Acoustic and Biomedical Engineering

Audibility of Tones in Equivalent Rectangular Bandwidth Wide Inharmonic Tone Complexes

P. Kleczkowski* and M. Pluta

AGH – University of Science and Technology, Faculty of Mechanical Engineering and Robotics
Department of Mechanics and Vibroacoustics, al. A. Mickiewicza 30, 30-059 Krakow, Poland

The threshold of discrimination of a pure tone probe in the background of a one ERB wide masker composed of inharmonic partials was investigated. The experiment was participated with a large group of 35 listeners. Very consistent results were observed across two independent variables: the number of partials (from three to five) and their amplitudes being equal or unequal. For all but one of the conditions, the threshold obtained was between –4 and –5 dB. The threshold slightly decreased with increasing number of partials and was slightly lower when their amplitudes were even.

DOI: 10.12693/APhysPolA.123.984

PACS: 44.66.Fe, 44.66.Hg, 44.66.Dc

1. Introduction

Perception of mixed sounds in a realistic environment has not found a proper attention [1]. An example of such an environment is a piece of multi-instrumental music. Simultaneous masking occurs primarily within the critical band (CB). This type of masking is usually referred to as the energetic masking. In a typical context of music, many sound sources are mixed within one CB and the authors are not aware of any studies on the number of instruments which can be simultaneously perceived. In this work, the authors have chosen an inharmonic tone complex as the masker and a pure tone as the signal for the simplest model to begin with. The justification is the following: in a realistic musical material in low or middle frequencies usually each instrument has only one harmonic in a particular CB.

According to [2], wideband harmonic complex tones (HCT) with spacing of components closer than one CB were about 10 dB less efficient in masking than inharmonic tone complexes. The authors argued that this was due to regularity in beatings between partials that caused masking release.

The threshold of detection of a stimulus in some conditions may be different than the threshold of discrimination of its pitch. For a 1000-Hz tone in noise this was investigated by Cardozo [3], who found that for Δf of 16 Hz at the threshold of detectability the pitch was recognized correctly in about 50% of trials and improved to 83% when the level of the tone was increased to 3 dB above the detection threshold. Scheffers [4] investigated low-pass filtered vowel sounds, pulse trains, and pure tones and found that for all these signals changes of 5% or more in fundamental frequency F_0 could be well detected at signal levels close to masked threshold. However, for smaller changes in F_0 like 1%, the F_0 discrimination required signal level of 5–10 dB above masked threshold. This was confirmed in [5] and [6]. It was found in [5] that when the harmonic components are resolvable, their F_0 is discriminated very close to the masked threshold.

Since according to the model proposed above, the inharmonic components within one CB investigated will not be resolvable, some effect of higher F_0 discrimination threshold than the detection threshold may be expected in view of the above works.

The authors have assumed that in the perception of music, F_0 discrimination is more important than the detection threshold. The assumption is based on more prominent and fundamental nature of the melody and harmony element over the timbre element. Detection itself can affect the timbre, but melody and harmony both require a discrimination. Therefore, the authors have chosen the model proposed in the first paragraph above, and decided to measure F_0 discrimination and not the detection threshold. The purpose of this research was to find the threshold of discrimination of the pitch of a pure tone at the background of an inharmonic tone complex, depending on the number of components in that tone, e.g. the density of components.

2. Stimuli

A set of stimuli consisting of pairs of 3-, 4- and 5-partial inharmonic tone complexes was used as maskers. In each case partials fell within an equivalent rectangular bandwidth (ERB). Within a pair, both complexes were identical, but to each of them a pure tone (signal, further referred to as a test tone) was added, also falling within the same bandwidth. The tone had a different frequency in each complex in a pair. If it was higher in the first complex, then it was lower by 50 ct (cents) in the second one, and vice versa (Fig. 1). Frequencies of partials are presented in Table I. The test tones had frequencies of

*corresponding author; e-mail: kleczkow@agh.edu.pl

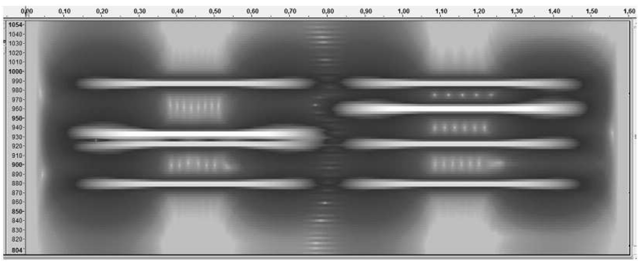

Fig. 1. Spectrogram of a 3-partial sample, with the test tones in lower (left) and higher (right) intervals, generated by Audacity 2.0.1, 16384-sample Hann window.

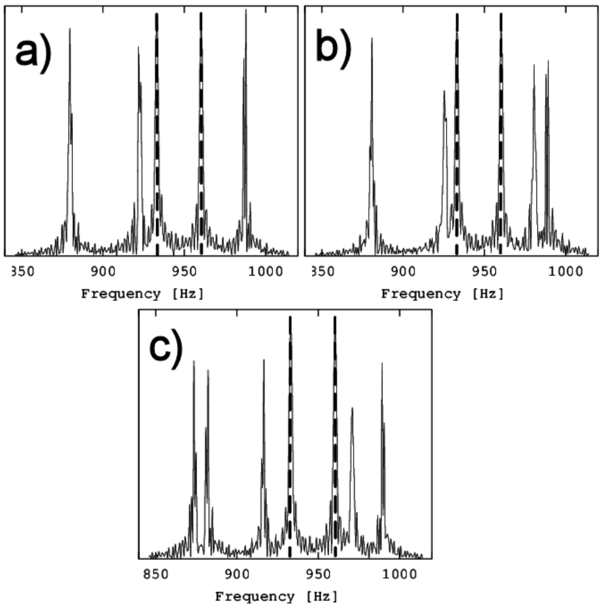

Fig. 2. Amplitude spectra of complexes with the lower and the higher test tones indicated by dashed lines for (a) – 3-tone, (b) – 4-tone, (c) – 5-tone complexes. Spectra generated by GNU Octave 3.6.2, 65536-sample rectangular window.

932.33 Hz and 959.65 Hz. Such values were selected in order to avoid amplitude modulation. There were additional conditions for the partials' frequencies. They had to be harmonics of some (various) musical pitches (in case of 4- and 5-partial complexes, the pitches were from a440-reference system, and in 3-partial complex they were from a443-reference). Finally, there had to be enough space for raising the frequency of the test tone without crossing the frequency of any partial (Fig. 2). In half of the cases partials had equal amplitudes, in the other half they were varied with a 1.5 dB step in random order (Table II). The variation was introduced in order to make the maskers closer to the conditions of a real multi-instrument context. The levels of all complexes were normalized to the same RMS value. The test tone level was measured relative to the RMS of the masker, i.e. the sum of partials, and ranged from –8 do +2 dB with 2 dB increment. All sounds had boxcar envelopes with 20 ms fade in and 20 ms fade out. Duration of the whole sample (two complexes) was 1.6 s, including 0.5 s for each complex, and 0.2 s of silence in between, in the beginning and in the end.

TABLE I
Frequencies of partials [Hz].

partial no.	3-tone	4-tone	5-tone
1	879.02	880.00	873.07
2	921.93	924.99	880.99
3	986.67	979.99	915.69
4		987.77	970.14
5			988.88

TABLE II
Relative amplitudes of partials in uneven amplitudes condition.

partial no.	3-tone	4-tone	5-tone
1	1.0	0.595	1.0
2	0.707	0.707	0.707
3	0.841	0.841	0.841
4		1.0	0.5
5			0.595

A set of 72 different stimuli was created and used in the test, by combining three types of complexes (3-, 4- and 5-partial), two types of spectrum envelope (equal and varied), six levels of the test tone, and two orders of pitches of the test tone (low-high or high-low).

3. Procedure

The test was conducted using a dedicated computer software, running on a personal computer, equipped with an external audio interface M-Audio Fast Track Pro USB, with stimuli reproduced by means of closed headphones Beyerdynamic DT 770 Pro in silenced rooms of a recording studio. The stimuli were presented diotically, as this was found to be more comfortable for listeners. The goal of the software was to present, upon the listener's request, subsequent sound stimuli, according to a prepared playlist and to record the answers. The stimulus, as described above, was a pair of inharmonic tone complexes with an additional test tone changing its frequency up or down from the first to the second complex. The listener had a choice between two possible answers: "the first is higher", and "the second is higher", referring to the test tone. Confirmation of the answer activated the next stimulus. There was no time limit, and no possibility to repeat a sound sample. Every stimulus was presented 11 times during the test, each time evaluated as an individual example by the listener. During the test the feedback was not provided.

The test consisted of a training session, with a limited set of stimuli, and the main session, with all of the stimuli, lasting about one hour. Stimuli were presented in random order, however the order was identical for each listener. The test was attended by 35 listeners, all of them 1st year students of the major in acoustical engineering. Their audiograms were not obtained, but none reported any problems in hearing.

4. Results

The results are presented in Table III in the form of percentage of correct recognitions for the particular stimulus. Averages of equal and varied amplitude versions of the same stimulus are also included, as well as linearly interpolated level for 75% recognition. The same data is presented in Fig. 3 in the form of psychometric functions with a fit of the sigmoidal function,

$$y = A_2 + \frac{(A_1 - A_2)}{1 + \exp\left(\left(x - x_0\right)/dx\right)}, \tag{1}$$

without weighting, using the Levenberg-Marquardt algorithm with tolerance of ≈ 0.0001. Table IV contains selected parameters of the fit: x_0 represents 50% threshold for the curves, and dx relates to the slope (steep curves have lower values).

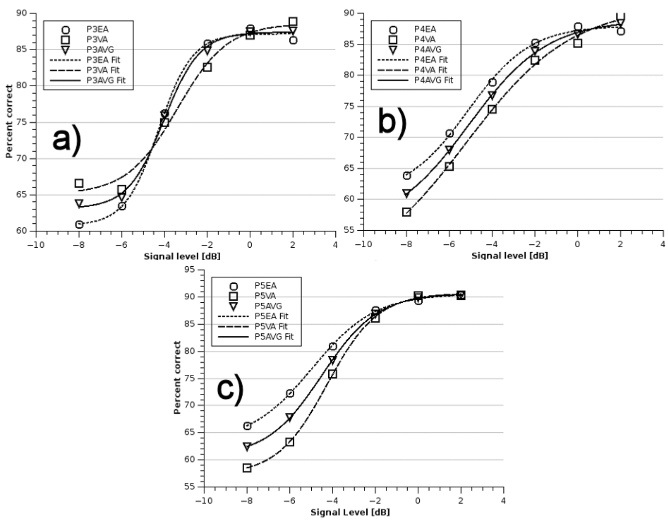

Fig. 3. Psychometric functions (EA – equal amplitudes, VA – various amplitudes, P – partials, AVG – average equal/unequal) for (a) – 3-tone, (b) – 4-tone, (c) – 5-tone complexes.

Statistical analysis was performed to find out whether (i) the number of tones in a complex, and (ii) their amplitudes being equal or unequal affected perception of the probe tone.

In (i) the analysis of variance (ANOVA) for correlated samples was used. The correlation was observed in Table III, where data in all rows followed a similar pattern of increase as the level of the probe tone was increased from -8 dB through +2 dB. The results averaged for pairs of equal/varied amplitudes of tone complexes (respective rows in Table III), were taken to the test and treated as correlated samples. The averages for samples were as follows: 77.35 for 3 tones, 77.42 for four tones, and 79.26 for five tones. ANOVA showed, that all these differences were significant: $\{F(2,10) = 4.82, p = 0.034\}$, where F was obtained from mean sum of squares within group equal to 7.03, and mean sum of squares of error equal to 1.49. This proved that the perception of probe tones increased with each increase of the number of inharmonic tones in the complex, from 3 to 5.

In (ii) the t-test for the difference in means of correlated samples was used. The first sample consisted of all percentages of perception in Table III in complexes with equal amplitudes, ordered accordingly (all three respectful rows in Table III), and the second sample consisted of results for complexes with unequal amplitudes (18 values in each sample). The average difference (2.027) showed that it was easier to recognize the probe tone in complexes with equal amplitudes. The t-test ($d.f. = 17$) proved that this result was significant, at $p < 0.025$.

TABLE III

Test results (EA – equal amplitudes, VA – various amplitudes, P – partials, AVG – average equal/unequal, LI – level interpolated for 75% recognition [dB]).

Tone complex	–8 dB	–6 dB	–4 dB	–2 dB	0 dB	+2 dB	LI
P3EA	60.91	63.51	76.23	85.84	87.92	86.36	–4.19
P3VA	66.62	65.71	74.94	82.60	87.01	88.83	–3.98
P3AVG	63.76	64.61	75.86	84.82	87.47	87.60	–4.15
P4EA	63.90	70.65	78.96	85.32	87.92	87.14	–4.95
P4VA	57.92	65.32	74.55	82.47	85.19	89.61	–3.89
P4AVG	60.91	67.99	76.76	83.90	86.56	88.38	–4.40
P5EA	66.23	72.34	80.91	87.66	89.35	90.39	–5.38
P5VA	58.57	63.25	75.84	86.10	90.26	90.26	–4.13
P5AVG	62.4	67.80	78.34	86.88	89.81	90.33	–4.63

TABLE IV

Selected parameters of psychometric function fit procedure.

Tone complex	x_0	dx
P3EA	-4.28 ± 0.13	0.79 ± 0.14
P3VA	-3.41 ± 0.42	1.18 ± 0.4
P3AVG	-4.03 ± 0.13	0.84 ± 0.15
P4EA	-5.13 ± 0.33	1.46 ± 0.27
P4VA	-5.23 ± 1.03	2.49 ± 0.82
P4AVG	-5.09 ± 0.2	1.84 ± 0.16
P5EA	-4.95 ± 0.15	1.47 ± 0.13
P5VA	-4.25 ± 0.08	1.17 ± 0.08
P5AVG	-4.49 ± 0.04	1.29 ± 0.03

5. Conclusions

The average threshold for the discrimination of pitch of a probe tone with an ERB wide inharmonic tone complex as masker is between –4 and –5 dB. The spread of results between 3, 4 and 5 partials in an inharmonic masker is low. The spread of results between the conditions of equal and unequal amplitudes of partials is also low, indicating good reliability of results.

The discrimination of pitch was investigated rather than the detection of the probe tone, and hence the threshold found can be considered representative for audibility of a musical instrument against the background of a group of other instruments.

A very small but statistically significant effect of the density of partials was observed. The threshold decreased by 0.5 dB with increase of density from three to five partials.

The threshold was lower when the amplitudes of the masker were even.

Acknowledgments

This research was supported by an AGH University grant no. 11.11.130.

References

[1] R. Plomp, *The Intelligent Ear. On the Nature of Sound Perception*, Lawrence Erlbaum Associates, Mahwah (NJ) 2002.

[2] W.C. Treurniet, D.C. Boucher, *J. Acoust. Soc. Am.* **109**, 306 (2001).

[3] B.L. Cardozo, *Acustica* **31**, 330 (1974).

[4] M.T. Scheffers, *J. Acoust. Soc. Am.* **76**, 428 (1984).

[5] H. Gockel, B.C.J. Moore, C.J. Plack, R.P. Carlyon, *J. Acoust. Soc. Am.* **120**, 957 (2006).

[6] P. Carlyon, *J. Acoust. Soc. Am.* **99**, 517 (1999).