Detecting of Residual Pitch in Two-Component Complexes

Adam Mielczarek, Andrzej Dobrucki
Wroclaw University of Technology, Institute of Telecommunications and Acoustics, 50-370 Wroclaw, Wybrzeże Wyspiańskiego 27, Poland,
{adam.mielczarek, andrzej.dobrucki}@pwr.wroc.pl

The paper presents results of research on residual pitch perception threshold in two-component complexes. In the adjustment procedure listeners had to specify the pitch of two- component stimuli. The subject listened to two components, which were the successive harmonics of 200 Hz. The frequency range was 400-2000 Hz. The component with lower frequency had constant level 50 dB SPL, while the level of the component with higher frequency was changing in the range of 5 - 50 dB SPL in 5 dB step. The first pitch shift effect was also examined. The same stimuli were used in the experiment, but the both components were equally shifted in frequency range. The 30 Hz shift was applied. The frequency range was 430-1630 Hz. The same pattern of presentation the stimuli was used in both experiments. Two models were used to confirm the psychoacoustic experiment results and explain the ability of perception the residual pitch: the auditory toolbox based on the Licklider duplex theory and the time structure model of residual pitch perception. They gave the evidence of existence of the residue pitch perception threshold in two-component complexes. The appearance of the first pitch shift effect was also confirmed by both models. The conclusion is that the nonlinearity of hearing system does not influence the perception of residual tone in two-component complexes.

1 Introduction

The earlier researches [1, 2] on the residue phenomenon did not give unambiguous answer that the missing fundamental is clearly perceived in a two-components complex. Many works [3,4] explained the possibility of defining the residual pitch on the basis of the nonlinearity of a hearing system. They suggested that the nonlinear components like $f_2–f_1$ or $f_1–k(f_1–f_2)$ dominate in residue perception. However, the discovery of the first pitch shift effect [5] gave evidence that the difference tone does not influence residual pitch perception. The frequency shift of the tone-complex components gave a linear residual pitch shift. The newest investigations [6, 7] on the residue reveal that the perception of the residue phenomenon in two-frequency stimuli is possible and depends strongly on the level of individual components. After crossing a particular level of the components, the effect becomes unambiguously defined.

This paper presents the research results regarding the residue phenomenon obtained by hearing experiments. The first pitch shift effect was also examined and provided evidence that the nonlinear components do not dominate in the residue perception in two-frequency stimuli. The following two models are used to confirm these results: the time structure model [8] created by Schouten in 1962, and the auditory model [9, 10] created by Slaney and Lyon based on the Licklider duplex theory [11] of pitch perception.

2 The hearing experiments

Ten people with normal hearing took part in the experiments. The experiments were conducted with the measuring set shown in Fig. 1.

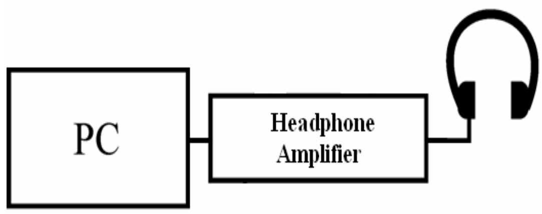

Figure 1: Measuring set

The two-component complexes were generated by a computer and passed through the headphone amplifier on headphones. The stimuli were composed of 200 Hz harmonics (400 Hz, 600 Hz …) for the residue phenomenon experiments. The frequency range was 400-2000Hz. In the first pitch shift experiment, the same harmonics were shifted up for Δf=30 Hz (430 Hz, 630Hz …) and the frequency range was 430-1630 Hz. Every two-frequency complex contained vicinal components: h_2 and h_3, h_3 and h_4, …, h_9 and h_{10}. The lower components had a constant level of 50 dB SPL, the level of the higher components was changed in the range of 5 - 50 dB SPL in 5dB step. Six measurement series were made for each subject. The level of higher components was increased from 5 up to 50 dB SPL in the odd series, and decreased from 50 down to 5 dB SPL in the even ones. During the adjustment procedure, listeners had to specify the pitch of two-frequency stimuli. The following pattern of presenting the complexes was applied in the experiments: the presentation time was 4 sec., followed by a 0.5 sec. break, and after that, the listeners had to specify the pitch of stimuli. The standard deviation value was used as a factor of certainty in defining the pitch [12]. The results of the experiment are shown in chapter 5.

3 The time structure model of excitation

According to this model, the residual pitch is defined on the basis of the time period between the auditory neuron impulses, which are generating for subtle maxima near the envelope maximum in the time structure of a sound wave. The reverse of this time gives the pitch of complex tones. The time structure of the stimulus is shown in Figure 2 and Figure 3. The time between the maxima in Figure 2 (Δt=5ms) determined the residual pitch corresponding to f=200 Hz. The perceived pitch changed when the pitch shift Δf=30 Hz was applied (Figure 3). The time between the maxima changed (Δt=4,7ms) and determined a new pitch corresponding to f=212 Hz. It can be said that the time structure model allowed the defining of the residual pitch in the two-component complexes.

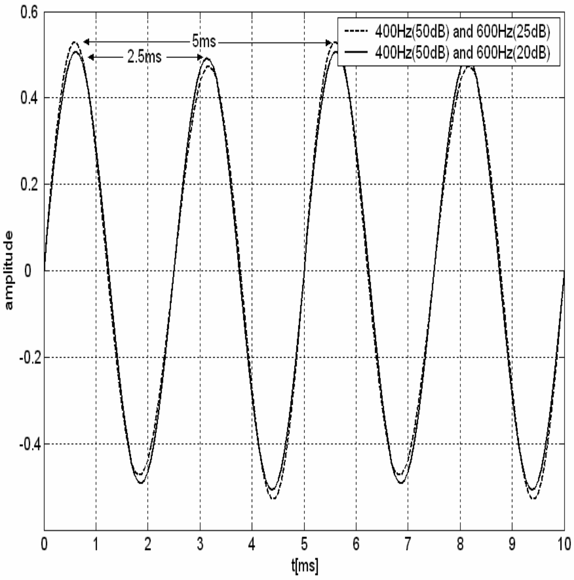

Figure 2: Time structure of harmonic stimuli (harmonic situation)

This model also explains the threshold formation mechanism of this phenomenon. If the level of higher harmonic L_2 = 20 dB SPL, the time period between the maxima in the time structure of stimulus is t =2,5 ms in Fig. 1 (or t =2,3 ms in Figure 2) and determines the pitch corresponding to f =400 Hz (or f =432 Hz in Figure 2). If the level increases to L_2 = 25 dB SPL, the time structure has a different periodicity. This causes amplitude differences and the appearance of a new envelope in the time structure. Apparently, the maxima near the envelope maximum are taken into consideration in the residual pitch perception process.

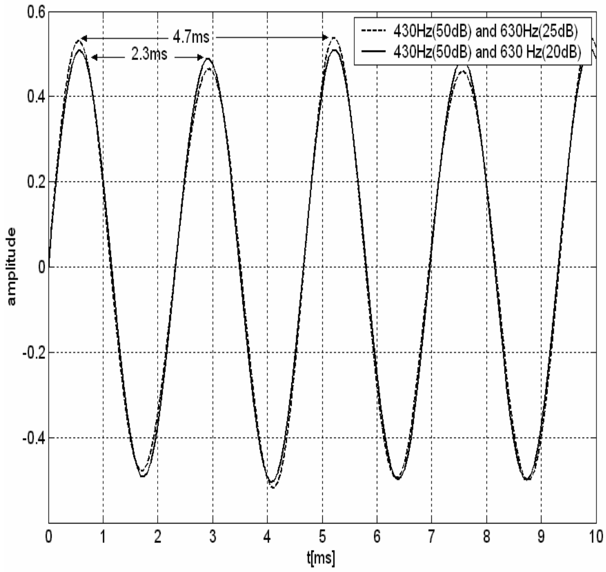

Figure 3: Time structure of harmonic stimuli (inharmonic situation)

The time between those maxima for L_2=25 dB SPL is t =5 ms in Figure 1 (or t =4,7 ms in Figure 2), which determines a new pitch (residual) of stimulus corresponding to f=200 Hz (or f=212 Hz in Figure 2). The introduced results concern h_2 and h_3 harmonics exclusively. The results for all harmonics with the range of frequency 400-2000 Hz are shown in chapter 5.

4 The auditory model

To confirm the hearing results, the model of a hearing system was also introduced. M. Slaney and R. Lyon [10, 11] created this model, which is based on the Licklider duplex theory of pitch perception [12]. The model performs autocorrelations on the activity of groups of stimulated nerve fibres (the hearing channels) and aggregates these autocorrelation functions (ACF) to produce a summary autocorrelation function. The maximum of the summary autocorrelation function provides the information regarding the perceived pitch expressed in a time period value.

After using the model and executing a suitable procedure, the pitch of the stimulus is received. For a two-frequency stimuli: f_1=400 Hz, f_2=600 Hz, L_1=L_2=50 dB SPL as a result model calculate pitch corresponding to f=200 Hz. After adding the pitch shift Δf=30 Hz, the new stimuli is created which is composed of the following components: 430 Hz, 630Hz … After using the model for f_1=430 Hz, f_2=630 Hz, L_1=L_2=50 dB SPL as a result, a new pitch is received which corresponds to f=213.2 Hz. The set of results is shown in chapter 5.

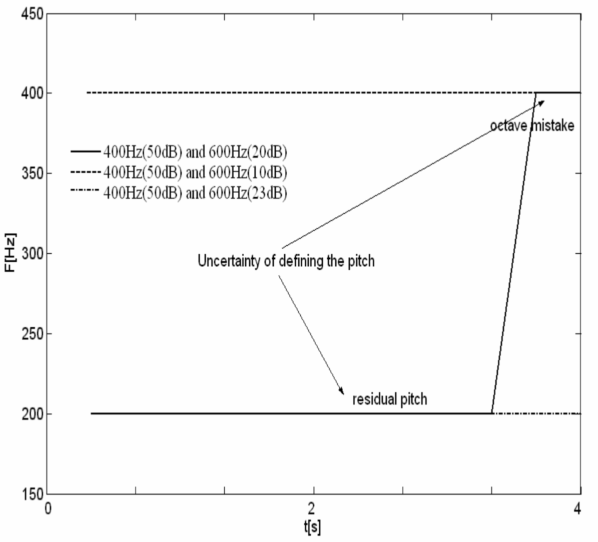

Figure 4: Auditory model response (harmonic situation)

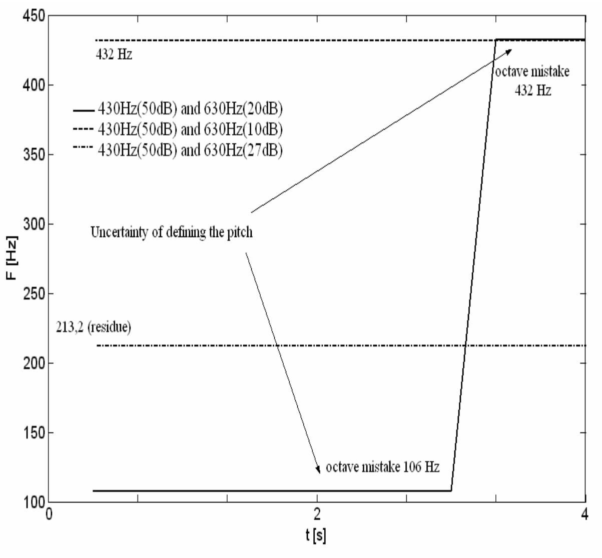

Figure 5: Auditory model response (inharmonic situation)

This modelling also shows the threshold formation mechanism of the residue phenomenon. Figure 4 shows that the level of higher harmonic in two-component complexes has an influence on pitch perception. When the level of higher harmonic is L_2=20 dB, an ambiguous pitch occurs. Because of calculating the pitch as a frequency of repetition for the ACF function, it can be observed as two values: 400 Hz and 200 Hz. The same situation was observed in psychoacoustics experiments because the subjects defined the pitch as an octave of 200 Hz. This effect is called an octave mistake. For L_2 =23 dB, uncertainty disappears and the model shows a clearly defined residual pitch corresponding to f=200 Hz. For inharmonic stimuli (after adding the pitch shift), the same behaviour can be observed. For L_2=20 dB, the model shows two perceived pitches corresponding to 106 and 432 Hz. After crossing L_2 =27 dB, one clear defined pitch is perceived, which corresponds to the f=213.2 Hz. On the basis of these results, the level of higher harmonic, for which the residue phenomenon is unambiguously perceived, one can consider as the threshold of residue phenomenon.

The model shows the uncertainty of defining the pitch accurately. However, the moment at which the pitch changes cannot be interpreted as the time after which the subjects perceived the new pitch. This is the result of the calculating procedure, which operates on time domain. The pitch change should be interpreted as a possibility to perceive two different pitches.

The behaviour of the threshold as a function of frequency is shown in the next chapter.

5 Results and discussion

The time structure model and the auditory model confirm the hearing experiments results. There is evidence (Fig. 4) to the existence of the residual pitch in the two-frequency stimuli. The first pitch shift effect is also perceived (Fig. 5) in these stimuli. There are small differences between the models and hearing experiments, but it is the influence of uncertainty in defining the pitch by the listeners. These differences are smaller than the standard deviation (<5 Hz) of pitch up to the ninth harmonics for harmonic stimuli and up to the seventh component for inharmonic, so the uncertainty of defining the pitch by the listeners does not influence the obtained results. It is very important that the models also gave the same results. The hearing experiments also confirmed that the tenth harmonic is the limit to the existence of residue phenomenon in two-component complexes for harmonic situations. Above this component, the standard deviation of defining the pitch exceeds 5 Hz, and it is shown that this phenomenon becomes ambiguous. It is a consequence of the virtual pitch timbre changes.

The experiments were conducted up to the seventh component for the inharmonic situation. The auditory model gives unexpected results above that number of component, which cannot be connected with listener's sensation. Because of the limit of residual pitch existence (tenth harmonic) in two-frequency stimuli, the first pitch shift effect also begins to disappear above the eighth component. In consequence, the listeners start defining the pitch corresponding to 200 Hz. This is the residual pitch for harmonic situation. It can be said that the eighth component is the limit of existence of the first pitch shift effect in two-component complexes.

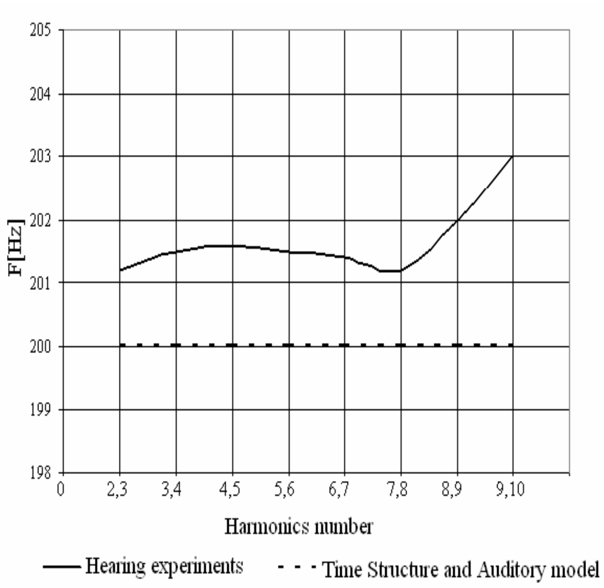

Figure 6: Relationship between two-tone complex pitch and the two presented components number (harmonic situation)

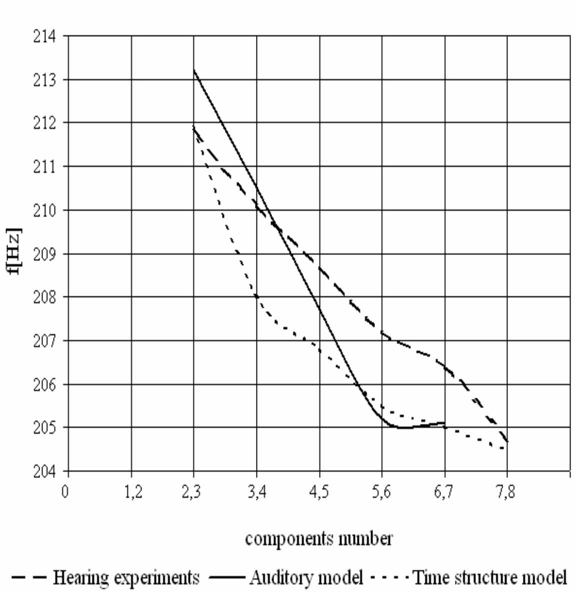

Figure 7: Relationship between two-tone complex pitch and the two presented components number (inharmonic situation)

As was assumed previously [6,7], the influence of timbre perception is greater when we analyze the behaviour of the threshold of residue phenomena. It is very important that the hearing experiments and auditory modelling results gave the evidence that the threshold of residual pitch perception increases along with the frequency increase (Fig. 6 and 7). These differences are the results of residual pitch timbre changes. When the frequency increases, the listeners had to increase to a higher harmonic level to be certain of residual pitch perception. The auditory model does not respect the uncertainty of defining the pitch by listeners, and this created the differences in the results between the hearing experiments and auditory modelling.

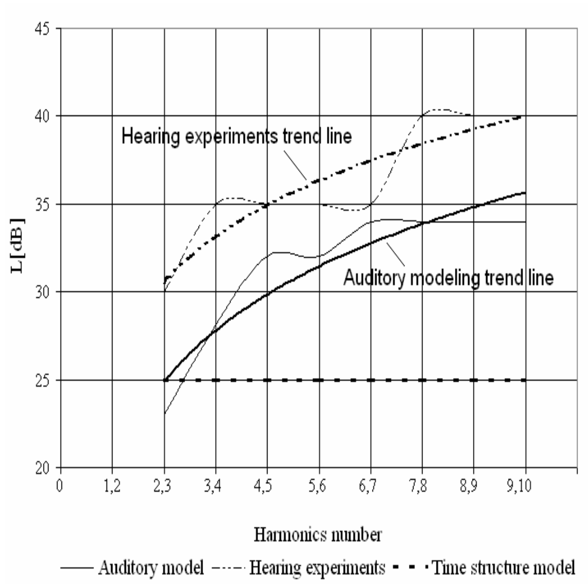

Figure 8: Relationship between the higher harmonics level that caused the residue phenomenon and the two presented components number (harmonic situation)

The time structure model also gave evidence to the existence of the threshold of residue phenomenon and first pitch shift effect, however, on the basis of that model, it is impossible to show an increase of the residue threshold in the frequency domain. It is a consequence of only getting the information from the time structure of stimuli. The appearance of a new periodicity, which has an influence on the appearance of the residual pitch, cannot be connected with the shift of the two-component complexes in the frequency domain.

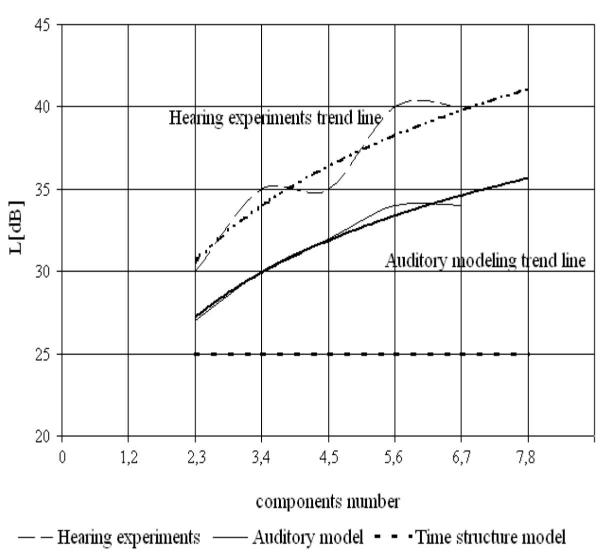

Figure 9: Relationship between the higher harmonics level that caused the residue phenomenon and the two presented components number (inharmonic situation)

6 Summary

- There is a clearly defined residual pitch and first pitch shift effect perception threshold in two-frequency stimuli.
- The threshold of the residue phenomenon and first pitch shift effect increases with the frequency increase of the two-frequency stimuli components.

References

[1] F. Wightman, “Pitch Perception: An Example of Auditory Pattern Recognition”, *Auditory and visual pattern recognition.* New Jersey, pp.3-25 (1981)

[2] A. Houtsma, “Effect of Signal Envelope on the Pitch of Short Complex Tones”, *JASA*, 81, pp.439-444 (1987)

[3] J. Goldstein, “Auditory Nonlinearity”, *JASA*, 41, pp.676-689 (1967)

[4] J. Goldstein, Y. Kijang, “Neural Correlates of the Aural Combination Tone $2f_1–f_2$, *Proc. IEEE*, 56, pp.981-992 (1968)

[5] J. Schouten, R. Ritsma,B. Cardozo, “Pitch of the Residue”, *JASA*, 34, pp.1418-1424 (1962)

[6] A. Mielczarek, “The Threshold Researches of Residue Phenomenon in Two- and Three-Tone Complex”, In *Proceedings* of 10^{th} *Symposium New Trends in Audio and Video,* Wrocław (Poland), pp.111-118 (2004), (in Polish)

[7] A.Mielczarek, “Residual Pitch of a Two-Component Complexes”, In *Proceedings of Subjective and Objective Assessment of Sound.* Poznań (Poland) (2004), (CD)

[8] J. Schouten, Proc. Kon. Akad. Wetenschamp. 43, pp. 991-999 (1940)

[9] M Slaney, R. Lyon, “On the Importance of Time—A Temporal Representation of Sound”, *Visual Representations of Speech Signal,* Sussex (England) (1993)

[10] M. Slaney, “Auditory Toolbox Version 2”, *Apple computer technical report #1,* Interval Research Corporation (1998)

[11] J. Licklider, “A Duplex Theory of Pitch Perception”, *Experientia 7*, pp.128-133 (1951)

[12] J Żera, “The Role of Harmonic Components in the Pitch Formation Process”, *PhD Thesis*, Wrocław (Poland) (1989), (in Polish)