ARCHIVES OF ACOUSTICS
23, 1, 67–85 (1998)

SYMMETRY AND ASYMMETRY AS A PHYSICAL AND PERCEPTUAL FEATURE OF THE COMPLEMENTARY PAIR OF BEATING SINUSOIDS
PART II. PITCH MATCHING EXPERIMENTS

L. RUTKOWSKI, A. WICHER

Institute of Acoustics,
Adam Mickiewicz University
(60-769 Poznań, ul. Matejki 48/49)

Two experiments were designed and performed in that the pitch of complementary pairs of beats signals, referred to as SL and SH, was evaluated. In experiment 1, the changeable signal consists of a pair of signals of equal sound pressure level and of the same separation of frequencies as this between SL and SH. Experiment 2 was made in the AM mode with a changeable signal whose amplitude envelope is identical with that of the SL/SH pair; the carrier frequency could be varied during the experiment. The results of the experiment 1 do not confirm the asymmetry in the evaluation of the pitch of the pairs of SL/SH signals with respect to the average frequency of the beats, a feature that has been reported earlier in the literature. On the order hand, the outcome of the experiment 2 point to certain asymmetry, moreover, the data reveal a large pitch difference of the SL/SH pairs; this difference (for which no explanation has been offered yet) is often much larger than that measured in experiment 1. The observed asymmetry, which is due to frequency variations determined with the frequency envelope (Part I) resulting from the broadband character of the beatings, does not quite justify the range of asymmetries reported in literature.

1. Introduction

From the physical properties point of view, beatings are the signal consisting of concurrent, synchronous variations of the sound pressure level (SPL) and frequency envelope [15]. The problem of perception of the beatings, regarded as an effect of superposition of a pair of tones of slightly different frequencies, mostly reduces to the definition of the resulting pitch and to the discrimination of the two tones depending on their SPL and frequency. So far, there is no doubt that a listener perceiving the pitch localizes it in the frequency range bound up by the components of the two-tone complex of beats. It has also been found that the possibility of discrimination between the two tones is ruled by their separation on the frequency scale relative to the critical bandwidth. All the other observations concerning the beats are to some degree ambiguous and the conclusions are often speculative. It seems that the problem remains in the finding of appropriate physical measures, either new ones or those being a combination of already known parameters, that should well correlate with the results of the perception of beatings. No model is available yet that could unambiguously relate a set of physical parameters and the perceived pitch of signals whose both envelope and frequency are variable in time. Some authors opt for a model of perception based on the time domain representation of a signal, others prefer a spectral mechanism of perception. Therefore, the problem arises what is more appropriate: to search the physical parameters relevant to the time domain realization of beatings, or investigating those associated with their spectral properties? IWAMIYA maintains that his experimental findings [7, 8] support the model according to which the perception of the sound envelope and frequency changes occurs independently in the process of hearing. But, he does not prompt to either the spectral or time-domain model. On the other hand, JENKINS *et al.* [11] highlight the time-variable attributes, namely the fine changes of temporal structure, which are important in pitch perception.

Considerable effort has been reported on investigations dealing with the perception of the amplitude and frequency modulated sounds. According to the arguments brought out in Part 1 [15]; the signal of beats can also be viewed as a sound featuring these two kinds of modulation. Therefore, references on the loudness and pitch perception can also be helpful in the interpretation of the phenomena observable for the beatings.

Certain difficulty in perceiving the signal frequency changes arises from the fact that the sound pressure level change itself creates an impression of frequency modification. STEVENS has demonstrated [17] that the pitch of low frequency signals decreases with the increase of SPL; within the medium range frequencies, the SPL modifies the pitch although slightly, while the pitch of the high frequency signals rises with the increase of SPL. Pitch changes described above are known as the *Stevens rule*. Similar conclusions have been reached by VERSCHUURE *et al.* [19] who analyzed the published data. These data had evidenced a large differentiation in the pitch assessment made by various listeners and the difficulty in obtaining reproducible results. It was experimentally shown that the *Stevens rule* is not a sufficiently accurate measure of the pitch changes. The absolute value of the pitch variations caused by SPL changes is small within the frequency limits of 1 to 2 kHz. This value increases with increasing as well as with decreasing frequencies. For an individual listener these changes may be positive, negative or do not exhibit monotonicity. However, only the data averaged over a large number of listeners corroborate the *Stevens rule*. Except for very high or very low frequencies, the perceived pitch changes can be insignificant in many cases. Verschuure explains the discrepancies between the results as due to the differences in the experimental procedures used by various authors.

Subsequently, IWAMIYA *et al.* [7, 9] states that the pitch perceived under simultaneous AM&FM modulation, depending on the phase shift between the two modulating functions, is determined by the amplitude weighted frequency variations. He assumes a $\{A_1(t)\}^\alpha$ amplitude envelope, where $A_1(t)$ describes relative changes of the amplitude envelope at the amplitude modulation index equal 1 and α is an increment which modifies physical changes of the envelope in order to obtain a loudness modulation function. The frequency shift P (measured from the carrier frequency) to the so called "principal pitch" localized by listeners is given by formula [7]

$$P = m^2 \int_0^T C(t)\{A_1(t)\}^\alpha dt \Big/ \int_0^T \{A_1(t)\}^\alpha dt . \tag{II.1}$$

The increments $\alpha = 0.3$ to 0.7 fit well to the pitch measure proposed by Iwamiya. He explains the differentiation of the values, occurring with individual listeners, by a personal ability of perception of amplitude modulation. What he also has noted is that the regression curves of the perceived pitch values vs the phase shift, mentioned above, exhibit a displacement, the latter being larger at the modulation index $m = 1$ than at $m = 0.71$ (he concludes that this is caused by envelope changes). Following Iwamiya, these displacements may be caused by timbre differences for AM+FM sounds the and sinusoidal tone. In majority, they are negative and are discriminately perceived by listeners (the pitch values are reported to be lower than those predicted by the accepted model). Iwamiya attributed the differentiation of those P shifts to a differentiated effect of the sound pressure level on the perceived pitch.

HARTMANN [6] has studied the influence of the amplitude envelope on the pitch of sinusoidal tones. He discovered that the pitch of a sinusoidal tone of an exponentially decaying amplitude envelope is higher than a tone of the same frequency but of rectangular envelope. He used signals of constant change of the envelopes rate of 1000 dB/s. The experiments have been made at selected frequencies in the frequency range from 412 to 3300 Hz. For three listeners of the group of 15 people, he has noted some differentiation of the asserted changes of the pitch. Averaging the results for these 3 listeners, he obtained the best fit value of the frequency shift equal to 16 Hz for decaying tones; this result does not depend on the tone frequency. It proves that only the envelope changes determine the pitch shift, while the frequency values of the separate tones have no effect on the observed shift. If the results are analyzed in detail for each particular listener in Hartmann's experiment, it seems that this conclusion is not entirely justified. Hartmann is attempting to find a set of physical attributes of a tonal signal of decremented amplitude envelope which properly correlate with the perceived shift of pitch. Discussing the time evolution of the spectrum, he quotes the observations of some listeners whose experience was that the pitch of the exponentially decaying tone appeared to rise in the process. Concluding, Hartmann says that the shift of the pitch does not depend on frequency. However, he underlines the speculative character of this reasoning and emphasizes the need of further investigation.

ROSSING *et al.* [14] run a similar experiment, but they used different initial reference levels, different rates of the envelope change, ranging from 500 to 8000 dB/s, with both incremented and decremented amplitude envelopes. They established unambiguously that both the envelopes, the decaying and the rising ones, produce a comparable pitch increase. This increase grows up with still higher frequencies. Such result eliminates the perception model based on the "periodicity in neural firings" predicting opposite pitch changes discriminating between up- or down-trends of the envelope amplitude. The authors come to the conclusion [14] that the experiments implicate that the envelope-dependent change of the pitch would not directly be caused by the envelope, but, preferably, may be due to a process the quantity of which, in turn, depends on the envelope; for example the averaged sound pressure level.

SCHOUTEN *et al.* [16] observed that the majority of natural sounds, like speech and sounds of musical instruments, feature a harmonic spectral structure; according to them, the joint perception of a number of spectral components, the so called *residuum*, determines the sounds pitch. There is however, as they remark, certain ambiguity in assessment of the pitch which may be due to the influence of the fine structure of the sound amplitude envelope. According to them, the pitch assessment is performed within the hearing system in the time-domain mode (*via* technique of delay lines), whilst the spectral model is less plausible.

RITSMA *et al.* [12] studied the pitch under conditions of *quasi-frequency modulation*, discovering its dependence on the modulation depth (*index of modulation*, crucial in determining the depth of the changes of the amplitude envelope, too). The measured values of the pitch were not correlated with the period of the amplitude envelope of the signals, but calculated assuming that the separation measured between two adjacent peaks of positive polarity of *temporal fine structure*, in the vicinity of the neighbouring maxima of the amplitude envelope, can be a measure of pitch.

JENKINS [11], summing up the discussion on the perception of the sound pitch, sound timbre and loudness, favours the model performing a broad-band spectral analysis inside the auditory system including a detection of envelope and detection of periodicity (based on a delay line). Such a mechanism of perception is featuring a short time response. Accepting that the spectral model is preferred in the perception processes, Jenkins argues that, on physical grounds, it is not possible to accomplish an exact frequency determination with the Fourier transform (pitch evaluation) while maintaining a fast time response.

ZWICKER [20] comments on the perception thresholds of frequency differences of tones that are essential to our analysis. Zwicker has found that the perception of frequency differences are much better for tones of constant amplitude than for a very narrow band noise, centred at the same frequency, which sounds like a tone of randomly variable amplitude. For example, the thresholds of the frequency difference perception for a narrow - band noise of carrier frequency of 1500 Hz and 10 Hz bandwidth are 6 times greater than for a tone of the same frequency. In particular, according to Zwicker, the diversions in the sensitivity of the hearing systems to perceive frequency differences are controlled by the rate and non-periodicity of the amplitude changes.

ROEDER [13] has found that the response of the hearing system to the signal of beatings follows the resultant waveform of the beatings. The ear "is not aware" of the perceived sound resulting from addition of the components. Simultaneous stimulation of the closely situated regions on the basiliar membrane is responsible for the resultant impression of a sound pitch. Such stimulation of the basiliar membrane involves overlapping of adjacent regions that effects indirectly the intermediate pitch impression. The SPL changes during beatings are perceived as a modulation of loudness.

JEFFRES [10] isolated fragments of beating sinusoids within the maximum and minimum of the amplitude envelope and could distinguish the pitch differences. For the pair of signals, in which the lower frequency tone is of smaller SPL value, Jeffres found an increase in frequency, when passing from the minimum to the maximum on the envelope, and the contrary for a pair in which the higher frequency tone is the stronger one. In the case of two signals of equal sound pressure levels, Jeffres pointed out the instantaneous phase reversal at each beats minimum. The phase reversal is of a jump wise character end involves an infinite frequency shift, but of zero duration.

The results of the referenced perception studies as well as the arguments regarding the mechanism of pitch perception only prove the complexity of the phenomenon of beats leaving many questions unresolved. Therefore, it is necessary to search for other properties of the beats signal that could contribute to a better understanding of the observed relations. The perception analysis, which will be pursuit in this part of the study, is confined to a situation in that the frequency separation of the components is so small that we can not resolve two tones of unique frequencies, but we may perceive a resultant (average) sensation of pitch or its change (beat components fit within the critical band). Attention will be focused on the responses of listeners who may testify either symmetry or asymmetry of pitch perception for the complementary pairs of beating signals. Two experiments and the results reported previously [2], provide a basis for the analysis of the problem. The interpretation of the results will also be made with reference to Part I in the scope relevant to the predefined physical parameters of beats.

2. Theorems related to the experimental procedure

Dai [2] reported the results of the perception of pitch of beatings by 3 listeners. His examination of the pitch was made in two combinations denoted SL and SH (see Fig. 1) for the following set of "fixed signals":

- two-tone complex with frequencies f_L = 480 Hz and f_H = 520 Hz (average frequency = 500 Hz),
- two-tone complex with frequencies f_L = 960 Hz and f_H = 1040 Hz (average frequency = 1000 Hz).

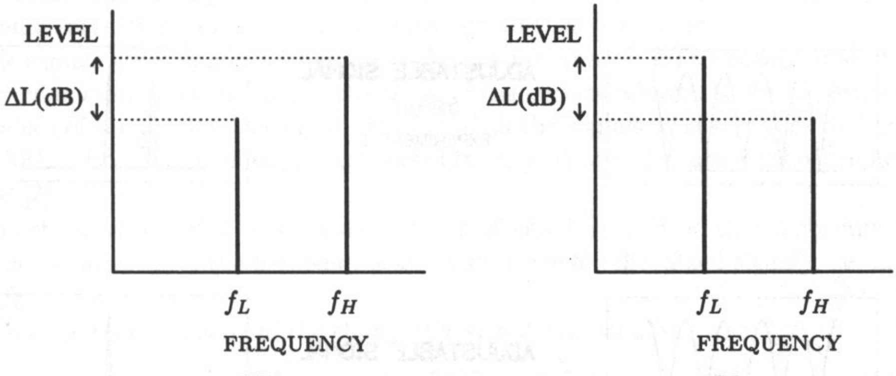

Fig. 1. Complementary pairs of the 2 tone complexes SH (stronger high) and SL (stronger low) used both in the study of Dai [2] and in the present experiments 1 and 2. The ΔL is the sound pressure level difference of the beats components f_L and f_H.

For each combination, the author checked three relative difference levels ΔL, i.e. 2.4, 3.4 and 4.4 dB.

The signal of variable parameters, the so called "adjustable signal", was a two-tone complex consisting of tones of equal sound pressure levels and a frequency difference equal to the separation of the beating frequencies. The level of the acoustic pressure was 65 dB SPL for all the signals applied in Dai's experiment, i.e. for the fixed signal and the adjustable one. The duration of the signals was 250 ms which corresponds to 10 beats at 500 Hz and 20 beats cycles at a frequency of 1000 Hz. Dai has performed his experiment by a pitch matching procedure. The listeners changed the adjustable signal frequency till perceiving the adjustable and the fixed signals as equaled in their pitches. Dai detected the occurrence of some asymmetry in the listeners valuation of the pitch for the complementary pairs of beating signals. With the aim to understand the possible causes of this asymmetry, in the present study two separate experiments were made in which the signal parameters assumed similar, if not identical, values to those reported by Dai. Experiment 1 (Sec. 3) was in principle a repetition of the experiment performed by Dai, although somewhat new methods were introduced. Experiment 2 (Sec. 4) consisted in applying as the signal of variable parameters, an amplitude modulated signal whose envelope was calculated as for the beats (the fixed signal). Due to this, the envelope changes of the two signals were identical in the course of the experiment. Basic differences regarding the signal parameters of the experiments reported in this study are shown in Fig. 2.

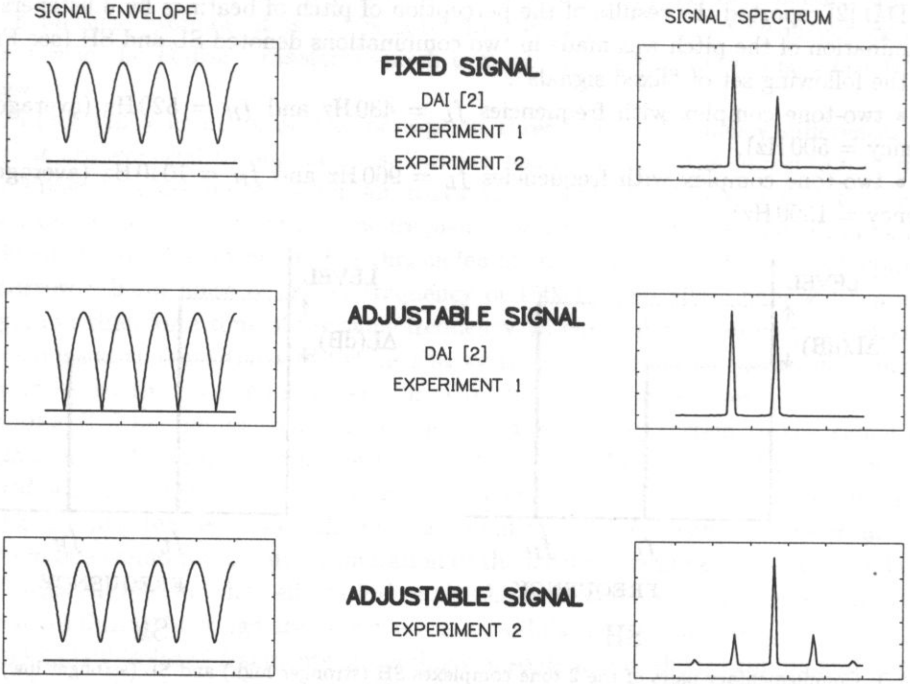

Fig. 2. Illustration of the basic signal parameter differences according to DAI [2] and to the experiments 1 and 2.

3. Experiment 1:
pitch matching with equal SPL beating sinusoids

A. Procedure

The method consisted in transmitting pairs of signals as a random sequence to the listener. The parameters of one of the signals, the "fixed signal", were kept unchanged during the whole experiment. The parameters of the other one, the "changeable signal", were regulated during the experiment following the two-alternatives forced choice (2AFC) paradigm with an adaptative procedure one – down, one – up. The listener was required to decide, for each pair of the randomly launched sequence, the signal (either the first or the second one) that could be heard at a higher pitch. The message whether the listener's answer was correct or wrong was displayed each time on the monitor screen. The test series was terminated after a sequence of 40 pairs of signals. For each test at least 8 turnpoints were obtained. An intermediate value of matched frequency was calculated as an arithmetic mean of the tone's center frequency of the last 8 turnpoints. Each listener took part in 5 such sessions. The final value of frequency – f_{adj} was an average taken from those five sessions. The measurements were performed in special sound-proof rooms properly insulated from the outside noise.

B. Stimuli

The test signals were digitally generated with an instrumental setup consisting of an Array Processor DSP32C connected to a 16-bit digital-analogue converter with optical fibres (lower noise level). Next, the signals were filtered with a low-pass filter of cut-off frequency 8 kHz and a slope of trailing edge of the filter characteristics of 90 dB per octave. The described setup for the generation of the signals (Tucker Davis Technologies – USA) and the experiment sequencing was controlled with a PC computer. Each signal lasted 1000 ms including the cosine rise and decay time, both 50 ms. The time period between a successive pair of signals depended on the listeners choice.

The signals were monaurally submitted to the listeners by DT48 headphones of frequency characteristics equalised in the range of frequencies relevant to the experiments. The values of the resultant acoustic pressure of all the signals were the same and equal 65 dB SPL. Also, the remaining parameters of the signals were the same as those reported by DAI [2].

The changeable signal was a two-tone one of equal SPL ($\delta = 1$) components and the same separation on the frequency scale as that one for the "fixed signal", i.e. $\Delta f = f_H - f_L$.

The amplitude envelope of the changeable signal was equal to

$$e(t) = x_L\sqrt{2}\sqrt{1 + \cos 2\pi\Delta f t}\,. \tag{II.2}$$

To simplify the formula (II.2), the cosine rise and fall of the signal was omitted. The complex instantaneous frequency (CIF) envelope of this signal varied according to the following dependence:

$$|\mathrm{CIF}(t)| = \sqrt{\left[\frac{\Delta f \sin(2\pi\Delta f t)}{2(1+\cos(2\pi\Delta f t))}\right]^2 + \left[\frac{\Delta f}{2} + f_{L\,\mathrm{adj}}\right]^2}\,, \tag{II.3}$$

where $f_{L\mathrm{adj}}$ is the lower frequency value of the two-tone of equal amplitudes.

C. Subjects

Three listeners participated in the measurements, all with audiometrically normal hearing. Their age was between 21 – 24. One of the listeners was the author of this study (AW). The other one, a female (MK), had some musical education, while the third one, (PR), was a musician. Each of the listeners was instructed and practiced before the experiments; the aim was to attest their understanding of the objectives. The listeners had good practice in this kind of experiments.

D. Results

Tables 1 and 2 contain the mean values (averaged over 5 series of measurements) of the center frequency of the changeable signal of pitches matched by the listeners to the pitch of the SL and SH pairs of average frequencies of 500 Hz (Table 1) and 1000 Hz (Table 2), respectively. Standard deviations of the matching frequencies are given in brackets. δ denotes the ratio of amplitudes of the SL and SH signals for the given ΔL (level difference) values determined in decibels. Frequency values listed in tables 1 and 2 equal the arithmetic average of the frequencies of the beating tones matched with respect to the signal pitch, i.e. $(f_L + \Delta f/2)_{\mathrm{adj}}$.

Table 1. Experiment 1. The center frequencies of the equal intensity adjusted signal, (the pitch was matched to the pitch of either the SL or SH signal) and the corresponding standard error. Data for $f_{\mathrm{av}} = 500$ Hz.

	500 Hz					
ΔL	2.4 dB		3.4 dB		4.4 dB	
	SL $\delta = 0.759$	SH $\delta = 1.318$	SL $\delta = 0.676$	SH $\delta = 1.479$	SL $\delta = 0.603$	SH $\delta = 1.66$
AW	494.9 (0.3)	506.3 (0.3)	492.6 (0.4)	507.4 (0.2)	490.1 (0.2)	509.6 (0.2)
MK	492.3 (0.5)	507.7 (0.1)	490.5 (0.2)	510.3 (0.5)	488.7 (0.3)	510.2 (0.2)
PR	493.4 (0.2)	504.8 (0.2)	493.8 (0.4)	507.5 (0.3)	492.9 (0.2)	508.9 (0.3)

Table 2. Experiment 1. The center frequencies of the equal intensity adjusted signal, (the pitch was matched to the pitch of either the SL or SH signal) and the corresponding standard error. Data for $f_{av} = 1000\,\mathrm{Hz}$.

1000 Hz						
ΔL	2.4 dB		3.4 dB		4.4 dB	
	SL $\delta = 0.759$	SH $\delta = 1.318$	SL $\delta = 0.676$	SH $\delta = 1.479$	SL $\delta = 0.603$	SH $\delta = 1.66$
AW	993.8 (0.6)	1007.9 (0.5)	990.4 (0.5)	1014.6 (0.6)	987.8 (0.4)	1019.5 (0.7)
MK	988.3 (0.6)	1010.4 (0.7)	984.5 (0.5)	1016.6 (0.6)	980.3 (0.5)	1021.6 (0.6)
PR	993.9 (0.5)	1006.8 (1.4)	992.9 (1.06)	1007.6 (0.3)	987.2 (0.8)	1010.3 (0.8)

4. Experiment 2:
pitch matching with amplitude modulated adjustable signal

A. Objective of the experiment

DAI [2] and FETH *et al.* [3] have found a characteristic asymmetry in pitch matching experiments when the pitch of the complex signal having two components of unequal levels, is compared to the arithmetic average of the frequencies of the two-tone SL and SH pairs. The amplitude envelopes of the SL and SH signals are identical and, in principle, the only difference between the signals is the pattern of their instantaneous frequency variations; however the symmetry of the Fourier spectra around an average frequency is not corroborated by the symmetry of the perceived pitch values. It should be noted that in the experiments of Dai and Feth the adjustable signal for which $\Delta L = 0$ ($\delta = 1$) had an amplitude envelope different from those of the tested signals (fixed signals of the SL and SH pairs, compare Fig. 2). Therefore, in spite of the different mean frequency values of the fixed and adjustable signals (which is obvious, especially in the initial phase of the experiment), the signals were also different envelope amplitude (in each phase of the experiment). This subsidiary discrimination between the fixed and matched signal was not, most certainly, a circumstance that made the listeners task of matching the pitch of pairs of signals easier. For this reason, new modified experiments have been suggested, in that the changeable signal had exactly the same amplitude envelope as the fixed one. This idea was implemented by the technique of the amplitude modulated (AM) signal in that the envelope was calculated according to equation (I.5), hence in the same way as the envelope of the fixed signal of beatings. The carrier frequency of the AM signal was controlled by the listener during the experiment. Consequently, the signal to be matched (changeable signal) was of the form

$$x_{\mathrm{adj}}(t) = x_L\sqrt{1 + \delta^2 + 2\delta\cos 2\pi\Delta f t}\,\cos(2\pi f_{\mathrm{adj}}t), \tag{II.4}$$

(the cosine rising and decay are omitted for simplicity) where $\Delta f = f_H - f_L$ – amplitude modulation frequency, f_{adj} – carrier frequency of the AM signal whose pitch was matched to the fixed signal. According to Eq. (I.14), the frequency envelope of the signal (II.4) was

$$|\mathrm{CIF}(t)| = \sqrt{\left[\frac{\delta \Delta f \sin(2\pi\Delta f t)}{1+\delta^2+2\delta\cos(2\pi\Delta f t)}\right]^2 + [f_{\text{adj}}]^2}\,. \tag{II.5}$$

An example of the matching Fourier spectrum (II.4) is shown in Fig. 2.

B. Results

In Tables 3 and 4 average values of the carrier frequency of the AM signal (5 measurement series) are presented; the pitch was matched by the listeners (the same group which took part in the experiment 1) to the pitch of the SL and SH signals of mean frequencies 500 Hz and 1000 Hz, respectively, for Table 3 and 4.

Table 3. Experiment 2. The carrier frequencies of the amplitude modulated adjusted signal, (the pitch was matched to the pitch of either the SL or SH signal) and the corresponding standard error. Data for $f_{\text{ad}} = 500$ Hz.

	500 Hz					
ΔL	2.4 dB		3.4 dB		4.4 dB	
	SL $\delta = 0.759$	SH $\delta = 1.318$	SL $\delta = 0.676$	SH $\delta = 1.479$	SL $\delta = 0.603$	SH $\delta = 1.66$
AW	492.6 (0.4)	508.1 (0.4)	491.9 (0.6)	509.8 (0.7)	490.3 (0.4)	513.4 (0.8)
MK	491.1 (0.3)	506.8 (0.2)	488.3 (0.5)	509.4 (0.6)	486.5 (0.6)	509.4 (0.3)
PR	489.9 (0.4)	501.2 (0.3)	488.7 (0.5)	504.6 (0.5)	487.3 (0.5)	507.7 (0.3)

Table 4. Experiment 2. The carrier frequencies of the amplitude modulated adjusted signal, (the pitch was matched to the pitch of either the SL or SH signal) and the corresponding standard error. Data for $f_{\text{av}} = 1000$ Hz.

	1000 Hz					
ΔL	2.4 dB		3.4 dB		4.4 dB	
	SL $\delta = 0.759$	SH $\delta = 1.318$	SL $\delta = 0.676$	SH $\delta = 1.479$	SL $\delta = 0.603$	SH $\delta = 1.66$
AW	988.7 (0.9)	1019.1 (1.2)	984.5 (1.4)	1017.7 (1.1)	977.1 (0.6)	1023.9 (0.5)
MK	982.5 (0.5)	1012.7 (0.3)	981.3 (0.2)	1018.0 (0.2)	981.9 (0.3)	1022.2 (0.5)
PR	970.4 (1.1)	1008.7 (2.1)	965.9 (0.7)	1014.4 (1.05)	967.1 (0.5)	1014.5 (0.6)

5. Discussion

In Figs. 3 and 4 the results of the two experiments are displayed together with the earlier results obtained by DAI [2]. The data in Fig. 3 a–c correspond to the beats average frequency of $f_{\mathrm{av}} = 500$ Hz and the frequency separation $\Delta f = 40$ Hz with SPL differences of the beating tones equal to 2.4 dB (Fig. 3a), 3.4 dB (Fig. 3b) and 4.4 dB (Fig. 3c). The data in Fig. 4 a–c correspond to the beats average frequency $f_{\mathrm{av}} = 1000$ Hz and the frequency difference $\Delta f = 80$ Hz, with the SPL differences 2.4 dB (Fig. 4a), 3.4 dB (Fig. 4b) and 4.4 dB (Fig. 4c). In both figures (Figs. 3 and 4), the frequencies f_{adj} of the changeable signal matching the pitch of fixed ones are given. The values are referred to the arithmetic average of the fixed signals, so the vertical axes are in $f_{\mathrm{adj}} - f_{\mathrm{av}}$ [Hz] units. The positive values of such a calculated frequency correspond to the SL pairs, while the negative ones represent the SH signal pairs. Filled circles are the data points obtained by the listeners for the SL and SH pairs. The empty ones correspond to the calculated average values of the matching frequencies for the SH and SL pairs.

The displacement of these points from the dotted straight line ($f_{\mathrm{adj}} - f_{\mathrm{av}} = 0$) means that the responses of the listeners featured an asymmetry.

Additionally, the values of the envelope weighted average of instantaneous frequency EWAIF – dashed line, and intensity (square envelope) weighted average of instantaneous frequency IWAIF – solid line, diminished by f_{av}, are given. The two values are calculated for the frequency envelope given by Eq. (I.14) and the amplitude envelope described by Eq. (I.5) of the beatings frequency variations. These two quantities are compatible with the loss of symmetry for complementary pairs of the SL and SH signals (see Fig. 4 in Part I and Fig. 5 in Part II).

The results obtained for the beatings sinusoids of mean frequency 500 Hz in the experiment 1 point to the existence of an almost ideal symmetry of the two complementary pairs of the signals SL and SH, as reported by the listeners. The largest departures from symmetry amount to ±0.9 Hz with an average error of ±0.23 Hz. The mean deviation $\langle \Delta f \rangle_{\mathrm{Sub},\Delta L}$ for all the listeners (Sub) and all the values of level differences ΔL were equal to +0.1 Hz. DAI [2] reported a much larger deviation from symmetry for the same value of the average frequency of beatings. Namely, each of his listeners claimed frequency variations of positive sign from the symmetry related pattern. The largest reported deviation was +4.8 Hz at ±0.55 Hz accuracy. The mean deviation $\langle \Delta f \rangle_{\mathrm{Sub},\Delta L}$ amounted to +2.9 Hz, for the Dai's listeners. The major part of the data from experiment, corresponding to 500 Hz average frequency, justifies the argument that the pitch is well correlated to the intensity weighted average of the instantaneous frequency value [1, 2], although the listener's MK results are closer to the value of the envelope weighted average of the instantaneous frequency.

The application of AM as the changeable signal (experiment 2) with the same envelope as that of the beatings just investigated resulted in a wider differentiation of the listeners responses gained at 500 Hz average beatings frequency. The listener AW localizes the pitch of the SH signal closer to envelope weighted frequency value, while for the SL pair of signals at level differences of 3.4 and 4.4 dB, his results approach the intensity weighted frequency value. Most of the data on the SL signal obtained from

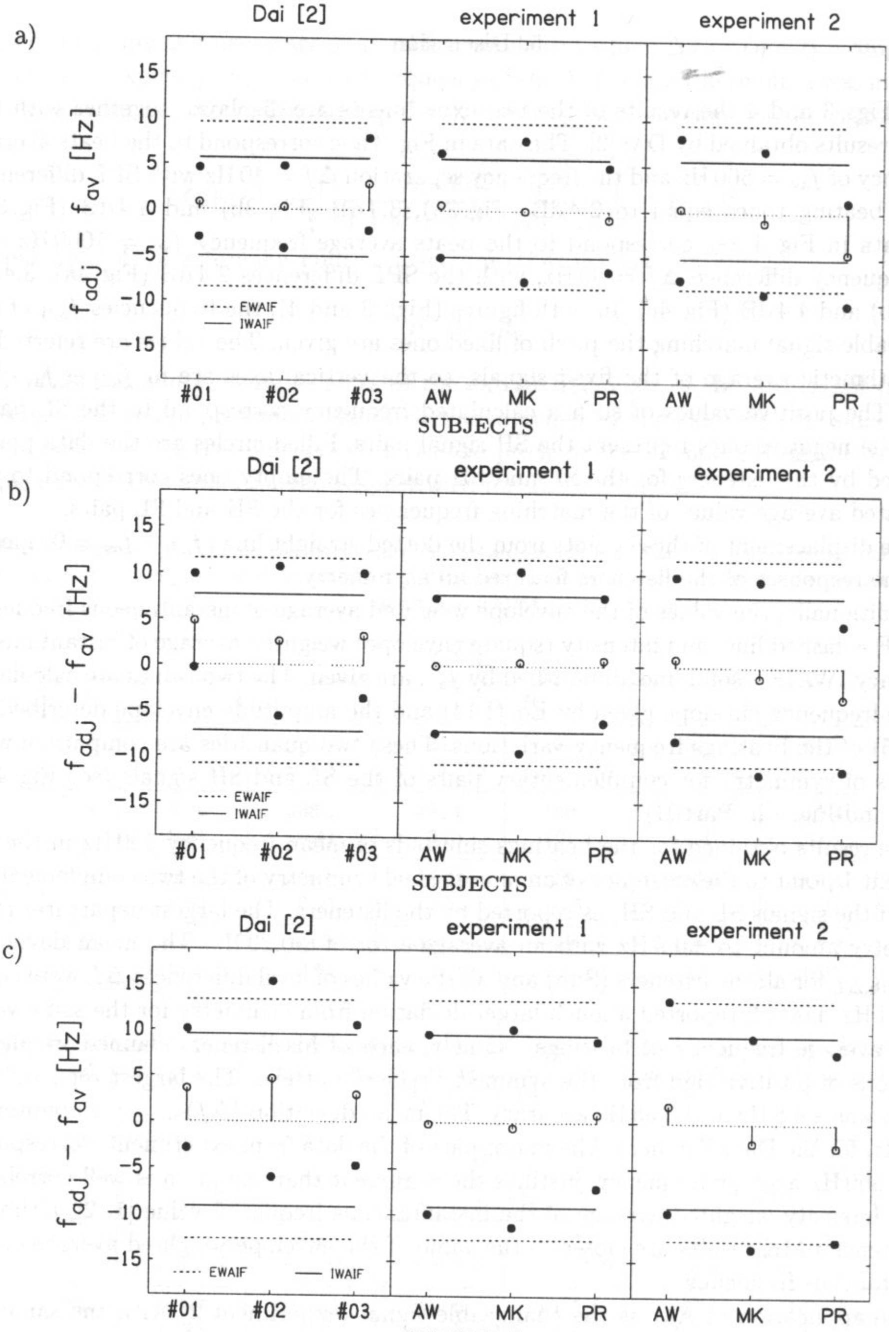

Fig. 3. Comparison of the results obtained by DAI [2] and in the experiments 1 and 2 for the average frequency of 500 Hz (closed circles) for the following SPL differences ΔL: a) 2.4 dB, b) 3.4 dB, c) 4.4 dB. Open circles – calculated average values for the complementary SL-SH pairs. EWAIF – calculated envelope weighted average of the instantaneous frequency. IWAIF – calculated intensity (squared envelope) weighted average of the instantaneous frequency.

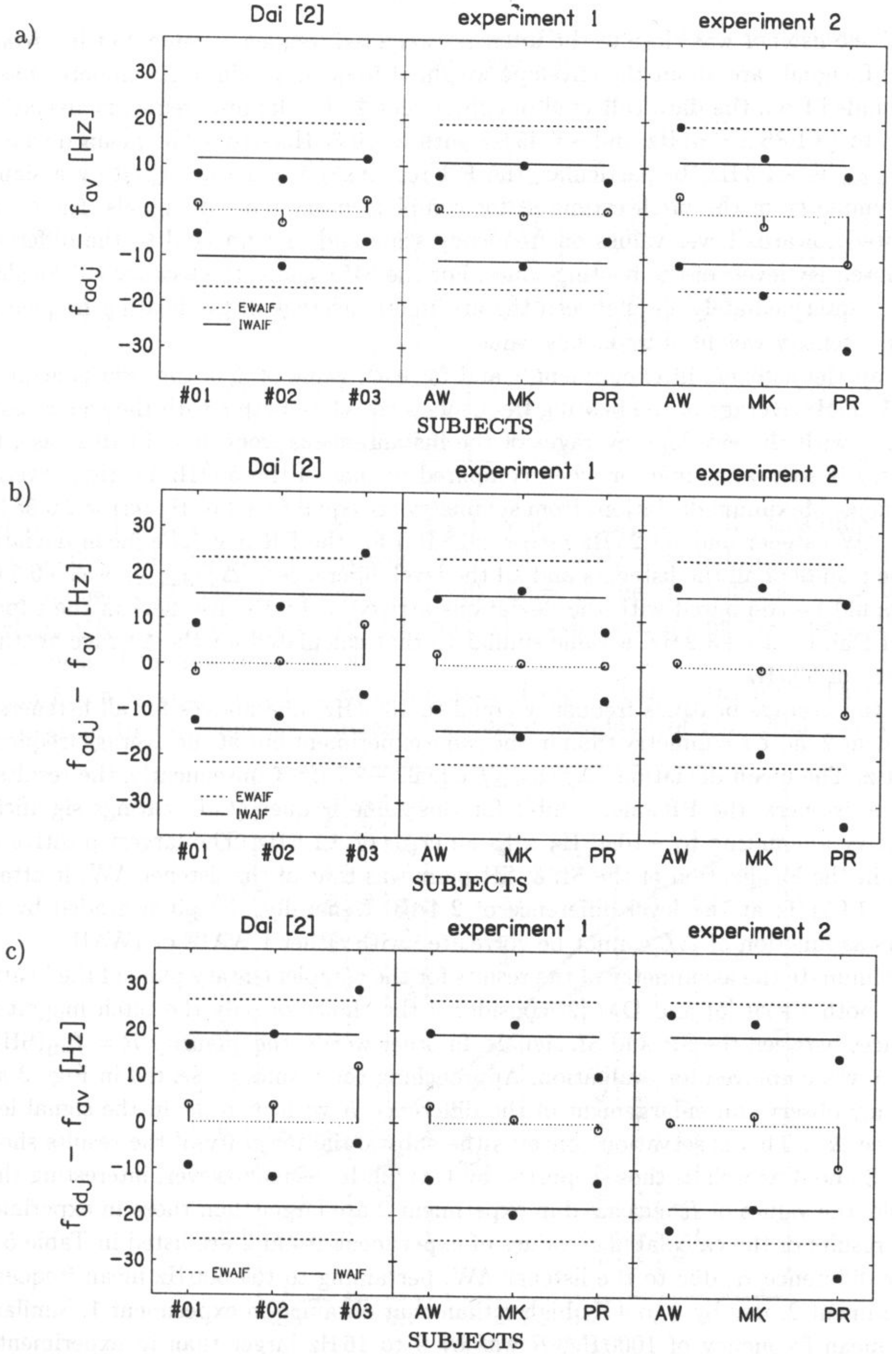

Fig. 4. Comparison of the results obtained by DAI [2] and in the experiments 1 and 2 for the average frequency of 1000 Hz (closed circles) for the following differences ΔL: a) 2.4 dB, b) 3.4 dB, c) 4.4 dB. Open circles – calculated average values for the complementary SL-SH pairs. EWAIF – calculated envelope weighted average of the instantaneous frequency. IWAIF – calculated intensity (squared envelope) weighted average of the instantaneous frequency.

the MK she-listener was close to the intensity weighted frequency value, but her results on the SL signals are about the envelope weighted frequency values. Asymmetry has to be concluded from the data collected in experiment 2. The largest frequency deviations amount to $(+1.85 \pm 0.6)$ Hz and -4.45 Hz with a ± 0.35 Hz error. The mean deviation $\langle \Delta f \rangle_{\mathrm{Sub},\Delta L}$ is -1.3 Hz. In particular, the PR responses are characterized by a significant asymmetry of the pitch estimates for complementary pairs of signals. His results are shifted towards lower values on frequency scale and unconnected to the difference in the intensity levels of the beating tones. For the SH signals, the listener PR localizes the pitch intermediately, i.e. between the arithmetic average of the beating frequencies and the intensity weighted frequency value.

For all the listeners in experiment 1 and for each value of ΔL, the results acquired at the 1000 Hz average of the beating frequencies correlate better with the square envelope than with the envelope averages of the instantaneous frequency. In this case, the symmetry is somewhat inferior when compared to that of the 500 Hz beatings average experiment. Maximum deviations from symmetry are equal to $+3.65$ Hz (error ± 0.55 Hz) for the AW listener and -1.25 Hz (error ± 0.8 Hz) for the PR one. The mean deviation over the results of all the listeners and all the level differences, $\langle \Delta f \rangle_{\mathrm{Sub},\Delta L}$ was $+0.8$ Hz. This should be compared with the deviations arrived at by the listeners in the experiment of Dai, i.e. to $+3.2$ Hz, a value similar to that calculated for the average beatings frequency of 500 Hz.

For the average beatings frequency equal to 1000 Hz, one notices for all listeners in experiment 2 more asymmetry than in the same experiment but at the average frequency of 500 Hz. The mean deviation $\langle \Delta f \rangle_{\mathrm{Sub},\Delta L}$ equals -2.7 Hz. Consequently, the results of only one listeners, the PR one, exhibit for this same frequency of beatings significant asymmetry amounting to -10.45 Hz with an error of ± 1.6 Hz. The largest positive deviation in the localization of the SL & SH pairs was that of the listener AW; it attains $(+3.9 \pm 1.05)$ Hz at the level difference of 2.4 dB. Generally, the pitch graded by the listeners as function of ΔL cannot be correlated with either EWAIF or IWAIF.

To eliminate the asymmetry of the results for the complementary pairs of the beating signals, both FETH [5] and DAI [2] considered the choice of only the pitch magnitude differences between the SH and SL signals. In other words, the quantity $R = f_{\mathrm{adj}}(\mathrm{SH}) - f_{\mathrm{adj}}(\mathrm{SL})$ was employed for evaluation. Approaching the results presented in Figs. 3 and 4, we may observe an enlargement of the difference R with increase in the signal level difference ΔL. This observation concerns the substantial majority of the results shown in Figs. 3 and 4 as well as those reported by DAI [2]. It seems however, interesting that, as a rule, the values of R measured in experiment 2 are larger than those in experiment 1. The results of the calculated R values of experiment 1 and 2 are listed in Table 5.

The difference R, due to the listener AW, pertaining to the 500 Hz mean frequency in experiment 2, was by 3 to 4 Hz higher than that obtained in experiment 1. Similarly, at the mean frequency of 1000 Hz, R was by 9 to 16 Hz larger than in experiment 2. The MK listeners results, pertaining to the mean frequency of 500 Hz, led to a difference $\Delta R = R(\mathrm{Expmt.}\,2) - R(\mathrm{Expmt.}\,1)$ equal to $0.3 - 1.4$ Hz. For the mean frequency of 1000 Hz in experiment 2, the magnitude of R was larger by 4.6 to 8 Hz, whereas, only at $\Delta L = 4.4$ dB, it was less by 1 Hz. For the average frequency of 500 Hz, the listener

Table 5. The differences between adjusted frequencies (in [Hz]) for the SH and SL complementary pairs for the experiments 1 and 2.

f_{av} [Hz]		500			1000		
ΔL [dB]		±2.4	±3.4	±4.4	±2.4	±3.4	±4.4
DAI [2]	#1	+0.8	+4.8	+3.6	+1.45	−1.8	+3.7
	#2	+0.3	+2.55	+4.65	−2.65	+0.55	+3.7
	#3	+2.85	+3.20	+2.95	+2.45	+8.95	+12.25
Exp. 1	AW	+0.6	+0.0	−0.15	+0.85	+2.5	+3.65
	MK	+0.0	+0.4	−0.55	−0.65	+0.55	+0.95
	PR	−0.9	+0.65	+0.9	+0.35	+0.25	−1.25
Exp. 2	AW	+0.35	+0.85	+1.85	+3.9	+1.1	+0.5
	MK	−1.05	−1.15	−2.05	−2.4	−0.35	+2.05
	PR	−4.45	−3.35	−2.5	−10.45	−9.85	−9.2

PR reported a difference $\Delta R = 2.2 - 4.4$ Hz; only at $\Delta L = 2.4$ dB, this difference was practically 0. Much higher differences: $\Delta R = 24 - 34$ Hz were obtained for this listener at a frequency of 1000 Hz. Relatively small values of the R – difference heard by the listener MK may result from her R(Expmt. 1) – valuations being systematically larger than those of the remaining listeners.

In Fig. 5 arranged together with the data of Fig. 4, Part I, the normalized (divided by beating tones frequency difference) envelope weighted of instantaneous frequency

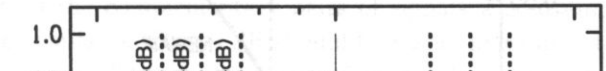

Fig. 5. Illustration of the asymmetry effect for the normalized (by dividing by beating tones frequency difference) *envelope weighted average of instantaneous frequency* (NEWAIF) and the *intensity (squared envelope) weighted average of instantaneous frequency* (NIWAIF). The signal amplitude ratios used in the experiments are denoted by vertical dotted lines. The lack of symmetry is observed for complementary pairs of the beating tones SL and SH.

NEWAIF, and squared envelope (intensity) weighted average of instantaneous frequency NIWAIF are plotted against the amplitude ratio δ; the δ – values used previously in the experiments for SL and SH pairs have been marked (dashed vertical lines).

In view of the asymmetry of the graphs shown in Fig. 5 an asymmetry for the SL/SH pairs may be anticipated. However, this asymmetry has no effect on the value of the R parameter discussed above. But we have to recall that, at $\delta = 1$ ($\Delta L = 0\,\mathrm{dB}$), i.e. for the matched signals changeable in experiment 1 and adjustable in Dai's investigations, there was some frequency shift due to the asymmetry of those graphs. Moreover, the results were referred to the arithmetic average of the two tones of the pitch matching signal. In order to estimate the frequency displacement (shift), at the two equal levels, i.e. $\Delta L = 0\,\mathrm{dB}$, the ratio $(\Delta f/f_L)\cdot 100$ [%] was introduced as an approximate measure of the bandwidth narrowness of the complex signals. For the two mean frequencies of the beats, i.e. 500 and 1000 Hz, this value was 8.33%. In Fig. 6, which is some modification of Fig. 5 (Part I), this value is marked as the vertical dashed line. The crossing points of this line with the EWAIF & IWAIF graphs provide a numerical reading of this asymmetry. The frequency shift values at $\Delta L = 0\,\mathrm{dB}$ and the mean frequency of the beats equal to 500 Hz amount to +1.65 Hz (EWAIF) and +0.38 Hz (IWAIF). At the 1000 Hz mean frequency, they are +3.27 Hz and +0.78 Hz, respectively. The correction of the pitch matching asymmetry may be accomplished by subtracting the above numbers from the listeners matched frequencies for the SL/SH pairs of signals.

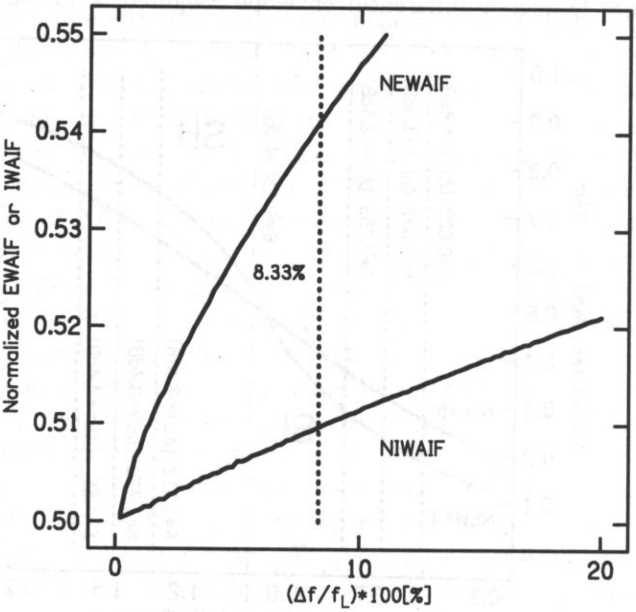

Fig. 6. Normalized (by dividing by beating tones frequency difference) *envelope weighted average of instantaneous frequency* (NEWAIF) and *intensity (squared envelope) weighted average of instantaneous frequency* (NIWAIF) shifts due to the not completely fulfiled narrow-band condition. The lower tone frequency $f_L = 480\,\mathrm{Hz}$, and SPL difference $\Delta L = 0\,\mathrm{dB}$. Δf is the frequency difference between the beating tone components. The value 8.33% (vertical dotted line) corresponds to signal parameters used in the experiments.

Using the arguments from Part I of this study, the asymmetry emerges as a consequence of the influence of the relative changes of the amplitude envelope on the frequency envelope. The greatest effect of these changes occurs at the highest rates of the amplitude envelope, i.e. for $\Delta L = 0\,\mathrm{dB}$, which corresponds to the signals used in the experiment 1 and those of DAI [2]. In experiment 2, the envelope of the changeable signal was the same as that of the fixed one. This means that the resultant frequency envelope depended on the relative changes of the amplitude envelope, but to a lesser degree. Therefore, the asymmetry related corrections, which have to be subtracted from the matched carrier frequencies f_c of the AM signal, will be smaller.

6. Conclusions

The discussed results of the investigations carried out in experiment 1 and experiment 2 together with those of DAI [2] yield evidence of a substantial pitch differentiation by individual listeners. Contrary to the listeners, whose results have been reported by DAI [2], the listeners involved in this study have performed a pitch evaluation following consequently an individually adjusted scheme. Owing to the aforementioned differentiation, a generalization of the conclusions of these investigations is not possible. Most likely, VERSCHUURE *et al.* [19] were right when arguing that a general evaluation of the pitch should be accomplished applying a larger number of listeners.

The asymmetry reported earlier by DAI [2] was not confirmed in experiment 1, where the pitch evaluation of the complementary pairs of signals was attempted. The small differentiation of the results for SL and SH should be rather attributed to the individual preferences of the listeners. The results obtained by the listeners can be corrected as described above, however, such a correction does not prove that the physically occurring asymmetry is the unique cause of the perceived asymmetry of the results.

If, as a valid proposition, same kind of independent "channels" controlling the perception of the amplitude and frequency changes [7, 8] were adopted, then some modification of Eq. (I.14), and consequently of equations (II.3) and (II.5), would perhaps be justified. This modification would consist in diminishing the role of the imaginary part of the complex instantaneous frequency (the phase changes) in shaping the resultant values of the frequency envelope variations. This could be achieved, for example, with a summation of the magnitudes of the real (related to to envelope changes) and imaginary parts of the complex instantaneous frequency $\mathrm{CIF}(t)$.

A set of interesting data was obtained in the experiment 2. Not only an asymmetry of the SL and SH signals was discovered, but quite often, values of R (the difference of the matched frequencies) much larger than those in experiment 1 were obtained. It can be presumed that, to some extent, due to the events of both the fixed and changeable signals, the listeners could concentrate better on their experimental task in experiment 2 than in experiment 1. At the present stage, however, it is difficult to give an explicit rationale for the cause of the pitch differences perceived in these two experiments.

There is no doubt that the idea of the amplitude weighted frequency variations as a means for the evaluation of the sound pitch has been already adequately established.

The problem which remains to be solved is the choice of appropriate weighting functions, for instance those proposed by IWAMIYA [7, 9] and, may be better, correlated with the loudness. Perhaps, the suggestions by ZWICKER [20], that changes of the amplitude are responsible for the perception of the frequency separations, including their dynamical variations, ought to be reconsidered. These amplitude changes are believed [20] to obscure the perception of frequency fluctuations (smoothing effect). The frequency variations to be accounted for in "calculating" the pitch, according to the authors of the present article, should be expressed by equation (I.14) determining the frequency envelope, and modified as explained above.

Acknowledgments

The authors wish to thank Prof. E. OZIMEK for the fruitful discussions and his numerous suggestions during the experiments.

References

[1] J.N. ANANTHARAMAN, A.K. KRISHNAMURTHY, L.L. FETH, *Intensity-weighted average of instantaneous frequency as a model for frequency discrimination*, J. Acoust. Soc. Am., **94**, 2, 723–729 (1993).

[2] H. DAI, *On the pitch of two-tone complexes*, J. Acoust. Soc. Am., **94**, 2, 730–734 (1993).

[3] L.L. FETH, H. O'MALLEY, *Two-tone auditory spectral resolution*, J. Acoust. Soc. Am., **62**, 4, 940–947 (1977).

[4] L.L. FETH, *Frequency discrimination of complex periodic tones*, Perception & Psychophysics, **15**, 2, 375–378 (1974).

[5] L.L. FETH, H. O'MALLEY, J. RAMSEY, *Pitch of unresolved, two-component complex tones*, J. Acoust. Soc. Am., **72**, 5, 1403–1412 (1982).

[6] W.M. HARTMANN, *The effect of amplitude envelope on the pitch of sine wave tones*, J. Acoust. Soc. Am., **63**, 4, 1105–1113 (1978).

[7] S. IWAMIYA, K. FUJIWARA, *Perceived principal pitch of FM-AM tones as a function of the phase difference between frequency modulation and amplitude modulation*, J. Acoust. Soc. Jpn. (E), **6**, 3, 193–202, 1985.

[8] S. IWAMIYA, K. KOSUGI, O. KITAMURA, *Perceived principal pitch of vibrato tones*, J. Acoust. Soc. Jpn. (E). **4**, 2, 73–82, (1983).

[9] S. IWAMIYA, S. NISHIKAWA, O. KITAMURA, *Perceived principal pitch of FM-AM tones when the phase difference between frequency modulation and amplitude modulation is in-phase and anti-phase*, J. Acoust. Soc. Jpn. (E), **5**, 2 (1984).

[10] L.A. JEFFRES, *Beating sinusoid and pitch changes*, J. Acoust. Soc. Am., **43**, 6, 1464 (1968).

[11] R.A. JENKINS, *Perception of pitch, timbre, and loudness*, J. Acoust. Soc. Am., **33**, 11, 1550–1557.

[12] R.J. RITSMA, F.L. ENGEL, *Pitch of frequency modulated signals*, J. Acoust. Soc. Am., **36**, 2, 1637–1644 (1964).

[13] J.G. ROEDER, *The physics and psychophysics of music*, 2nd Ed., Springer, Berlin 1975.

[14] T.D. ROSSING, A.J.M. HOUTSMA, *The effect of amplitude envelope on the pitch of short sine-wave and complex tones*, ICA 12 - Toronto 1986, B2-1.

[15] L. RUTKOWSKI, *Symmetry and asymmetry as a physical and perceptual feature of the complementary pair of beating sinusoids, Part I. Amplitude and frequency envelope relations*, Arch. Acoust., **23**, 1, 51–66 (1998).

[16] J.F. Schouten, R.J. Ritsma, B.L. Cardozo, *Pitch of the residue*, J. Acoust. Soc. Am., **34**, 8, Part 2, 1418–1424 (1962).

[17] S.S. Stevens, *The relation of pitch to intensity*, J. Acoust. Soc. Am., **6**, 3, 150–154 (1935).

[18] N.J. Versfeld, A.J.M. Houtsma, *Discrimination of changes in the spectral shape of two-tone complexes*, J. Acoust. Soc. Am., **98**, 2, 807–816 (1995).

[19] J. Vershuure, A.A. van Meeteren, *The effect of intensity on pitch*, Acustica, **32**, 33–44 (1975).

[20] E. Zwicker, *Direct comparisons between the sensations produced by frequency modulation and amplitude modulation*, J. Acoust. Soc. Am., **34**, 8, Part 2, 1425–1430 (1962).