INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2024, VOL. 70, NO. 3, PP. 587-595
Manuscript received June 17, 2024; revised July, 2024. doi: 10.24425/ijet.2024.149583

Profile Analysis: A framework for the study of auditory sound spectrum analysis

Jan Żera

Abstract—**Profile Analysis (PA) is a research approach developed by David M. Green and his associates, aimed at measuring listeners' ability to discern changes in the spectral envelope shape of complex tones. PA introduces an innovative method by separating across-frequency spectral envelope shape comparisons within a complex tone and across-time level comparisons within specific frequency channels. This paper revisits the outcomes of several PA studies and examines the correlations between PA and the results obtained from experiments involving harmonic signals and listening tests designed to enhance timbre evaluation skills among sound engineering students at the Chopin University of Music in Warsaw.**

Keywords—**sound spectrum; auditory profile analysis; formants; detection and discrimination**

I. INTRODUCTION

THE purpose of this paper is to recall a research line known as *Profile Analysis* (PA) which was introduced and extensively explored by professor David M. Green and his co-workers. These studies involved highly formalized and methodologically precise experimental work on the human ability to detect spectral changes in multitone complexes. The primary focus of this research was to investigate how the hearing system analyzes the spectrum of sound; beyond mere level comparisons between simple stimuli. Groundwork for this idea and research began in the early nineteen-eighties and continued with significant research activity for the next 15-20 years, resulting in a comprehensive collection of data gathered under various stimulus conditions.

The major conditions encompassed several aspects, including the basic detection of a change in the amplitude of a single component of an equal-amplitude multitone complex [1]-[8], the effects of additional pedestal [2], rippled spectrum [2],[5]-[7],[9], step and tilted spectrum [2],[7],[10], the uncertainty of spectral components in frequency and amplitude [11]-[15], harmonically spaced components in frequency [16], dichotic listening conditions [2],[6], the influence of background noise [17], and listeners' experience in PA task [4],[13],[15].

The research analyzed a number of signal parameters occurring in the PA, such as frequencies and number of components [1],[3]-[5],[7],[8],[10], components' phases [4], pitch [18], range of level variation [19], signal duration [20]. Among the major issues addressed were the significance of a simultaneous (among spectral components) vs. successive (single frequency but between stimuli) comparisons [1],[21], the construction of specific models for the PA [3],[7],[15],[18],[22], the applicability of Weber's law [4], and determination of the psychometric function governing the PA task [5],[14].

In more recent times, based on the author's knowledge, research on the PA has not been widely pursued or extensively applied in connection with other applications. Nevertheless, several interesting approaches can be found in the literature concerning speech [23], harmonic vs inharmonic signals [24], auditory grouping [25], and certain properties related to the PA [26]-[28].

The PA, nevertheless, exhibits certain relations with research that possess a more applied focus. For instance, notable work was conducted, at the Sound Engineering Department of the Chopin University of Music (CUM) in Warsaw where specific experiments on the PA were conducted in relation to harmonic complexes [29] and formants imposed on noise and natural music signals [30]-[34].

The primary objective of this article is threefold: firstly, to revisit the fundamental assumptions that form the basis of PA; secondly, to review and highlight some of the most significant findings by Green and his associates; and lastly, to present the research conducted at CUM that directly aligns with the PA framework.

II. FUNDAMENTALS OF AUDITORY PROFILE ANALYSIS

Auditory profile analysis in its form led by David M. Green originated from his earlier involvement with C. Watson and his colleagues' research on the perception of spectrally complex multitones (for more details refer to [35]). It is reasonable to assume that when the hearing system makes comparisons between spectrally complex sounds it likely performs both simultaneous cross-spectrum analysis and between-stimuli level comparisons. In successive comparisons, a change of energy within a specific passband, where the difference between stimuli occurs, is compared. This process commonly occurs in traditional masking experiments conducted with the use of narrowband stimuli. Such successive narrowband energy comparisons involve short-term memory usage.

In simultaneous comparisons, there exists a form of cross-spectrum analysis within each stimulus, which may play a role in timbre recognition. This information is stored in long-term memory before the decision is made about the difference

---

Work supported by the grant 504/04064/1034/40.00 from the Warsaw University of Technology.

Author is with Warsaw Institute of Technology, Faculty of Electronics and Information Technology, Institute of Radioelectronics and Multimedia Technology, Poland (jan.zera@pw.edu.pl).

between the two stimuli. Green et al. [1] demonstrated the significance of simultaneous cross-spectrum comparisons in an experiment where the delay between stimuli was extended from 250 ms to 8 s. They observed that when cross spectrum simultaneous comparisons were not available, the thresholds increased by 12 dB (ΔA/A). In contrast, when these comparisons were available, the threshold only increased by 3 dB after the time interval between stimuli was increased to 8 s. This observation, confirmed later for both simultaneous and successive broadband changes in signal spectrum by Dai and Green [21] highlights the importance of conducting comprehensive studies on PA.

The key methodological challenge in PA is to enable simultaneous comparisons across spectrum while eliminating potential successive comparisons. This is done by introducing overall level variation between consecutive signal presentations (within-trial roving of level), as shown in Fig. 1 taken from [1].

During the PA experiments, the overall level is varied and randomly chosen for each signal presentation irrespective of the increase in the level of the tested component (the signal to be detected). A typical roving range used in the PA was from 20 to 40 dB with median level of 45 to 55 dB SPL per component. When the signal to be detected is an increase in level ΔA added to the standard component level A, a certain minimum roving range is necessary to ensure that ΔA remains undetectable through successive between-trial comparisons of level A vs. level A+ΔA. Consequently, the variation in overall level must be sufficiently large to avoid any between-trial successive comparisons. Green's statistical analysis in Appendix A of [35] provides the statistical background for determining the minimum necessary level roving range required for a given maximum ΔA added as a signal.

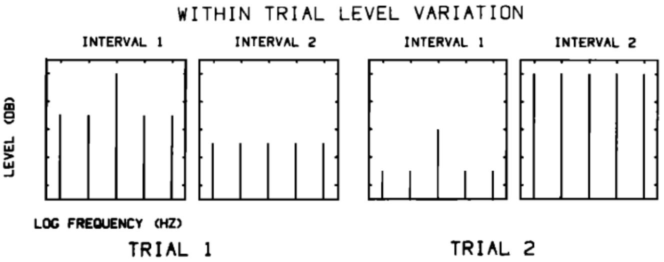

Fig. 1.Within trial roving of level. At each stimulus presentation overall level is changed randomly within assumed roving range. Reproduced from [1], with the permission of the Acoustical Society of America

For the sake convenience, most experiments on PA use the ratio of ΔA/A to describe signal levels and express thresholds values. Thus, a signal level of 0 dB (ΔA=A) corresponds to a doubling (+6 dB) of the component amplitude. Signal levels with ΔA/A of -5 dB, -10, -15 or -20 dB correspond to respective increases in component levels ((A+ΔA)/A) of merely 3.9, 2.4, 1.4 and 0.8 dB. This specific decibel scale of ΔA/A was consequently employed in all the PA research. It allowed for the use of larger differences in numbers to express thresholds. Moreover, it was chosen due to specific experimental setup at that time, where the signal (an increase in component(s) level) was created by summing two separately generated signals in phase: the standard signal of amplitude A and an added signal of amplitude ΔA.

Another crucial aspect in PA experiments is the use of logarithmically spaced components in frequency. While such a spectrum may not replicate the natural distribution of components in real sounds, its use is well-justified as logarithmically spaced components within a complex are spread evenly along the auditory filters on the logarithmic frequency scale. Figure 2 presents a logarithmic complex (depicted as circles) consisted of 21 components extending between 200 and 5000 Hz, with a ratio of 1.1746 between frequencies of consecutive spectral components. It is clear that, on average, there is one component per auditory filter width, expressed as equivalent rectangular bandwidth (ERB), with the actual change ranging from 0.72 to 1.42 components per filter. For a comparison, a similar change for a 25-component harmonic complex of 200-Hz fundamental is shown (represented as squares in Fig. 2). Here the change varies from one component at every 3.54 filters (low number harmonics) to 2.77 components per filter at 5000 Hz. The difference between logarithmic and harmonic stimuli leads to vastly different excitation patterns evoked in the hearing system.

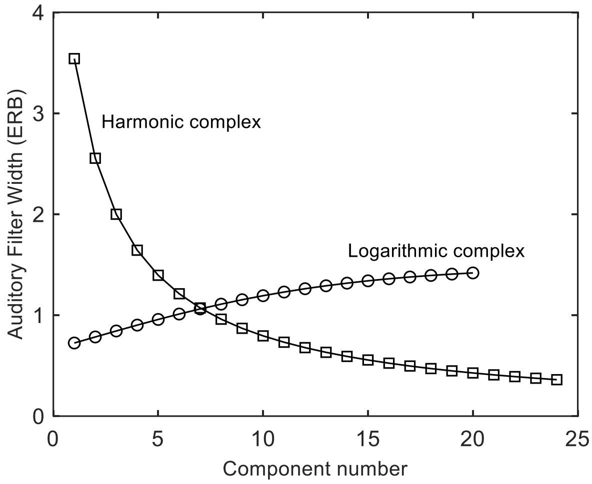

Fig. 2. Spacing between 21 spectral components of a logarithmic complex spread between 200 and 5000 Hz with a ratio of 1.1746 and between 25 components of harmonic 200-Hz complex expressed in the auditory filter widths at corresponding frequencies

III. MAIN BODY OF WORK ON PROFILE ANALYSIS

This section presents a significant portion of the work conducted on PA under the leadership of David. M Green during the lab activities in the 1980s and 1990s.

A. *Essential basic experiment – detecting a change in a single component of an otherwise flat spectrum*

Considerable efforts were directed towards exploring the basic properties of the PA using the simplest stimuli possible. This involved investigating the detection threshold for an increase in the amplitude of a single frequency component in a multi-tone logarithmic complex [1]-[8]. The task was to compare a stimulus with one component having an amplitude of A+ΔA with another stimulus containing all components of equal amplitudes A (see interval 1 and interval 2 in Fig. 1). This task was repeated for several group of subjects at different time with slightly different spectral content of the complex. Fig. 3 from [5] shows a typical result obtained for a 21-component complex with components spanned from 200 to 5000 Hz (frequency ratio of 1.1746 between adjacent components, see also Fig. 2). The overall signal level was randomized following a rectangular distribution within a 20 dB range.

The threshold curve shown in Fig. 3 exhibits what is commonly referred to as a “spectral bowl”. In this pattern, the middle component of the complex stands out as the most detectable, likely due to its optimal cross-channel comparisons with other components within the complex. Irrespective of the bowl shape, the threshold remains relatively flat across different frequencies as level differences between the signal component and standard components (20log10((A+ ΔA)/A)) of about 1.3 dB for the 1000-Hz component is sufficient to detect a change. The difference in levels amounts to approximately 3 dB for the lowest and highest components, which represent the spectral bowl.

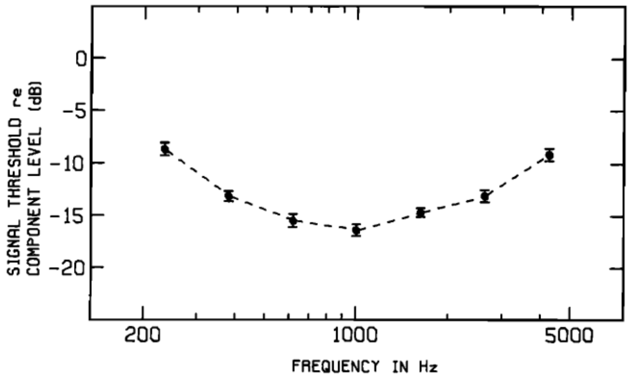

Fig. 3. Threshold for detecting of an increase ΔA in a single component of the 21-component complex with logarithmically spaced components. The ordinate is the relative ΔA/A in decibels. Reproduced from [5], with the permission of the Acoustical Society of America

The phenomenon of thresholds favoring the middle component of a logarithmic complex also applies to complexes with fewer than 21 components, even when positioned in different frequency ranges from 200 Hz to 5000 Hz [5]. This observation confirms that the detection of the central component of a multitone complex occurs under optimal conditions up to a frequency of 5000 Hz. Above 5000 Hz, the thresholds for detecting the signal component in the complex increased to approximately 6 dB (ΔA/A) [5].

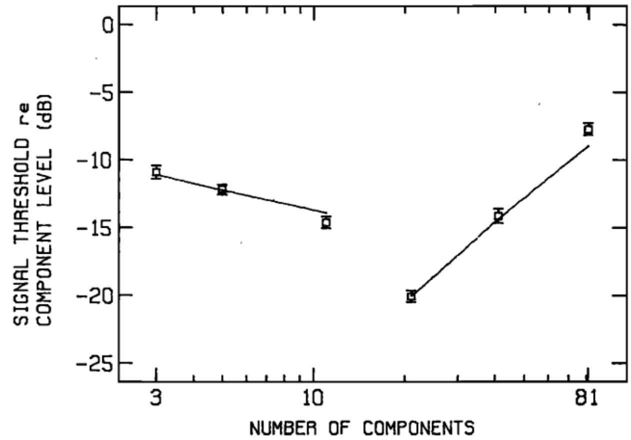

Fig. 4. Thresholds (ΔA/A) for detecting an increase ΔA in a 1000-Hz component of logarithmic complexes with the number of components varied from 3 to 81 in a frequency range of 200-5000Hz. Reproduced from [6], with the permission of the Acoustical Society of America

Several studies [1],[3],[5],[6] have examined how the detection of a single central component in PA depends on the number of logarithmically spaced components within a complex. An example of results taken from study [6] is shown in Fig. 4. As the number of components is increased from 3 to 21 the threshold (ΔA/A) decreases from about -11 to -20 dB (from 2.2 dB down to 0.8 dB as difference in level (A+ΔA)/A). This decrease in threshold is attributed to an expanded number of frequency channels from 3 to 21, as shown in Fig. 1, which facilitates simultaneous cross-spectrum comparisons. When the number of components is further increased to 41 and 81, the threshold increases to about -8 dB (or 2.9 dB). This increase in threshold is an effect of masking. In an 81-component logarithmic complex, on average, there are four components within each auditory filter, leading to this masking phenomenon.

B. *Pedestal experiment – effect of non-uniform spectrum*

In the experiments described so far, all components were of equal amplitude and increase ΔA in amplitude was solely added to one component. However, in the experiments described here, the initial amplitude of the signal component differed from the amplitudes of other components which was referred to as a pedestal. An additional increase ΔA to be detected was then applied to this particular signal component. The condition was used to replicate a simple case of a non-uniform spectrum in a standard, akin to what might be encountered in natural stimuli.

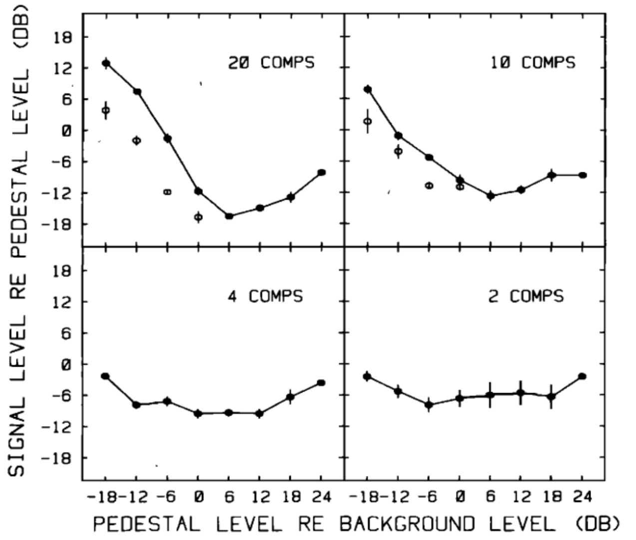

Fig. 5. Pedestal experiment by Green and Kidd [2] for 20, 10, 4 and 2 non signal components in the spectrum. Signal component initial level ranges from -18 dB (negative pedestal) to +18 dB (positive pedestal). Reproduced from [2], with the permission of the Acoustical Society of America

In the pedestal experiment conducted by Green and Kidd [2], the level difference between the pedestal and the other components varied from 6 to 24 dB for positive pedestals and from -6 to -24 dB for negative pedestals. The results for logarithmic complexes with *n* = 3, 5, 11, and 21 components equally spaced in log frequency from 300 to 3000 Hz, are shown in Fig. 5. The pedestal was applied to the middle signal component at 948.6 Hz and the additional change in level was to be detected, while the remaining *n*-1 (used as labels of individual panels in Fig. 5) components were kept at equal amplitudes.

There are the two major findings in this experiment. The positive pedestal that is initially larger amplitude of middle signal component than remaining components had little effect on the threshold, even for a large pedestal of 24 dB. The thresholds are smaller, of about -18 and -12 dB (ΔA/A), for a large number of 10 and 20 non-signal components in the complex, compared to thresholds of about -7 dB for 2 and 4 non-signal components in the complex. This finding confirms, under different experimental conditions, what was observed in Fig. 4, where increasing the number of components to 21 facilitates the detection of the signal component, even for stimuli with a pedestal. For all stimuli, there is some increase in threshold for 18- to 22-dB pedestal.

Comparing the pedestal experiment with the results presented in Fig. 3, for all component amplitudes equal, it is evident that the PA remains effective even if the signal component has an initially elevated amplitude, as it may occur for formants in the spectrum of natural sounds.

In the case of a negative pedestal, the strong increase in threshold for 10 and 20 non-signal components was attributed to the increased masking caused by the surrounding components closely spaced in frequency. To confirm this observation Green and Kidd [2] conducted additional measurements (represented by open circles in the left upper panel in Fig. 5) in which two components nearest to the signal component were removed from the spectrum. This resulted in a decrease in thresholds demonstrating a release from the masking effect.

C. Rippled, step and tilted spectra

Sinusoidal ripple across all amplitudes of components in the complex [2],[6],[7],[9],[17], step spectra [2],[7], and up- and down-tilted spectra [7],[10] were among the types of signal introduced to test wide-band spectral changes in the PA.

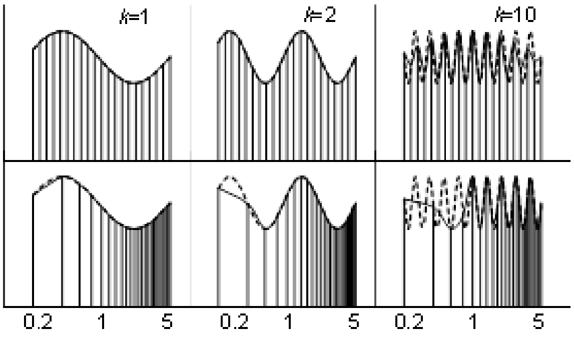

Fig. 6. Rippled imposed on logarithmic (upper panels) and harmonic spectrum (lower panels). Adapted from [29]

A wide-band ripple in the signal was introduced by applying sinusoidal variation in the amplitude of all components, with a parameter *k* defining the number of ripples, typically extending from 1 to 10. The sketch of a rippled spectrum imposed on a 21-component logarithmic complex is shown in Fig. 6, upper panels. This rippled spectrum served as the signal to be distinguished from the flat spectrum in the standard. The lower panels in Fig. 6 show the rippled spectrum imposed on the harmonic complex used in study [29], which will be discussed later in the text.

A rippled spectrum, serving as a wide band spectrum change, was used for 21 components in the complex (e.g. study [5]). However, in Fig. 7, taken from study [6], thresholds are shown for a considerably larger number of 161 components in the logarithmic complex, to test, on extreme, as many as 81 ripples in spectrum. It can be observed that the threshold slightly decreases with an increase in the number of ripples up to 10 ripples, and then it significantly increases. This shows that regardless of whether the number of components is 21 [5] or 81 [6], ten ripples provide the optimum conditions for detection, with the threshold ΔA/A at about -25 dB (corresponding to ripple amplitude by just 0.5 dB).

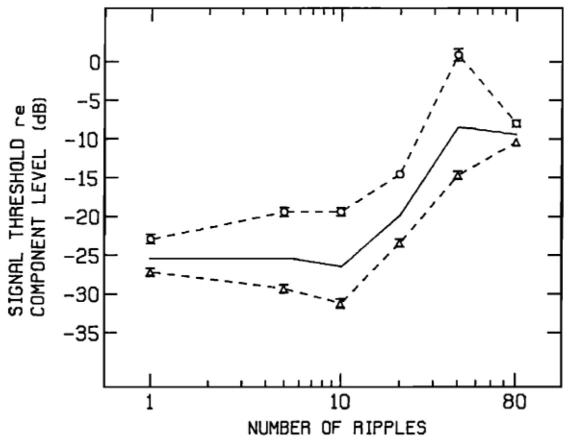

Fig. 7. Detection of rippled spectrum. Average (solid line) of thresholds for two groups of listeners (triangles and open circles). Reproduced from [6], with the permission of the Acoustical Society of America

Step spectra were a type of wide-band spectral change, with larger amplitudes introduced starting from component number n (*step-up* tilt) or with larger amplitudes up to component number n (*step-down* tilt) [2],[7].

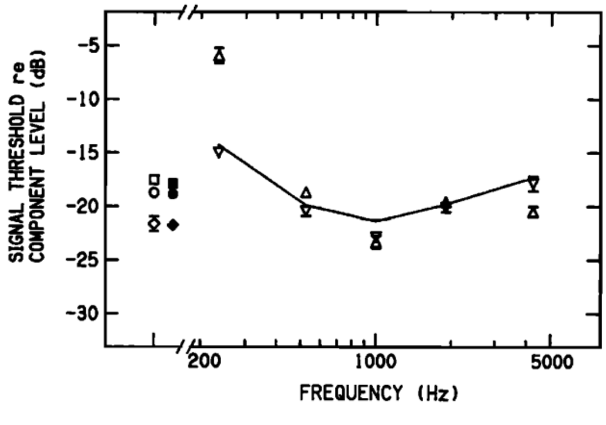

Fig. 8. Thresholds for step-up (triangles) and step-down (inverted triangles) spectra as a function of the frequency of the step. Open squares and circle show thresholds for tilted spectra with flat spectrum in standard. Open diamond represents threshold of alternating decreased and increased amplitudes (not discussed in the text). Solid line and symbols represent model predictions not discussed in the text. Reproduced from [7], with the permission of the Acoustical Society of America

Thresholds for step spectra are shown in Fig. 8. The triangle symbols represent step-up or step-down frequencies, while the solid line represents model predictions, not discussed here. The results demonstrate a similar spectral bowl as seen in Fig. 3, but with thresholds (ΔA/A) being 5-8 dB lower, indicating that wide-band spectral change are more effective in PA than single component changes. Other symbols refer to the detection of tilted spectra, which was conducted in the same study (see Fig. 9).

A wide-band tilt in the spectrum of a logarithmic complex, considered in studies [7] and [10] by Bernstein and Green, was created by a linear increase or decrease of amplitudes with increasing component number. The scheme of the tilted-down logarithmic spectrum is shown in Fig. 9, left panel. The right panel in Fig. 9 shows the same tilt imposed on the harmonic spectrum. The simplest subject task was to detect a positive or negative spectrum tilt in comparison to the flat spectrum. Additionally, four more advanced conditions, in which the standard reference spectrum was initially tilted up or down by ±6 or ±12 dB, were also studied by Bernstein and Green [10]. For the initially tilted spectrum, the added tilt to be detected in the signal caused either a further enlargement of tilt in the standard (both initial tilt and addition either positive or negative) or a decrease of tilt in the standard (addition of a different sign than the original tilt in the standard).

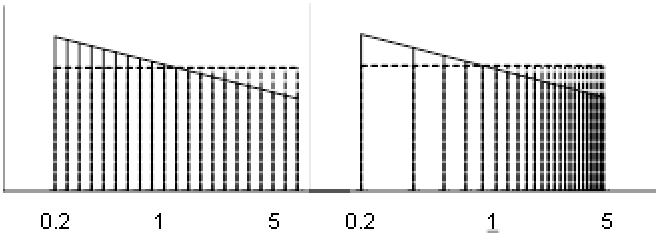

Fig. 9. Tilted spectrum imposed on logarithmic (left panel) and harmonic spectra (right panel). Adapted from [29]

The results of study [7] demonstrated that the initial flat spectrum in the standard required a minimum change in tilt. A downward tilt yielded the smallest threshold of about -18.8 dB (ΔA/A), while an upward tilt a threshold of -17.5 dB. As expected, the existence of an initial tilt of ±6 dB or ±12 dB in the standard studied in [10] increased the thresholds to about -8 dB and 6 dB, respectively. Subjects were consistently slightly more sensitive to opposite direction of the tilt to that set in the standard which is consistent with the general effect of lower thresholds occurring for less steep tilts in the spectrum.

D. *Harmonic complexes*

Logarithmic complexes are very uncommon for natural sound sources. Sound sources typically produce periodic signals, unless it is random noise, whose spectra consist of harmonically related components. Harmonic signals specifically excite the hearing system, producing a clear unambiguous pitch percept related to the fundamental frequency of the complex tone. Such signals are also common in music. Therefore, investigating harmonic stimuli spectra in was considered as an important issue in PA as it was expected that their thresholds might significantly differ from those measured for logarithmic complexes.

A study conducted by Henn and Turner [24] compared PA thresholds for a 1000-Hz middle component of 3, 5, 7, or 9-component logarithmic (275-3625 Hz) and harmonic (200-1800 Hz) complexes. Surprisingly, they found comparable thresholds, approximately -10 dB (ΔA/A), for both harmonic and logarithmic complexes when subjected to a 20-dB overall level roving.

A study of harmonic spectral stimuli in the context of PA was subsequently undertaken by Żera et al. [16] who tested 100- , 200- and 400-Hz harmonic complexes with component frequencies extending up to 6000 Hz (60, 30, or 15 harmonics). In this study, a single harmonic in the spectrum was increased in amplitude and the threshold ΔA was determined while applying a 20-dB roving of the overall level.,

The results of this study are presented in Fig. 10. The thresholds exhibit an increase from approximately -15 dB for harmonic frequencies below 1000 Hz (about a 1.5-dB difference in level between signal and non-signal components), to as much as +15 dB (more than a 16-dB difference in level of signal and non-signal components) at frequencies above 5 kHz.

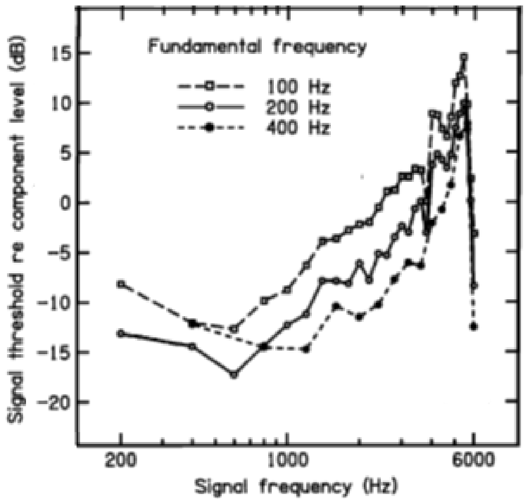

Fig. 10. Threshold for detection of an increment to a single component of harmonic complex as a function of frequency of the component. Averago for three subjects. Reproduced from [16], with the permission of the Acoustical Society of America

Above about 1 kHz, the larger the distances between harmonic (100, 200, or 400 Hz depending on fundamental frequency) the lower the threshold. The substantial drop in threshold values at very high frequency was attributed to the end of the spectrum effect (see [16] for more information).

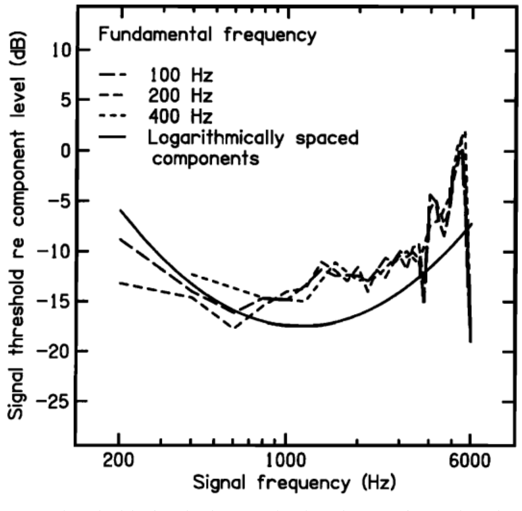

Fig.11. Thresholds for the harmonic signals transformed to the ERB bandwidths compared against data for logarithmic complex. Reproduced from [16], with the permission of the Acoustical Society of America

An important issue was to clarify why the PA thresholds for a single harmonic increase in amplitude differed so significantly from the thresholds observed for logarithmic complex, as shown in Fig. 3. Figure 1 shows that the density of harmonic components per auditory filter differs considerably between harmonic and logarithmic complexes, with a higher number of harmonics present in each filter at high frequencies.In study [16], a simple energy model was used which involved summing up the energy of all harmonics that fell into a single filter. This approach assumes that only energy changes in consecutive auditory filters count in the PA task.

The results of model calculations are shown in Fig. 11. Two effects are observed in these calculations. Firstly, the thresholds for 100-, 200-, and 400-Hz complexes which exhibited significant differences in Fig. 10 coincide after equalization for the width of the auditory filters. This finding shows that considering energy within the filter passband is important and not contradictory to other specific phenomena, such as phase locking in response to harmonics at the output of the particular auditory filter [36]. Model allowed to connect the highly divergent results obtained for harmonic complex with the previous data concerning the logarithmic spacing of components in a complex.

Secondly, overall shape of threshold for harmonic complexes expressed as a change of energy in auditory filter in Fig. 11 is not much different from the thresholds obtained earlier for logarithmic complexes and closely corresponds to the bowl-like shape of the threshold change with frequency. This reinforces the validity of adopting the model of auditory filters for the PA process. The fact that the thresholds measured for 100-, 200-, and 400-Hz fundamentals, align almost perfectly, and exhibit a clear fit to thresholds measured previously for logarithmic complexes further supports this notion.

IV. ATTEMPTS OF WORK TOWARDS REAL WORLD SIGNALS

The major body of experimental work on PA was carried out under the mentorship of Professor David M. Green at the Psychoacoustics Laboratory, University of Florida in Gainesville. Subsequent research conducted by others has been relatively limited. An extensive discussion of various factors affecting the thresholds in PA, was done by Drennan and Watson [37] who discussed earlier studies in reference to the effects of component spacing (harmonic vs non-harmonic), masking, absolute component frequency, possible pitch cues (pitch strength), and temporal factors. These contributing components were discussed with pointing out their relevance to natural signals in contrast to the laboratory created artificial constructs, such as the logarithmic spectrum. In other work [38] Drennan and Watson focused on the effects of training and extensive training on the APA thresholds. The discussion of findings from studies [37] and [38] provided substantial justification for the research undertaken at the Chopin University of Music in Warsaw (CUM), which will be presented in the subsequent three sections.

A. Harmonic complexes and rippled spectra

In an experiment conducted by Ciesielski [29] at CUM in Warsaw, dedicated efforts were made to use using harmonic signals in PA experiments on rippled spectra. The experimental conditions, including the imposition of ripples on the spectrum, signal level (50 dB per component), roving range, and step (20 dB, 1 dB), as well as adaptive psychoacoustic procedure were same as in study conducted by Green et al. [5].

The difference was that a 25-component harmonic complex with a 200-Hz fundamental extending in frequency up to 5000 Hz was used. As a reference condition exact replication of the stimulus used in study [5] was also employed, that is a logarithmic complex with 21 components ranging from 200 to 5000 Hz. The amplitude ripple imposed on the harmonic complex for k = 1, 2, and 10 was used as shown in the lower panel of Fig. 6. The upper panel of Fig. 6 shows the ripple used by Green et al. [5], imposed on logarithmic components and replicated in the study at CUM. Consequently, the stimuli (harmonic and logarithmic) were identical to those for which the frequency differences between consecutive components expressed in the ERB filter widths are shown in Fig. 1.

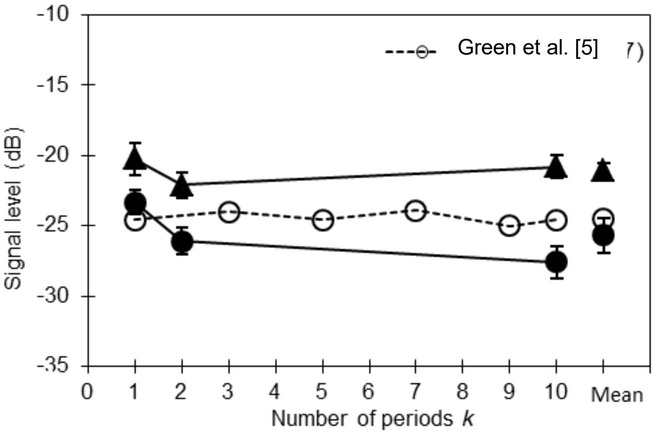

Fig.12. Thresholds for the detection of ripple in the spectrum for k = 1, 2, and 10 imposed on 200-Hz harmonic complex (filled triangles). Mean for three subjects. Thresholds for the replicated experiment of Green et al. [5] for logarithmic complex (filled circles) and original data of Green et al. [5] (open circles) are shown. The error bars represent the standard error of 12 threshold estimations per data point. Data from [29]

The results of the study [29] are shown in Fig. 12. Consistently across various k values, thresholds for harmonic complexes are higher by approximately 3 to 7 dB (ΔA/A) than for logarithmic complexes. However, the (ΔA/A) levels of -30, and -20 dB, represented in Fig. 12 correspond to an increase in level by merely 0.3, and 0.8 dB, respectively. These values are the level changes necessary to detect amplitude increment in a complex, regardless of whether it is logarithmic or harmonic. Thresholds for logarithmic complexes measured in [29] and in the earlier study of Green et al. [5] are quite similar. This finding is noteworthy since both experiments were conducted at vastly different times and in entirely different laboratory environments. The thresholds measured at CUM are only slightly lower by 1 to 3 dB, with an average difference by 1 dB, in terms of ΔA/A measure. This consistent difference may be attributed to the extensive listening experience and expertise in judging sound stimuli by all subjects from CUM who were musically trained sound engineering students. Extensive training in psychoacoustic tasks has long been recognized as significant in PA studies and has been specifically investigated in various studies [4],[13],[15],[38].

B. Harmonic complexes and tilted spectra

The experiment involving tilted harmonic spectra conducted in study [29] was an exact replication of that conducted by

Bernstein and Green [10]. However, in this case, only the flat spectrum and the tilted down spectrum were tested (see Fig. 9). The consideration of a tilted-up spectral change was excluded since such a spectral envelope is not common in musical signals. Typically, the music signal spectrum exhibits above 600-800 Hz a downward tilt with a slope of either 9, 12 or even 18 dB per octave, depending on the physical conditions imposed on the played instrument. Four conditions were tested: a flat initial spectrum with the signal adding either positive or negative tilt (denoted by '+' or '-' signs in Fig. 13), and -6- or -12-dB negative initial tilt with the addition of positive tilt in the signal (denoted by the '+' sign in the axis description in Fig. 13).

There were 25 harmonic components of a 200-Hz fundamental that extended in frequency up 5000 Hz in harmonic. As a reference condition, an exact replication of the stimulus employed in a study [10] was employed (21 components extending from 200 to 5000 Hz in frequency).

The results presented in Fig. 13 demonstrate a much stronger influence of subjects' experience in listening tasks compared what was observed in Fig. 12 for rippled spectra. The overall tendencies in both logarithmic and harmonic tilted complexes are same as those observed by Bernstein and Green [10]. However, it is worth noting that experienced subjects from CUM exhibited tilt detection thresholds for logarithmic complexes that were notably superior to those measured in study [10], by approximately 6 to almost 10 dB. Specifically, the measured thresholds ranged from -18 to -22 dB, while the thresholds in study [10] ranged from -11 to -17 dB. This difference in results is in line with the suggestions made by Drennan and Watson [38].

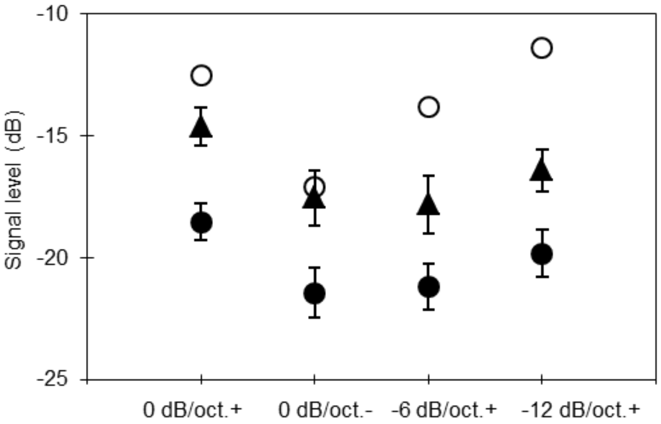

Fig.13. Thresholds for detecting tilt in the spectrum, with initial tilt of 0, -6, and -12 dB imposed on a 200-Hz harmonic complex (filled triangles). Mean of three subjects. Thresholds for the replicated experiment of Bernstein and Green [10] for the logarithmic complex are shown (filled circles), and original data from Bernstein and Green [10] (open circles). The error bars represent the standard error, calculated from 12 threshold estimations per data point. Data from [29]

In the case of harmonic complexes, the thresholds shown in Fig. 13 are approximately 4 dB higher than the thresholds obtained for logarithmic complexes. However, these thresholds are still lower than the thresholds obtained for unexperienced subjects in Bernstein and Green's study [10], and fall within a range of -14 to -18 dB (for ΔA/A) which corresponds to an increase in the level by only 1 to 1.5 dB.

C. *The Timbre Solfege course*

This section addresses an aspect of musical/sound engineering practice, comparing it with the fundamental findings of PA. The comparison of formant detection and profile analysis was previously discussed in detail by Żera [39]. This section will be focused on the essential concluding results.

Listening tasks involving formant discrimination, are integral components of the *Timbre Solfege* program, a course designed to provide comprehensive training in auditory sound evaluation to sound engineering students taught at the Music Acoustics Laboratory, Chopin University of Music in Warsaw [40] In contrast to the experiments concerned with PA, level roving is not used in *Timbre Solfege* exercises, therefore the students may make use of both simultaneous cross spectrum and successive level comparisons in their judgments. By comparing the data from PA experiments with the results of formant recognition within the *Timbre Solfege* program we can assess the extent to which the process of formant recognition is influenced by simultaneous comparisons in the tasks performed at CUM by sound engineering students.

In the experiment by Letowski and Rogala [41], detection thresholds for single 1/3-octave formants within wideband signals were measured. In Fig. 14, the thresholds for formant detection in white noise reported in [41] are compared with those obtained in a PA task involving the detection of an amplitude increment of a single component tone in a 21-component logarithmic complex [5],[35].

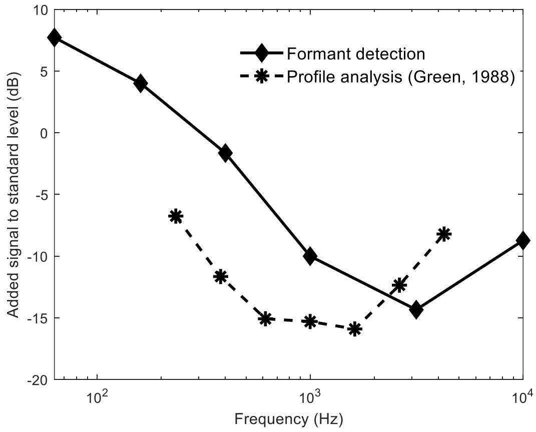

Fig. 14. The 1/3-octave formant detection in white noise (solid line, diamonds, 75% correct detection) and the detection thresholds for a single component of a complex with 21 logarithmically spaced components (dashed line, asterisks, 70.7% correct detection). Figure adapted from [39] with permission

The thresholds for formant detection in white noise, represented by the solid line in Fig. 14, are higher than those obtained for the detection of a single component in a 21-component logarithmic complex (dashed line). Formant detection thresholds decrease from +7.7 dB down to -14.3 dB as the formant frequency is increased from 63 to 3150 Hz. The wideband stimulus employed for the measurement of formant detection spans a broader frequency range compared to the range of a 21-component logarithmic complex in the PA task (200-5000 Hz). Both curves in the graph show a similar minimum of about –16 to –14 dB, although at different frequencies. This difference can be potentially linked to the constant power spectrum density of the white noise used in the formant detection task. Energy increment of +1 dB per filter (+3 dB per octave) as the centre frequency of the 1/3-octave bands increases bears a resemblance to the spectral tilt-up effect of the standard, a condition that was not explored as a single-component PA condition. In contrast, log-spaced components used in the auditory profile tasks produce approximately uniform energy distribution across the logarithmic frequency spectrum, in the entire frequency range (see Fig. 1). Rogala and Śliwka [32]-[34] measured formant discrimination thresholds, that is the minimum increase in level essential to detect a level change in an already existing 3- or 12-dB 1/3-octave formant in pink noise, for formant frequencies:125, 315, 1000, 3150, and 8000 Hz.

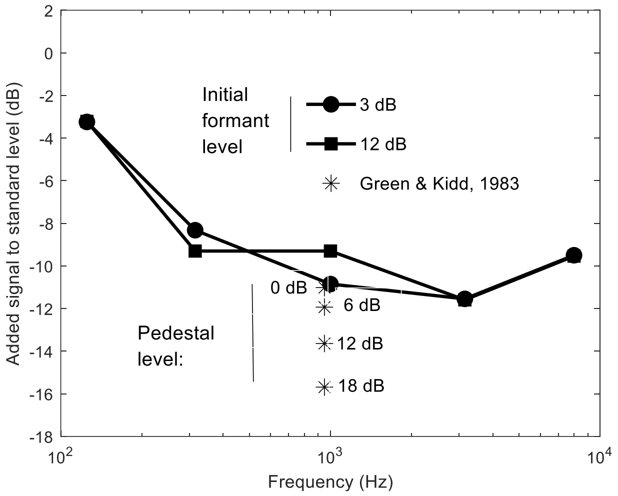

Fig. 15. The 1/3-octave formant discrimination (solid lines) and the results of the pedestal experiment conducted by Green and Kidd [2] (asterisks). Adapted from [39] with permission

The discrimination of formant level can be compared with a profile analysis task consisting in the detection of a level increment of a component within a pedestal [2]. Such an analogy mirrors the process of formant level discrimination and this comparison is shown in Fig. 15.The solid lines in Fig. 15 presents the formant detection thresholds for formants recalculated to the ΔA/A level used in profile analysis and due to differences in experimental procedures (oddity procedure in formant discrimination, 2AFC in PA) adjusted based on the detectability index d' (see [39] for details). The results of the pedestal experiment [2] are shown in Fig.15 as asterisks. The thresholds for pedestals of 0 and 6 dB correspond to the condition in which the initial formant level (formant pedestal) was 3 dB whereas the 12-dB pedestal corresponds to the 12-dB formant pedestal.

Formant discrimination thresholds are similar to thresholds in the PA tasks with 0- and 6-dB pedestals but are higher than the thresholds for the 12-dB pedestal. Interestingly, formant level discrimination thresholds do not change even when the initial 3-dB level (pedestal) is substituted with a 12-dB initial level. This observation had previously been linked to the principles of Weber's law.

A deeper exploration of the original data in [32]-[34]shows that subjects focus on the individual frequency channels relevant to formant discrimination and perform sequential comparisons rather than across-channel profile analysis. Such a listening strategy seems highly plausible, particularly the there was no application of signal level roving in the experiments concerning formant discrimination.

V. SUMMARY

This study provided an overview of experiments that established highly formalized methods, creating a distinct research area developed under the supervision of Professor David M. Green. The research focused on measuring the human ability to distinguish changes in the frequency content of sound.

The study addressed three key aspects. Firstly, it presented a summary of assumptions regarding the process involved in PA to separate the simultaneous cross-spectrum comparisons from successive single-channel level comparisons. Secondly, an overview of major experiments was provided, highlighting the variety of signals used, in terms of their narrowband and wideband spectral properties. Finally, the study presented examples of experiments conducted entirely separately from the main body of work on PA. For instance, formant detection studies were conducted among music students, to demonstrate the close relation of these studies with the PA research framework.

During the period of mentoring by Professor David A. Green, research on PA was lively, however, over time, it was partly abandoned. The present review aims to revive and highlight at least some of the achievements that were attained during that active period. It also emphasizes the importance of these achievements for other studies, particularly those of application nature, which have become increasingly common in various research fields nowadays.

REFERENCES

[1] D. M. Green, G. Kidd Jr., Maria C. Picardi, Successive versus simultaneous comparison in auditory intensity discrimination, J. Acoust. Soc. Am., vol. 73, pp. 639-643, 1983. https://doi.org/10.1121/1.389009

[2] D. M. Green, G. Kidd Jr., Further studies of auditory profile analysis, J. Acoust. Soc. Am., vol. 73, pp. 1260-1265, 1983. https://doi.org/10.1121/1.389274

[3] D. M. Green, Ch. R. Mason, G. Kidd Jr., Profile analysis: critical bands and duration, J. Acoust. Soc. Am., vol. 75, pp. 1163-1167, 1984. https://doi.org/10.1121/1.390765

[4] D. M. Green, Ch. R. Mason, Auditory profile analysis: frequency, phase, and Weber's law, J. Acoust. Soc. Am., vol. 77, pp. 1155-1161, 1985. https://doi.org/10.1121/1.392179

[5] D. M. Green, Z. A. Onsan, T. G. Forrest, Frequency effects in profile analysis and detecting complex spectral changes, J. Acoust. Soc. Am., vol, 81, pp., 692-699, 1987. https://doi.org/10.1121/1.394837

[6] L. R. Bernstein, D. M. Green, The profile-analysis bandwidth, J. Acoust. Soc. Am., vol. 81, pp. 1888-1895, 1987. https://doi.org/10.1121/1.394753

[7] L. R. Bernstein, D. M. Green, Detection of simple and complex changes of spectral shape, J. Acoust. Soc. Am., vol. 82, pp. 1584-1592, 1987. https://doi.org/10.1121/1.395147

[8] D. M. Green, 'Frequency' and the detection of spectral shape change, In. Auditory frequency selectivity, Eds.: B. C. J. Moore, R. D. Patterson, Plenum Press, NATO Series, 1988.

[9] G. Kidd, Jr.LC. R. Mason, A new technique for measuring spectral shape discrimination, J. Acoust. Soc. Am., vol. 91, pp. 2855-2864, 1992. https://doi.org/10.1121/1.402966

[10] L. R. Bernstein, D. M. Green, Detection changes in spectral shape: Uniform vs. non-uniform background spectra, Hear. Res., vol. 32, pp. 157-166, 1988.

[11] M. F. Spiegel, M. C. Picardi, D. M. Green, Signal and masker uncertainty in intensity discrimination, J. Acoust. Soc. Am., vol. 70, pp. 1015-1019, 1981. https://doi.org/10.1121/1.386951

[12] M. F. Spiegel, D. M. Green, Signal and masker uncertainty with noise maskers of varying duration, bandwidth, and center frequency, J. Acoust. Soc. Am., vol. 71, pp. 1204-1210, 1982. https://doi.org/10.1121/1.387769

[13] G. Kidd, Jr., C. R. Mason, D. M. Green, Auditory profile analysis of irregular sound spectra, J. Acoust. Soc. Am., vol. 79, pp. 1045-1053, 1986. https://doi.org/10.1121/1.393376

[14] J. J. Raney, V. M. Richards, Z. A. Onsan, D. M. Green, Signal uncertainty and psychometric functions in profile analysis, J. Acoust. Soc. Am., vol. 86, pp. 954-960, 1989. https://doi.org/10.1121/1.398730

[15] G. Kidd, Jr., C, R, Mason, R. M. Uchanski, M. A. Brantley, P. Shah, Evaluation of simple models of auditory profile analysis using random reference spectra, , J. Acoust. Soc. Am., vol. 90, pp. 1340-1354, 1991. https://doi.org/10.1121/1.401926

[16] J. Żera, Z. A. Onsan, Q. T. Nguyen, D. M. Green, Auditory profile analysis of harmonic signals: critical bands and duration, J. Acoust. Soc. Am., vol. 93, pp. 3431-3441, 1993. https://doi.org/10.1121/1.405673

[17] D. M. Green, T. G. Forrest, Profile analysis and backgroung noise, J. Acoust. Soc. Am., vol. 80, pp. 416-421, 1986. https://doi.org/10.1121/1.394092

[18] V. M. Richards, Z. A. Onsan, D. M. Green, Auditory profile analysis: Potential pitch cues, Hear. Res., vol. 39, pp. 27-36, 1989.

[19] C. R. Mason, G. Kidd Jr., T. E. Hanna, D. M. Green, profile analysis and level variation, Hear. Res., vol. 13, pp. 269-275, 1984.

[20] H. Dai, D. M. Green, Discrimination of spectral shape as a function of stimulus duration, J. Acoust. Soc. Am., vol. 93, pp. 957-965, 1993. https://doi.org/10.1121/1.405456

[21] H. Dai, D. M. Green, Auditory intensity perception: successive versus simultaneous across-channel discriminations, J. Acoust. Soc. Am., vol. 91, pp. 2845-2854, 1992. https://doi.org/10.1121/1.402965

[22] B. G. Berg, D. M. Green, Spectral weights and profile listening, J. Acoust. Soc. Am., vol. 88, pp. 758-766, 1990. https://doi.org/10.1121/1.399725

[23] D. M. Green, L. R. Bernstein, Profile analysis and speech perception, In. The psychophysics of speech perception, Ed. M. E. H. Schouten, NATO ASI Series, pp. 314-326, 1987. DOI:10.1007/978-94-009-3629-4

[24] C. C. Henn, C. W. Turner, Pure-tone increment detection in harmonic and inharmonic backgrounds, J. Acoust. Soc. Am., vol. 88, pp. 126-131, 1990. https://doi.org/10.1121/1.399958

[25] N. I. Hill, P. J. Bailey, Profile analysis with an asynchronous target: evidence for auditory grouping, J. Acoust. Soc. Am., vol. 102, pp. 477-481, 1997. https://doi.org/10.1121/1.419720

[26] W. Ellermeier, Detectability of increments and decrements in spectral profiles, J. Acoust. Soc. Am., vol. 99, pp. 3119-3125, 1996. https://doi.org/10.1121/1.414797

[27] H. Gockel, H. Colonius, Auditory profile analysis: is there perceptual constancy for spectral shape for stimuli roved in frequency? J. Acoust. Soc. Am., vol. 102, pp. 2311-2315, 1997. https://doi.org/10.1121/1.419640

[28] J. J. Lentz, V. M. Richards, M. R. Matiasek, Different auditory filter bandwidth estimates based on profile analysis, notched noise, and hybrid tasks, J. Acoust. Soc. Am., vol. 106, pp. 2779-2792, 1999. https://doi.org/10.1121/1.428137

[29] A. Ciesielski, Audibility of changes in timbre of sound made by formants, Master's Thesis, Chopin University of Music, Warsaw, pp.1-64, 2007 (in Polish).

[30] T. Letowski, Development of technical listening skills: Timbre solfeggio, J. Aud. Eng. Soc. 33, 1985, 240-244.

[31] T. Letowski, A. Miskiewicz, Developing of technical listening skills for sound quality assessment, Proceedings of Inter-Noise '95, Newport Beach, FL, 1995, 917-920.

[32] T. Rogala, P. Śliwka. Discrimination of formant amplitude in noise, Audio Engineering Society 138th Convention, May 7-10, Warsaw, 2015, Paper #9282.

[33] T. Rogala, Pink noise bandwidth discrimination, Audio Engineering Society 142th Convention, May 20-23, Berlin, 2017, Paper #9777.

[34] T. Rogala, Discrimination of formant frequency in pink noise. Audio Engineering Society 140th Convention, June 4-7, Paris, 2016, Paper #9583.

[35] D. M. Green, Profile analysis, Oxford University Press, New York – Oxford, 1988.

[36] E. Boer, "On the 'Residue' and the auditory pitch perception." In.: Handbook of Sensory Physiology. Auditory system. Clinical and special topics., Wolf D. Keidel and William D. Neff (Eds.), Vol. V/3, Chapter 13, pp. 479-583, 1976.

[37] W. R. Drennan, C. S. Watson, Sources of variation in profile analysis. II. Component spacing, dynamic changes, and roving level, J. Acoust. Soc. Am., vol. 110, pp. 2498-2504, 2001. https://doi.org/10.1121/1.1408311

[38] W. R. Drennan, C. S. Watson, Sources of variation in profile analysis. I. Individual differences and extended training, J. Acoust. Soc. Am., vol. 110, pp. 2491-2497, 2001. https://doi.org/10.1121/1.1408310

[39] J. Żera, Timbre Solfege and Auditory Profile Analysis, Vibrations in Physical Systems, 30, 2019122, 2019, pp. 1-8, 2019.

[40] A. Miśkiewicz, Timbre solfege: A course in technical listening for sound engineers, J. Aud. Eng. Soc. 40, 1992, 621-625.

[41] T. Letowski, T. Rogala. Formant perception: single formant. In: Sztuka słuchania (The Art of Listening), Chopin University of Music, Warszawa, 2015, 45-63.