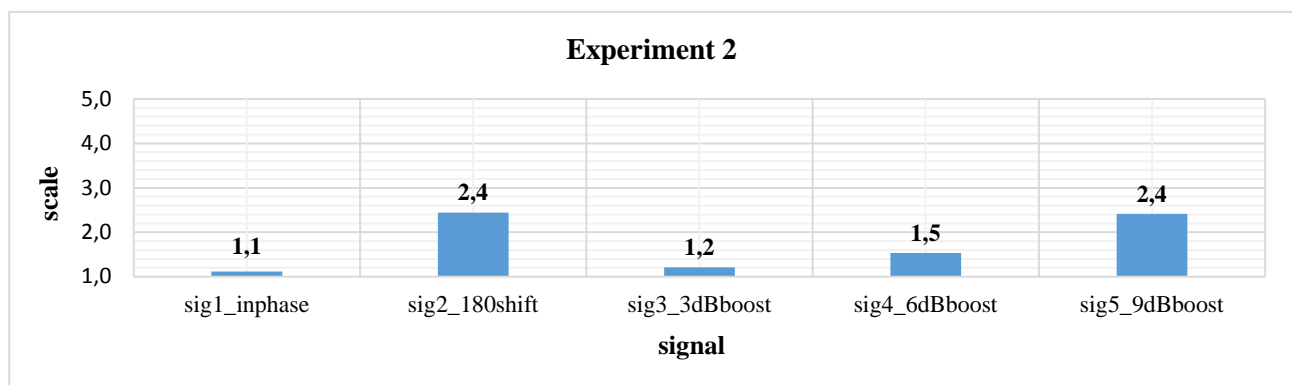


Fig. 4: The results of the Experiment 2.

Experiment 3

Experiment 3 dealt with the effects of different levels of phase modification. The parameters of the signals are presented in Table 3. The function U in the Table 3 means the uniform distribution. Applying ANOVA to the results of Experiment 3 was obtained $F(4, 65) =$

205.6; $p < 0.05$ and $F > F_{critical}$. According to the results gained by multicomparative tests, a statistically significant difference was not discovered just between one pair of the signals - sig4_090degrees and sig5_180degrees. The results of Experiment 3 are presented in Fig. 5.

Table 3: Parameters of the signals used in Experiment 3.

Signal	f_0 (Hz)	g_n	ϕ_n	τ_n (ms)
sig1_inphase	100	1	$(i - 1) \cdot \pi/2$	0
sig2_20degrees	100	1	$(i - 1) \cdot \pi/2 + \sim U(-\pi/9, \pi/9)$	0
sig3_45degrees	100	1	$(i - 1) \cdot \pi/2 + \sim U(-\pi/4, \pi/4)$	0
sig4_90degrees	100	1	$(i - 1) \cdot \pi/2 + \sim U(-\pi/2, \pi/2)$	0
sig5_180degrees	100	1	$\sim U(0, 2\pi)$	0

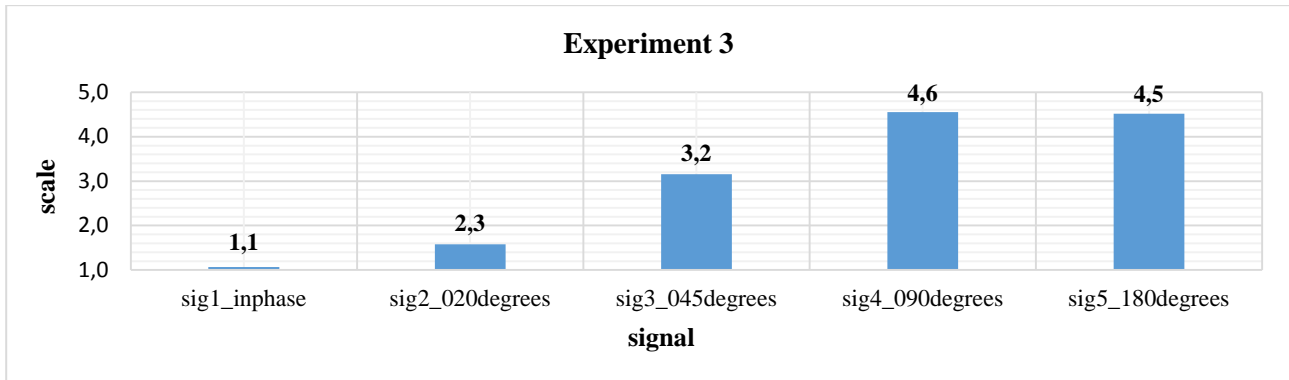


Fig. 5: The results of Experiment 3.

Experiment 4

Experiment 4 contained real audio signals with modified phase spectrum, as is described in section 4.2. The parts of the phase spectrum, which were replaced by exponential function are presented in Table 4. Applying ANOVA to the results of Experiment 4 was

obtained $F(4, 65) = 161.7$; $p < 0.05$ and $F > F_{critical}$. According to the results gained by multicomparative tests, a statistically significant difference was discovered between all pairs of the signals used in this experiment. The results of Experiment 4 are presented in Fig. 6.

Table 4: Modified parts of the phase spectrum of the signals used in Experiment 4.

Signal	music	music_exp_lf	music_exp_mf	music_exp_mf2	music_exp_hf
Modified part of the phase spectrum (Hz)	–	0 – 100	100 – 500	500 – 8000	8000 – 24000

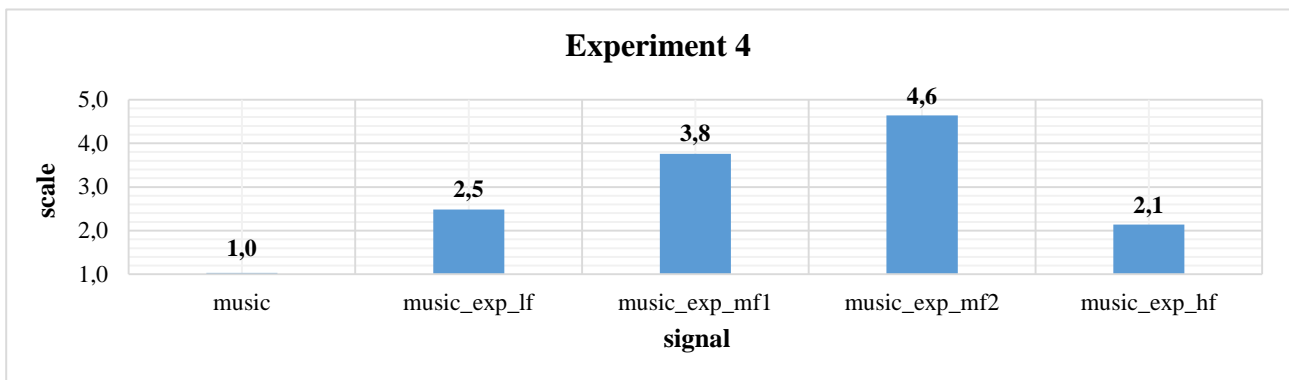


Fig. 6: The results of Experiment 4.

Experiment 5

Experiment 5 contains real audio signals with modified phase spectrum, as is described in section 4.2. The parts of the phase spectrum, which were replaced by exponential function are presented in Table 5. Applying ANOVA to the results of Experiment 5 was

obtained $F(4, 65) = 67.1$; $p < 0.05$ and $F > F_{critical}$. According to the results gained by multicomparative tests, a statistically significant difference was not discovered just between one pair of signals - applaus_exp_mf and applaus_exp_hf. The results of Experiment 5 are presented in Fig. 7.

Table 5: Modified parts of the phase spectrum of the signals used in Experiment 5.

Signal	applause	applaus_exp_lf	applaus_exp_mf	applaus_exp_mf2	applaus_exp_hf
Modified part of the phase spectrum (Hz)	–	0–100	100–500	500–8000	8000–24000

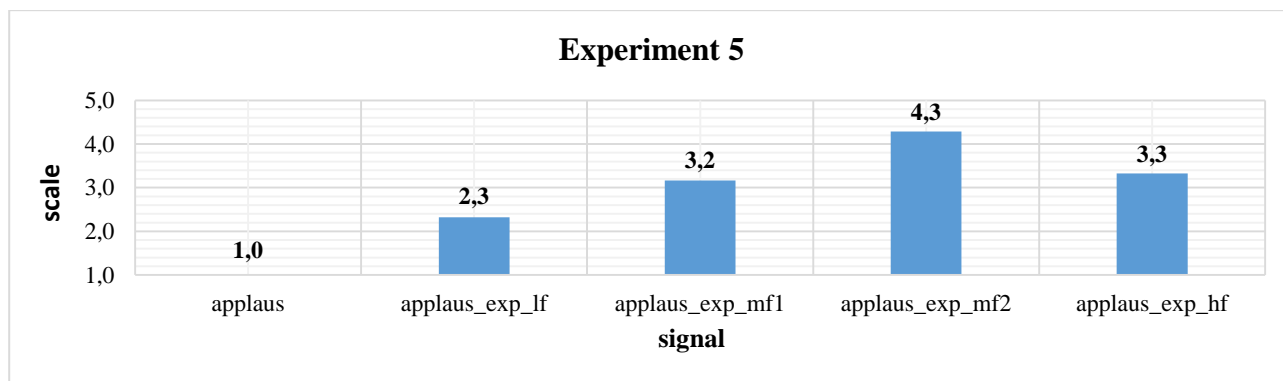


Fig. 7: The results of Experiment 5.

Experiment 6

Experiment 6 contains real audio signals with modified phase spectrum, as is described in section 4.2. The parts of the phase spectrum, which were replaced by exponential function are presented in Table 6. Applying ANOVA to the results of Experiment 6 was obtained $F(4, 65) = 128.6$; $p < 0.05$ and $F > F_{critical}$.

According to the results gained by multicomparative tests, a statistically significant difference was not discovered between two pairs of signals - speech_exp_mf and speech_exp_mf2, speech and speech_exp_lf. The results of Experiment 6 are presented in Fig. 8.

Table 6: Modified parts of the phase spectrum of the signals used in Experiment 6.

Signal	speech	speech_exp_lf	speech_exp_mf	speech_exp_mf2	speech_exp_mf3
Modified part of the phase spectrum (Hz)	–	0–100	100–300	300–500	500–8000

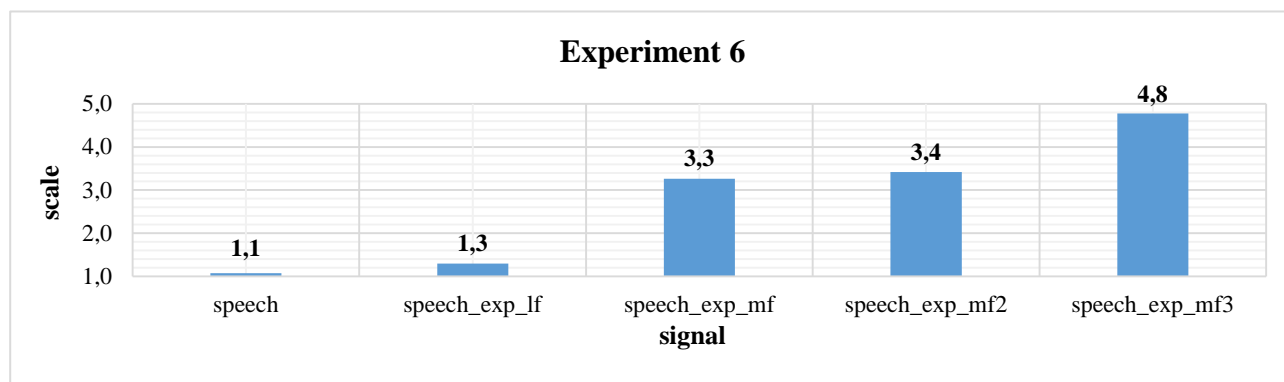


Fig. 8: The results of Experiment 6.

Conclusion

The aim of this study was to verify human ability to perceive differences in sound due to the modification of the phase spectrum. To accomplish this aim, the psychoacoustic test was prepared. The psychoacoustic test consisted of two groups of the signals – synthetic signals and real signals. For every group, three sets of signals were created. Thus, six psychoacoustic experiments were prepared in total and fourteen subjects participated in all experiments. The results of the

psychoacoustic test verified the claim of recent studies and proved that the human ear is not phase deaf and that the change of the phase spectrum has a significant effect on auditory perception.

Experiment 1 compared the effect of phase and amplitude modification. The results of this experiment (Fig. 3) showed that the randomization of the phase spectrum has larger effect on auditory perception than the randomization of the magnitude spectrum. As can be seen in Fig. 3, the difference of auditory perception increases with the size of the standard deviation of the magnitude spectrum used in the test.

Experiment 2 was focused on the effect of phase shift of one harmonic. Due to the results in Fig. 4, Signal 2 (all the harmonics are in phase except the harmonic at 3 kHz, which is shifted by 180 degrees) was evaluated the same different from the reference signal as Signal 5 (all the harmonics are in phase, but the magnitude of the harmonic at 3 kHz is amplified by 9 dB). Hence, the amplification of the harmonic component was noticed to cause a similar effect as the phase shift of the harmonic.

Experiment 3 dealt with the effects of different levels of the phase modification. It follows from Fig. 5 that two signals with the largest level of randomization are the most different from the reference signal. Specifically, Signal 4 (the randomization is restricted to within $\pm 90^\circ$) and Signal 5 (the phases are completely randomized). Thus, we can claim that the size of perceptual difference depends monotonically on the level of randomization.

These three experiment consisted of synthetic signals. This method and parameters of these signals were taken from the study [6], which used these signals to create auditory model for analysing phase perception. The results obtained from these experiments were similar to the results obtained in the study [6] and also similar to the output of the auditory model.

Experiment 4, 5 and 6 contained real audio signals with modified phase spectrum. In these cases, parts of the phase spectrum were replaced by the exponential function. As can be seen from the results (Fig. 6, Fig. 7, Fig. 8), in every experiment the biggest difference was noticed between reference signal and the signals with 500–8000 Hz range of modified part of phase spectrum. Particular results of these three experiments do not differ from each other. The cause of this result is probably the fact that applause and speech are known as the phase sensitive signals (according to recent studies) and were chosen for this psychoacoustic test. On the other hand, the instrumental music proved similar results as these two types of signals. In the upcoming study, the author is going to test other types of real signals and try to determine the just noticeable difference of phase change.

Acknowledgement

Study described in the paper was supervised by Dr. L. Husník, FEE CTU in Prague and supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS17/190/OHK3/3T/13.

References

- [1] OHM, G. S.: *Über die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen*, *Ann. Phys. und Chemie*, 1843.
- [2] HELMHOLTZ, H. L. F.: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Music*, 1863.
- [3] LAITINEN M.-V., PULKKI, V.: *Utilizing Instantaneous Direct-to-Reverberant Ratio in Parametric Spatial Audio Coding*, 2012.
- [4] HOTHO, G., VAN DE PAR, S., BREEBAART, J.: *Multi-channel Coding for Applause Signals*, 2008.
- [5] LAITINEN, M.-V., DISH, S., PULKKI, V., KUECH F.: *Reproducing Applause-Type Signals with Directional Audio Coding*, 2011.
- [6] LAITINEN, M.-V., DISH, S., PULKKI, V.: *Sensitivity of Human Auditory to Changes in Phase Spectrum*, *JAES*, vol. 61, no. 11, 2013.
- [7] VILKAMO, J., LOKKI, T., PULKKI, V.: *Directional Audio Coding: Virtual Microphone Based Synthesis and Subjective Evaluation*, 2009.
- [8] HUDDE, H.: *Communications Acoustics*, 2005.
- [9] MOORE, B. C. J., GLASBERG, B. R.: *Suggested Formulae for Calculating Auditory-Filter Bandwidths and Excitation Patterns*, 1983.
- [10] WOJTCZK, M., BEIM, J. A., MICHEYL, C., OXENHAM, A. J.: *Perception of Across-Frequency Asynchrony and Role of Cochlear Delays*, 2012.
- [11] SANTALA, O.: *Perception and Auditory Modeling of Spatially Complex Sound Scenarios*, 2015.

Ivana Štěpánková
Department of Radioelectronics
Faculty of Electrical Engineering
Czech Technical University in Prague
Technická 2, 166 27 Prague, Czech Republic

E-mail: stepaiv1@fel.cvut.cz