ARCHIVES OF ACOUSTICS
16, 3–4 489–500 (1991)

# AUDIBILITY OF CHANGES IN SPECTRAL STRUCTURE OF SAW-TOOTH STIMULI

H. HARAJDA AND W. MIKIEL

Department of Acoustics, Institute of Physics, Pedagogical University
(65-069 Zielona Góra, Plac Słowiański 6)
Institute of Fundamental Technological Research, Polish Academy of Sciences
(00-049 Warszawa, Świętokrzyska 21)

This paper presents the experiments which provided an answer to the question: what structural changes in the saw-tooth stimuli are likely to affect changes in the colour perception? The sound material of the experiments consisted of signals with structural features which can be found in the signals of natural string instruments. The signals were assumed to be characteristic of the following constant quantities:

– harmonic structure composed of successive component tones,

– triple time structure: attack-steady state-decay,

– gradually decreasing tendency in the level of the spectrum envelope. In the experiments the ATARI computer system was utilized. The groups of observers consisted of 5 subjects well-educated in music and acoustics who went through a special pre-test training. The research findings on the dependence of changes in auditory sensation on the changes in the structure of the saw-tooth stimuli being compared indicate that this relationship assumed a simple form when the variables were either the number of component tones or the decreasing tendency in the level of spectrum envelope.

## 1. Introduction

Since Helmholtz [10] advanced the problem of the dependence of an aural sensation of sound on the structure of a given signal, a number of suggestions have been put forward to solve this problem including the concept of a tone colour unit the "chrome" [2]. Nevertheless, the problem remains an open one. As the study of the structure complexity of signals produced on musical instruments becomes more precise, the need to define the most important distinguishing features of the structure and the need to move from descriptions of a qualitative relationship between "changes in structure" and "changes in sensation" to quantitative descriptions stands out more clearly [9, 13, 14, 15].

As our knowledge of dynamic spectral structures and the minute structure of natural instrumental and vocal stimuli is still insufficient, it becomes necessary to orient our experiments to stimuli which are structurally less complicated.

For several years experiments on the audibility of structural changes in signals have been carried out using computer techniques. During the last decade, a series of experiments has been regularly carried out by means of profile analysis [4, 6, 7, 8]. The main emphasis has been placed on short multitone complex stimuli modelled following the behaviour of the background of uniform amplitude and the central component of ca. 1000 Hz, distinctive for its amplitude level. A variety of changes have been introduced into this model: changes in the central tone level and the background level, in the number and density of component tones on the frequency scale [4, 5, 7], changes in the frequency and phase [6]. The dependence of change detection on the number of trials [11] has been investigated. The thresholds of difference detection in the spectrum envelope have been determined and, simultaneously, both the background and the central component have been modified [1]. The sound material of the experiments has comprised signals with flat and rippled spectrum envelopes of the background [12], [8].

The spectrum envelope of the natural sounds of string instruments is characteristic of the general decreasing tendency whereas the authors of the above papers of reference assume that the amplitude level is equal for all components.

## 2. Aim and method

The basic aim of the present experiments is to provide an answer to the question: what structural changes in the saw-tooth stimuli are likely to affect changes in the tone colour sensations?

The sound material of the experiments consisted of signals with the structural features which can be found in the signals of natural string instruments. The signals were assumed to be characteristic of the following constant qualities:

– harmonic structure composed of successive component tones (16 tones in the case of standards);
– a constant spectrum envelope slope;
– signals of 100 ms duration;
– triple time structure; attack-steady state-decay;
– time of increase up the amplitude level of to its steady value in: 50 ms;
– formation time of required frequency; 15 ms;
– signal decay of 50 ms duration.

Altered signals differed from standard ones in one of the following characteristics:

– number of harmonic tones
– value of the spectrum envelope slope.

In the experiments the ATARI computer equipment was used. The stimuli were tones generated by the POKEY computer system and the only controlled parameter was the tone intensity level. To design the test material, a modified set of SOFTSYNTH programes was used The SOFTWARE SYNTHESIZER, Ch. Nieber, Berlin 1985.

The method was as follows. The analogue image of a stimuli was transformed into a graphic one to set a pattern for the headphone membrane control. Fifteen aural detection tests of three types: N–tests, Tn–tests and R–tests were designed for the experiment. Each test comprised from 10 to 30 tasks to solve by a group of listeners (50 per cent of them were tasks of "standard-standard" type and 50 percent – of "standard-alteration" type).

The time paradigm for all the test tasks is shown in Table 1.

The experiment was carried out in an audiometric booth and the stimuli were presented to the listeners over TDH–49 headphones. The intensity level of signals was varied over 65 ± 3 dB.

The group consisted of 5 subjects well-educated in music and acoustics who went through a special pre-test training session. The task assigned to the subjects was to determine whether the stimuli were the same or different in terms of their tone colour qualities. The listeners were given response sheets on which they marked "+" for "I can hear a difference", "–" for "I can't hear a difference" and "x" for "I can't make a clear statement".

### *N–tests*

At this stage of experiments there were three tests with altered stimuli different in the number of tone components. In the tasks of the "standard-standard" type the stimuli comprised 16 successive harmonic tones whereas in the "standard-alteration" tasks the second stimulus contained from 1 to 15 successive tones. The slope of the amplitude the envelope was – 3 dB. The fundamental frequencies were 196 Hz, 392 Hz and 659.3 Hz. Each set comprised 30 tasks to perform.

### *Tn–tests*

Two sets of six tests were designed for the use in this experiment. The test stimuli consisted of 16 component tones and the fundamental frequencies were 196 Hz, 392 Hz and 659.3 Hz. In the first set, standard stimuli were characterized by the slope of the spectrum envelope of – 3 dB/oct (W.I). In altered stimuli the slopes were of –0.5, –1.0, –1.5, –2.0, –2.5 and – 3.5 dB/oct. Each test contained 14 tasks to perform (7 of the "standard-standard" type and 7 of the "standard-alteration" type).

In the second set the slope of the spectrum envelope for standard stimuli was – 5 dB/oct and in the case of altered stimuli: –1.0, –2.0, –3.0, –4.0 or 6.0 dB/oct

(WII). Each test contained 10 tasks to perform (5 of the "standard-standard" type and 5 of the "standard-alteration" type).

*R–tests*

The test material was composed of stimuli of fundamental frequencies 196 Hz (g), 293.7 Hz ($d^1$), 392 Hz ($g^1$) and 440 ($a^1$).

Seven tests were designed for this experiment, each of them comprising 16 tasks to perform (8 of "standard-standard" type and 8 of "standard-alteration" type). In the altered stimuli, amplitudes of the maxima located on one of the harmonics (from one of order 3 to 12) were varied, assuming thus different values. The maxima were also set at other places. Changes of the level of maximum placed on tones 3 and ranged from −0.83 to −2.41 dB, on tones from 5 to 8 ranged from −0.36 to −5.11 dB, on tone 9 from −1.85 to −5.11 dB, on tone 10 from −2.41 to 5.11 dB, on tones 11 and 12 from −3.01 to −5.11 dB (Fig. 4).

Independently of the baseline experiment, the subjects performed 58 individual tasks of the "standard-alteration" type. In altered stimuli, the following variations of the maximum level were introduced:

tone 3 – −0.36 dB and −3.01 dB
tone 4 – ranging from 0.36 to 1.32 dB, and −3.65 dB
tone 9 – ranging from 0.36 to −2.41 dB
tones from 12 to 16 – ranging from −0.36 to −5.11 dB.

This procedure broadened the limits within which the differences between the maximum level and the component tone level in the standard signal were contained from 1.49 dB to 11.64 dB.

In such a distribution of the maximum level variations differences between the maximum level and the amplitude level of the preceding tone were contained within the limits of 0.01 to 2.64 dB when the maximum was located on tone 3, and ranged from 6.31 to 11.14 dB when the maximum was located on tone 16.

To sum up, the differences between the maximum level and the amplitude level of a given component tone in the standard signal ranged from 2.09 dB to 11.64 dB.

## Results

*N–tests*

Totally, 2700 raw data were obtained from the subjects in this type of tests (2.312 correct and 372 incorrect detections, one percent of incorrect detections were false alarms. There were also 112 responses of ambiguity ("I can't make a clear statement").

**Table** 1. N-tests results

| Number of eliminated harmonic tones $\Delta$ on | Number of responses N | | | Percentage of responses: [%] | | |
|---|---|---|---|---|---|---|
| | + | — | X | + | — | X |
| 15 | 88 | 2 | 0 | 98 | 2 | 0 |
| 14 | 90 | 0 | 0 | 100 | 0 | 0 |
| 13 | 90 | 0 | 0 | 100 | 0 | 0 |
| 12 | 89 | 1 | 0 | 99 | 1 | 0 |
| 11 | 88 | 2 | 0 | 98 | 2 | 0 |
| 10 | 85 | 5 | 0 | 95 | 5 | 0 |
| 9 | 74 | 16 | 0 | 82 | 12 | 0 |
| 8 | 74 | 16 | 0 | 82 | 12 | 0 |
| 7 | 69 | 21 | 0 | 77 | 23 | 0 |
| 6 | 62 | 28 | 0 | 69 | 31 | 0 |
| 5 | 66 | 24 | 0 | 74 | 26 | 0 |
| 4 | 47 | 52 | 0 | 43 | 48 | 0 |
| 3 | 34 | 54 | 2 | 38 | 60 | 2 |
| 2 | 17 | 70 | 3 | 19 | 78 | 3 |
| 1 | 7 | 79 | 4 | 8 | 83 | 0 |
| 0 | 1330 | 17 | 3 | 99 | 1 | 0 |
| Σ | 2312 | 376 | 12 | 86 | 13 | 1 |

Table 1 shows the analysis of the data obtained as dependent on the number of higher successive component tones which have been eliminated, i.e., as dependent on the difference in the number of component tones between standard and altered signals being compared.

Indications of "I can't hear a difference" occurred in cases of slight structural changes (from 1 to 3 component tones).

Table 2 shows the distribution of correct (O.K.) responses in percentage as dependent on the stimulus pitch. Three segments can be seen:

– when the differences are of 10 to 15 tones, a very satisfactory level of agreement (over 95 percent) has been achieved; these changes are clearly audible;

– when the difference is in 5 to 9 component tones, the number of correct directions is from 70 to 95 percentage; changes within this limit are also audible;

– when 1 to 4 tones have been eliminated, the number of correct responses is considerably lower; these changes are very difficult to detect or almost impossible to discriminate.

The results for the highest pitches of the investigated stimuli are definitely lower as compared with the other signals to detect.

**Table 2.** R-tests results

| Harmonic $H$ | Sound | Intensity level of maximum [dB] | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | −5.11 | | | −4.35 | | | −3.65 | | | −3.01 | | | −2.41 | | | −1.85 | | | −1.32 | | | −0.83 | | | −0.36 | | |
| | | Number of responses | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | + | − | × | + | − | × | + | − | × | + | − | × | + | − | × | + | − | × | + | − | × | + | − | × | + | − | × |
| 5 | g | 8 | 20 | 2 | 0 | 30 | 0 | 0 | 30 | 0 | 13 | 3 | 14 | 10 | 8 | 12 | 20 | 1 | 9 | 26 | 1 | 3 | 30 | 0 | 0 | 30 | 0 | 0 |
| | $d^1$ | 0 | 30 | 0 | 30 | 0 | 0 | 27 | 3 | 0 | 18 | 1 | 11 | 26 | 0 | 4 | 22 | 1 | 7 | 29 | 0 | 1 | 20 | 1 | 9 | 27 | 0 | 3 |
| | $g^1$ | 1 | 28 | 1 | 16 | 14 | 0 | 15 | 10 | 5 | 20 | 4 | 6 | 22 | 2 | 6 | 28 | 0 | 2 | 28 | 0 | 2 | 27 | 0 | 3 | 25 | 0 | 5 |
| | $a^1$ | 16 | 14 | 0 | 28 | 2 | 0 | 28 | 0 | 2 | 27 | 0 | 3 | 28 | 0 | 2 | 25 | 0 | 1 | 30 | 0 | 0 | 30 | 0 | 0 | 30 | 0 | 0 |
| 6 | g | 0 | 30 | 0 | 0 | 30 | 0 | 12 | 16 | 2 | 6 | 18 | 6 | 10 | 8 | 12 | 18 | 0 | 12 | 24 | 1 | 5 | 28 | 0 | 2 | 30 | 0 | 0 |
| | $d^1$ | 6 | 14 | 10 | 12 | 10 | 8 | 25 | 4 | 1 | 30 | 0 | 0 | 30 | 0 | 0 | 26 | 1 | 3 | 30 | 0 | 0 | 30 | 0 | 0 | 30 | 0 | 0 |
| | $g^1$ | 10 | 10 | 10 | 20 | 6 | 4 | 21 | 7 | 2 | 10 | 12 | 8 | 20 | 2 | 8 | 28 | 0 | 2 | 30 | 0 | 0 | 27 | 0 | 3 | 30 | 0 | 0 |
| | $a^1$ | 20 | 10 | 0 | 25 | 5 | 0 | 30 | 0 | 0 | 30 | 0 | 0 | 30 | 0 | 0 | 28 | 0 | 2 | 30 | 0 | 0 | 28 | 0 | 2 | 30 | 0 | 0 |
| 7 | g | 7 | 18 | 5 | 11 | 14 | 5 | 21 | 7 | 2 | 20 | 7 | 3 | 19 | 6 | 5 | 22 | 2 | 6 | 22 | 2 | 6 | 24 | 0 | 6 | 30 | 0 | 0 |
| | $d^1$ | 7 | 20 | 3 | 28 | 0 | 2 | 26 | 0 | 4 | 20 | 3 | 7 | 18 | 5 | 7 | 22 | 3 | 5 | 22 | 1 | 7 | 23 | 0 | 7 | 30 | 0 | 0 |
| | $g^1$ | 20 | 5 | 5 | 14 | 8 | 8 | 12 | 8 | 10 | 22 | 3 | 5 | 28 | 0 | 2 | 23 | 2 | 5 | 22 | 1 | 7 | 25 | 0 | 5 | 30 | 0 | 0 |
| | $a^1$ | 22 | 5 | 3 | 24 | 4 | 2 | 29 | 0 | 1 | 26 | 0 | 4 | 29 | 0 | 1 | 25 | 2 | 3 | 26 | 0 | 4 | 27 | 0 | 3 | 30 | 0 | 0 |

## *Tn–tests*

In these tests, 1.620 raw data were obtained (1.134 correct and 291 incorrect detections). In the number of incorrect responses there were 16 false alarms. In 15 cases the subjects reported that they could not make a clear statement.

The distribution of correct (O.K.) responses as regards the stimulus pitch is shown in Fig. 3.

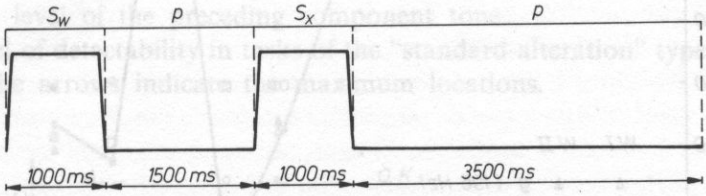


FIG. 1. Time paradigm of test task. $S_w$ — standard signal, $S_x$ — changed signal, $p$ — interval between signals, $P$ — interval for response

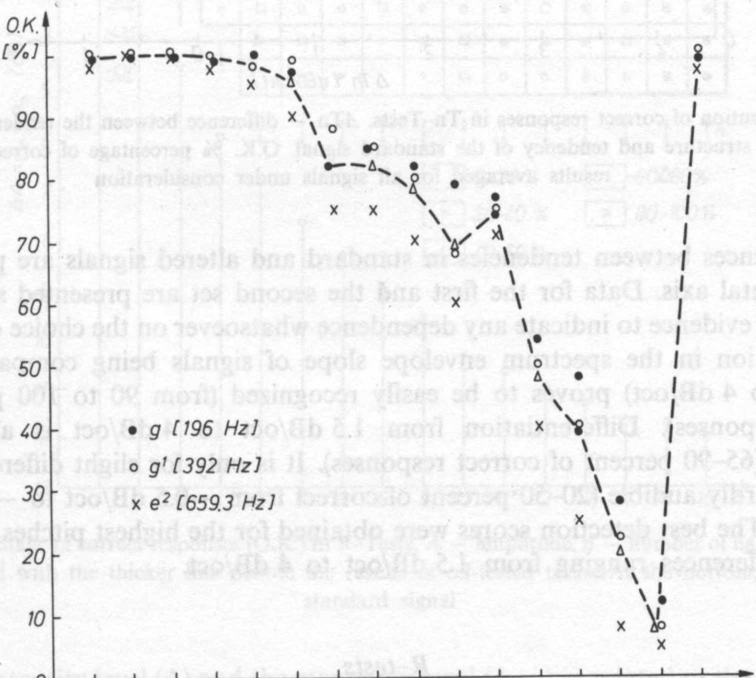


FIG. 2. Distribution of correct responses in N–Tests. $\Delta n$ — loss in the number of harmonic tones. O.K. [%] — percentage of correct responses, $\Delta$ — results averaged for all signals under consideration. Measuring point calculated out on 90 responses

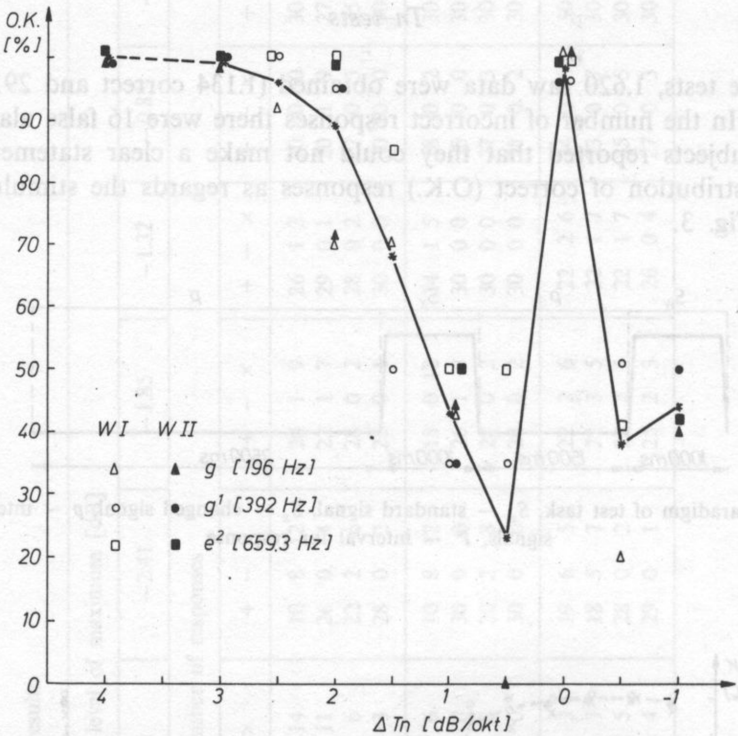


FIG. 3. Distribution of correct responses in Tn–Tests. $\Delta$Tn — difference between the tendency of signal with changed structure and tendency of the standard signal. O.K. % percentage of correct responses. $\triangle$ — results averaged for all signals under consideration

Differences between tendencies in standard and altered signals are plotted on the horizontal axis. Data for the first and the second set are presented separately. There is no evidence to indicate any dependence whatsoever on the choice of pattern. Differentiation in the spectrum envelope slope of signals being compared (from 3 dB/oct to 4 dB/oct) proves to be easily recognized (from 90 to 100 percent of correct responses). Differentiation from 1.5 dB/oct to 4 dB/oct is also easily detectable (65–90 percent of correct responses). It is only for slight differences that they are hardly audible (20–50 percent of correct from −0.5 dB/oct to −1 dB/oct) responses. The best detection scores were obtained for the highest pitches of signals and for differences ranging from 1.5 dB/oct to 4 dB/oct.

## *R–tests*

A total of 9.720 responses was obtained from the subjects for this type of tests and 1.044 responses from solving 58 individual test tasks, 73 percent of correct answers were obtained for both the test and individual tasks. Eighteen percent of them were incorrect, including 2.2 percent of false alarms. In 11 cases, the listeners reported that they were incapable of making a clear statement.

We can further consider the obtained data on the following grounds:

– we can assume that the intensity level of the fundamental tone is a constant level of reference;

– we can analyse differences in the maximum level with respect to a given level of the component tone in the standard stimulus;

– we can try to find out to what extent the level of a given maximum exceeds the intensity level of the preceding component tone.

The level of detectability in tasks of the "standard-alteration" type is presented in Fig. 4. The arrows indicate the maximum locations.

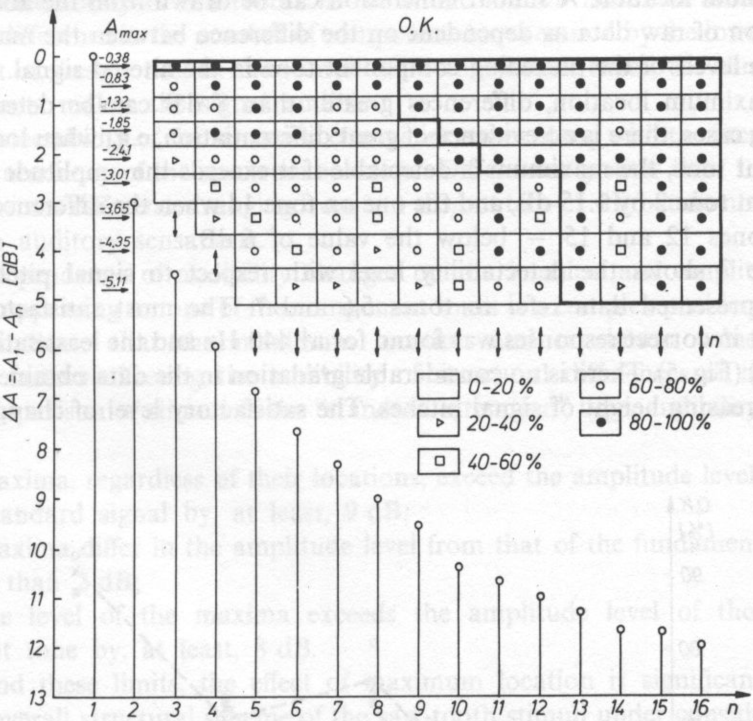


FIG. 4. Percentage of correct responses (O.K.) in R–Tests. *A* – amplitude, *n* – number of harmonic tones. Fields circled with the thicker line denote the results of off-tested tasks. At the bottom, spectrum of standard signal

The intensity level (A) and the maximum level ($A_{max}$) as related to the level of the fundamental tone are plotted on the vertical axis. The distribution of correct (O.K.) responses within the whole range of the alteration field points to a dependence of auditory sensation detectability on both the range of alteration and the maximum location. Overall structural changes comprising a maximum higher than the fundamental tone by − 2 dB regardless of its location, are easily detected (60 to 100 percent of correct responses).

From the analysis of data distribution as dependent on the difference between the maximum and amplitude levels of a given component tone in the standard stimulus it follows that the differences from 9 dB are satisfactorily detected regardless of the maximum location within the spectrum. In the case of smaller differences there is no evidence of a regular relationship between the maximum level and the percentage of detectable sound differences in signals being compared. For instance, it can be easily calculated from the data in Fig. 4 that the maximum located on tone 7 is audible only if its level exceeds the amplitude level of a given component tone in the standard by 3.15 dB, the maximum on tone 8 − by 5.84 dB, and the maximum on tone 9 − by 4.65 dB. And this point to the dependence of auditory sensation on the maximum location. A similar conclusion can be drawn from the analysis of the distribution of raw data as dependent on the difference between the maximum and amplitude levels of the preceding component tone in the altered signal. Irrespective of the maximum location, differences greater than 8 dB can be detected. In the remaining cases, there is no evidence of great differentiation, e.g., when located on the component tone, the maximum is detectable if it exceeds the amplitude level of the component tone 2 by 1.15 dB, and the one on tone 14 when the difference is 8.59 dB, but on tones 12 and 15 − below the value of 6 dB.

Table 2 shows the detectability level with respect to signal pitches.

The presented data refer to tones 5,6 and 7. The most satisfactory level of agreement in correct responses was found for $a^1$ 440 Hz and the least satisfactory for g (196 Hz) (Fig. 5). There is no considerable gradation in the data obtained as related to the increasing height of signal pitches. The satisfactory level of change detection

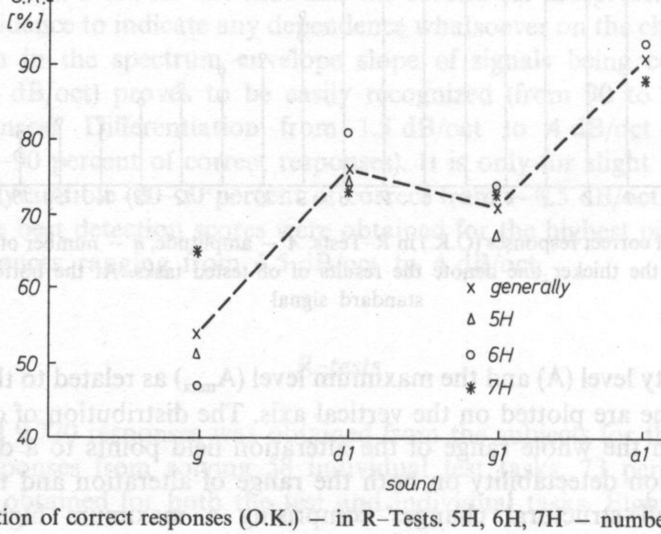


FIG. 5. Distribution of correct responses (O.K.) − in R–Tests. 5H, 6H, 7H − number of harmonic tones

for a$^1$ and, simultaneously, the low level of detectability of the same changes for g$^1$ is certainly due to the fact that a$^1$, a tuning fork sound, is particularly familiar to all those who are professionally concerned with music.

## Conclusions

The research findings on the dependence of changes in auditory sensations on the changes in the structure of the saw-tooth stimuli being compared indicate that this relationship assumes a simple form when the variables are either the number of component tones or the decreasing tendency in the level of the spectrum envelope.

The detectability of differences in the sound stimuli increases along with the increasing difference in the number of component tones caused by the lowering of the upper limit of these tones. In the stimuli of 16 component tones, the elimination of 5 components makes the signal audible.

The detectability of the sensation change increases along with the increase in the decreasing tendency of the spectrum envelope. Differences greater than 1 dB/oct are audible whereas differences ranging from +1 dB/oct to − 1 dB/oct do not affect any changes in auditory sensations.

The dependence of audibility changes on the location and amplitude of the maximum appearing on one of the component tones is a complex problem. The data obtained indicate that the maximum location and maximum value are very important factors affecting the audibility of structural differences of this type.

In comparison of stimuli of the "standard-alteration" type, audibility is secured when:

– maxima, regardless of their locations, exceed the amplitude level of a given tone in standard signal by, at least, 9 dB;

– maxima differ in the amplitude level from that of the fundamental tone by not more than 2 dB;

– the level of the maxima exceeds the amplitude level of the preceding component tone by, at least, 8 dB.

Beyond these limits, the effect of maximum location is significant.

The overall structural scheme of the saw-tooth stimuli under consideration can be treated, according to the model known as "profile analysis" [4, 8], as a set of elements forming the background with a signal (i.e., a change in amplitude forming a maximum) set against that background. The experiment data indicate then that the ratio of signal amplitude and the background is not the only factor which (when the background is not altered) results in aural sensation occurrence. The low audibility level of signals located on overtones 4 and 8 (octave positions) suggests the hypothesis that one of the decisive factors in this respect is likely to be the interval function of the component tone forming a signal. This problem is, however, beyond the purpose and scope of this paper. The factors which affect the ability of change discrimination when the amplitude location is changed in the signals with a decrease in the spectrum envelope have not yet been fully explored and therefore further experiments are called for to clarify the problem.

### References


[1] L. R. Benstein, D. M. Green, *Detection of simple and complex changes of spectral shape*, J. Acoust. Soc. Am, **82** (5), 1587–1592, (1987).

[2] F. Dacos, *Étude fondamentale du violon*, Proc. X Jub. Seminarium Otw. z Akustyki, Poznań, 65–75 (1967).

[3] DE RE ATARI, Software Development Support Group, COMPTON, USA, 1981.

[4] D. M. Green, G. Kidd, Jr, *Further Studies of auditory profile analysis*, J. Acoustic. Soc. Am. 73 (4) 1260–1265, (183).

[5] D. M. Green, G. Kidd, Jr., M. C. Picardi, *Successive versus simultaneous comparison in auditory intensity discrimination*, J. Acoust. Soc. Am. 73(2), 639–643, (1983).

[6] D. M. Green, Ch. R. Mason, *Auditory profile analysis: Frequency, phase, and Weber's Law*, J. Acoust. Soc. Am. 77 (3), 1155–1161, (1985).

[7] D. M. Green, Ch. R. Mason, G. Kidd, Jr., *Profile analysis: Critical bands and duration*, J. Acoust. Soc. Am. 75 (4), 1183–1167, (1984).

[8] D. M. Green, Z. A. Onsan, T. G. Forrest, *Frequency effects in profile analysis and detecting complex spectral changes*, J. Acoust. Soc. Am. 81 (3), 692–699, (1987).

[9] H. Harajda, *Projet de classification du timbre du violon*, Bulletin de la Soc. des Amis des Sciences de Poznań, 35–41, (1967).

[10] H. Helmholtz, *Sensations of Tone*, A. J. Ellis, London (1985).

[11] G. Kidd, Jr., Ch. R. Mason, *Evidence for sensory-trace comparison in spectral shape discrimination*, J. Acoust. Soc. Am. 84 (1), 144–149, (1988).

[12] G. Kidd, Jr., Ch. R. Mason, D. M. Green, *Auditory profile analysis of irregular sound spectra*, J. Acoust. Soc. Am. 79 (4), 1045–1053, (1986).

[13] Koeng-Mo Sung, *String instrument tone synthesis using nonlinear distortion technique*, International Symposium on Musical Acoustics, Mittenwald, Abstracts, 135 (1989).

[14] M. Kwiek, *Resonant factors in sound colour* (in Polish), Polski Rocznik Muzykologiczny, 2, 1–8 (1963).

[15] M. A. Müller, M. Schleske, *Is Scientific Research useful for Violin Makers*, International Symposium on Music Acoustic, Mittenwald, Astracts, 151 1989.


This work was supported by the Central Research Project CPBP 0203/7.10 – 1989

*Received November 13, 1989*