

Musical note recognition based on the upper adjacent harmonics without the presence of the fundamental frequency

Original

Musical note recognition based on the upper adjacent harmonics without the presence of the fundamental frequency / Albera, R.; Urbanelli, A.; Lucisano, S.; Aprigliano, A.; Morando, L.; Amoroso, A.; Alexeev, M.; Albera, A.. - In: SCIENTIFIC REPORTS. - ISSN 2045-2322. - 15:1(2025). [10.1038/s41598-025-89454-7]

Availability:

This version is available at: 11583/3002520 since: 2025-08-24T09:43:37Z

Publisher:

Springer - Nature Research

Published

DOI:10.1038/s41598-025-89454-7

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)



OPEN Musical note recognition based on the upper adjacent harmonics without the presence of the fundamental frequency

Roberto Albera¹, Anastasia Urbanelli¹✉, Sergio Lucisano¹, Alessandra Aprigliano¹, Luca Morando², Antonio Amoroso², Maxim Alexeev² & Andrea Albera¹

Musical signals are complex periodic waveforms characterized by the sum of different frequencies. In a harmonic complex tone, the lowest frequency is called fundamental frequency (f_0), while the other frequencies are called harmonics, and their frequencies are integer multiples of the fundamental. The perceived pitch of a sound is correlated with the fundamental frequency, even though it may be impossible to hear f_0 in many situations. In these cases, it is possible to identify the pitch based on the upper consecutive harmonics. This study aimed to evaluate the identification of the notes based on the presence of consecutive harmonics only and to determine the importance of their distance from the fundamental frequency. The study was carried out on 30 normally hearing amateur musicians without perfect pitch. The acoustic signal was characterized by the association of four consecutive and two consecutive harmonics of the middle region notes of the piano keyboard. The correct identification rate ranged between 8 and 100%, with better identification occurring when more harmonics and lower frequencies were present. The results confirm that it is possible to identify a note solely based on the presence of harmonics near the fundamental frequency, especially if it is under 2000 Hz.

Keywords Harmonics, Fundamental note, Pitch, Musical note identification

Pitch identification of an acoustic signal is made possible by the precise capacity of the organ of Corti to identify all the frequencies composing the sound^{1,2}. This is possible thanks to the resonance characteristics of the basilar membrane; at this level, each frequency causes it to oscillate at different points with a distribution that goes from the apex to the base, respectively, for low and high frequencies³. Furthermore, the oscillation mechanism concerning the frequency is made even more precise by the contractile activity of the external hair cells^{1,4}.

Based on the basilar membrane vibration, it is possible to explain the pitch identification of a pure (sinusoidal) tone characterized by the presence of a single frequency. According to the temporal theory of pitch perception, the periodicity that characterizes the periodic waveform of every musical note is directly related to the pitch of the note itself. Thus, the pitch is often described as “the perceptual correlate of the periodicity of the sound’s waveform”⁵.

Musical note and voices are characterized by a more complex acoustic signal defined as complex periodic waveform (CPW)⁶. In CPW signal, the waveform is characterized by the presence of multiple frequencies that sound together. The lowest frequency is called fundamental frequency (f_0), which represents, in the case of chordophone instruments and of the vocal folds, the vibration of the entire chord. The other higher frequencies are called harmonics and are integer multiples of the fundamental frequency^{7,8}. In this case, the basilar membrane vibrates at different points simultaneously, and each point is related to one of the harmonics composing the CPW. In this way, the cochlea perfectly represents the musical or vocal acoustic signal^{9,10}.

In the CPW, the perceived pitch is mainly defined as the fundamental frequency (f_0) expressed in Hz. In the current convention, the tuning reference of a piano is the keyboard’s fifth A note (A4), which has a fundamental frequency ca. 440 Hz. Based on this value and the mathematical relationship between the notes, it is possible to determine the fundamental frequency of each note. Theoretically, every octave interval (frequency ratio 2:1) is divided into 12 equally sized semitones (in equal-tempered tuning). However, in practice, the tuning of the piano is slightly stretched^{11,12}. The 88 keys of the piano keyboard cover the fundamental frequencies from 27.5 to 4186 Hz (a0 to C8)^{2,13}.

¹Otorhinolaryngology Unit, Department of Surgical Sciences, University of Turin, Via G. Verdi, 8, 10124 Turin, Italy.

²Department of Physics, University of Turin, Turin, Italy. ✉email: anastasia.urbanelli@gmail.com

Pitch perception may also be influenced by timbre and loudness due to filtering imposed by different vocal tracts of instrument bodies. About this point, McPherson et al. demonstrated that pitch discrimination was less accurate when musical notes derived from different instruments compared to when the instrument was the same and was biased by the associated timbre differences. However, they also demonstrated that relative pitch judgments are not invariant to timbre, even when such judgments are based on representations of f_0 ¹⁴.

According to some previous authors, pitch estimation relies on two components: spectral pitch (which involves analytically listening to individual harmonics) and virtual pitch (which involves holistic listening of a single evoked pitch). Moreover, there are several neurophysiological processes that may influence the perceived pitch of a complex tone; they consist in, among others, pitch bending of an individual harmonic, masking of harmonics, and irregular auditory sensitivity in listeners^{15–17}.

Due to masking or filtering, it can be difficult to hear the fundamental frequency of a note in the presence of background noise, where the acoustic pressure is concentrated on low frequencies^{1,6}, or in listening to filtered acoustic signals such as MP3 recordings, radio, or telephone conversation¹⁴. In addition, the hearing sensitivity of a human is considerably attenuated in low frequencies¹⁸.

In these cases, however, it is possible to identify the pitch of the notes based on the perception of the consecutive harmonics, which, being higher in frequency, are less easily masked. Since the harmonics are in a mathematical relationship with the fundamental, being its integer multiples (in the case of harmonic tones), it is possible to perceive the pitch of the fundamental frequency, even if it is not hearable, based on the consecutive harmonics. For example, in the 500–600 Hz frequency pair, the perceived fundamental is 100 Hz¹⁴ (greatest common divisor).

This study aimed to evaluate the pitch identification of acoustic signals characterized by the presence of two or four consecutive harmonics without fundamental with amateur musicians without perfect pitch¹⁹. Moreover, we aimed to evaluate the influence of the distance between the theoretical fundamentals and the presented harmonics.

Materials and methods

This study was performed following the ethics standards laid down in the 1964 Declaration of Helsinki and informed written consent was obtained from all subjects. The study was approved by the University of Turin's ethics committee, and the study's aim was clearly explained to each participant to obtain informed consent to be subjected to the test.

The study was performed on a group of 60 participants, 26 (43%) male and 34 (57%) female, aged between 21 and 75 years (mean of 37.2 years). Participants with non-professional ability were recruited to play an instrument but lacking perfect pitch.

In the design phase of the study, we first asked 6 musicians with perfect pitch to identify the presented notes with harmonics without fundamentals. In these participants, unlike what happens in people without perfect pitch, the sensation generated by the stimulus sent was that of the simultaneous presence of single notes having the same fundamental frequency of the harmonics presented. Those who have perfect pitch, therefore, have the same perceptive behavior as a person without this characteristic but to whom the different frequencies sent are not sent simultaneously²⁰. Since the aim of the study was to verify the possibility of note identification based on the harmonics without fundamentals, subjects with perfect pitch were excluded from the study.

We also tested some participants lacking the ability to play an instrument but for many of these participants the identification of the note on the keyboard caused significant tension and anxiety, making the note was identified with considerable difficulty and the outcome could not be considered reliable. Therefore, even these participants were excluded by the study group.

Hence, we decided to test people who were able to play on a keyboard but did not have perfect pitch. To avoid interference with cochlear function we included in the study group only participants free from ear pathologies and with normal audiometric threshold.

Among the inclusion criteria, there were also a negative history of ear diseases and an audiometric threshold equal to or less than 25 dB at frequencies between 125 and 8000 Hz.

The sound stimuli were presented by a tone generator that produced a complex waveform made up of two or four pure tones with the frequencies chosen by the Authors. The generator is based around a SGTL5000 stereo codec with headphone amplifier (<https://www.nxp.com/docs/en/data-sheet/SGTL5000.pdf>), mounted on an audio adaptor (https://www.pjrc.com/store/teensy3_audio.html), and controlled over a Freescale based (<https://www.nxp.com/docs/en/data-sheet/K20P64M72SF1.pdf>) Teensy 3.2 board (<https://www.pjrc.com/store/teensy32.html>). The SGTL5000 in our configuration is powered by 3.3 V and is driving the analog output to the Audiotechnica headphones (model: ATH-M50X) in a 24bit data regime with a 44.1 kHz sampling frequency. In this configuration, 32-Ohm headphones achieved a signal-to-noise ratio (SNR) of 100 dB with a total harmonic distortion plus noise (THD + N) of -88 dB and a frequency response of ± 0.11 dB.

All frequencies started simultaneously had the same sound pressure level (SPL) at 65 dB.

The test was divided in two sessions. In the first session, the harmonics 2–3–4–5 of the notes C3 and G3 were presented to all participants, so they were invited to identify the notes listened on an electronic keyboard without temporal limit (Table 1). Then, to the same participants were presented the harmonics 3–4–5–6 of the notes E3 and A3 and asked to identify the note following the same protocol (Table 2). In the second session, two consecutive harmonics, 3–4, 4–5, 5–6, 6–7, 7–8, 8–9 and 9–10 of the notes Ab4, D4, E4, F4, A4, Eb4, Db4 (Table 3) were presented to all participants and they were asked to identify the note with the same protocol similar to the first session.

The explanation for the choice of notes used in the study lies in the fact that these notes, being in the middle region notes of the piano keyboard, are more pleasing to auditory perception and facilitate the listening and the recognition of the presented harmonics.

Note	Fundamental (Hz)	Harmonics 2–3–4–5 (Hz)
C3	130.81	261.62
		392.43
		523.24
		654.05
G3	195.99	391.98
		587.97
		783.96
		979.95

Table 1. Notes utilized for the test, frequency of their fundamentals according to the actual musical codification and frequency of the harmonics 2–3–4–5 present in the acoustic signal of the test.

Note	Fundamental (Hz)	Harmonics 3–4–5–6 (Hz)
E3	164.81	494.43
		659.24
		824.05
		988.86
A3	220	660
		880
		1100
		1320

Table 2. Notes utilized for the test, frequency of their fundamental according to the actual musical codification and frequency of the harmonics 3–4–5–6 present in the acoustic signal.

Note	Fundamental (Hz)	Presented harmonics	Frequencies of the presented harmonics (Hz)
Ab4	415.30	3	1245.90
		4	1661.20
D4	293.66	4	1174.64
		5	1468.3
E4	329.63	5	1648.15
		6	1977.78
F4	349.23	6	2095.38
		7	2444.61
A4	440	7	3080.00
		8	3520.00
Eb4	311.13	8	2489.04
		9	2800.17
Db4	277.18	9	2494.62
		10	2771.8

Table 3. Notes utilized for the test, frequency of their fundamental according to the actual musical codification, progressive number of the harmonic association presented and frequency of the two harmonics presented.

Results

In Tables 4 and 5, the rate of correct identification of the notes in relation to the group of harmonics presented is reported. The correct identification rate of the note in the absence of the fundamental is between 88 and 100% listening to harmonics 2–5 and between 82 and 96% listening harmonics 3–6. The correct identification rate is higher in the presence of harmonics nearer to fundamental and of notes with a higher fundamental frequency. Table 6 shows the correct identification rate of the note in relation to the harmonics presented, respectively 3–4, 4–5, 5–6, 6–7, 7–8, 8–9 and 9–10. The correct identification ranges from 6 to 76%. Even in this case, the higher rate of correct identification was obtained by listening to a couple of harmonics nearer to the fundamental and lower frequencies. This pattern is also reported also in Fig. 1, where the rate of correct identification is related to the frequency of the lowest couple of harmonics presented to participants. The slope of the line clearly shows

Note	Fundamental (Hz)	Harmonics 2–5 (Hz)	Rate of correct identification of the notes Harmonics 2–5 (%)
C3	130.81	261.62	100
		392.43	
		523.24	
		654.05	
G3	195.99	391.98	94
		587.97	
		783.96	
		979.95	

Table 4. Rate of correct identification of the notes C3 and G3 on the basis of the presentation of harmonics 2–3–4–5.

Note	Fundamental (Hz)	Harmonics 3–6 (Hz)	Rate of correct identification of the notes Harmonics 3–6 (%)
E3	164.81	494.43	94
		659.24	
		824.05	
		988.86	
A3	220	660	82
		880	
		1100	
		1320	

Table 5. Rate of correct identification of the notes E3 and A3 on the basis of the presentation of harmonics 3–4–5–6.

Note	Fundamental (Hz)	Harmonics	Frequency of the two presented harmonics (Hz)	Rate of correct identification of the Notes (%)
Ab4	415.30	3–4	1245.90	76
			1661.20	
D4	293.66	4–5	1174.64	71
			1468.30	
E4	329.63	5–6	1648.15	35
			1977.78	
F4	349.23	6–7	2095.38	12
			2444.61	
A4	440.00	7–8	3080.00	35
			3520.00	
Eb4	311.13	8–9	2489.04	24
			2800.17	
Cb4	277.18	9–10	2494.62	6
			2771.80	

Table 6. Rate of correct identification of the notes in the different couple of harmonics presented to the participants.

that the identification of the note is easier if the frequency of the first harmonic presented has a frequency less than 1500–2000 Hz. Then, the success rate of identification falls to lower values.

Discussion

The results obtained from the study confirm that it is possible to identify a note based on the presence of four harmonics¹⁹. The correct note identification was higher if the harmonics presented were nearer to the fundamental (2–5 versus 3–6). If the signal is composed by two harmonics only, the identification is still possible, but the rate of correct identification dropped to lower values (6% to 76%), and even in this case, it is higher if tones presented were closer to the fundamental.

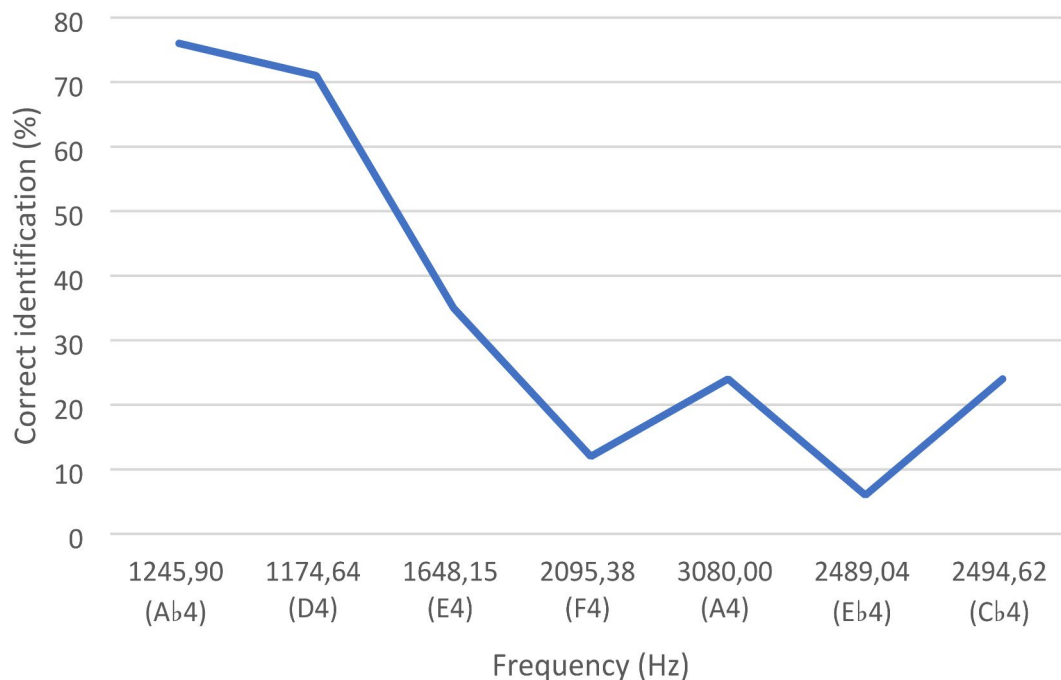


Fig. 1. Rate of correct identification of the note in relationship to the frequency of the lowest harmonic presented. The numbers in the brackets refer to the frequency of the first harmonic.

The identification of the note without the perception of the fundamental is possible since two or more frequencies correlated with each other as consecutive integer multiples of a lower frequency allow the reconstruction of the missing fundamental of the signal and confers the tonal characteristic perceived²⁰. This modality of perception requires the simultaneous presentation of the single frequency components (harmonics). Otherwise, these are identified as signals not correlated with each other and heard each one as single associated note^{21–23}.

Previous studies on pitch note discrimination based on harmonics listening have been conducted, so leading to the concept that pitch discrimination abilities and pitch salience decrease dramatically when harmonics of a complex below the tenth are removed^{24,25}. In the recent past, also Graves et al. demonstrated that functional pitch perception was possible within combinations and mixtures of different harmonics, even when the stimuli were filtered to fall within the same overlapping spectral region²⁶.

The concept of the “critical band” was introduced for the first time in 1933 by Harvey Fletcher and defines the frequency range of the so-called “auditory filter” created by the cochlea²¹. The critical bandwidth refers to the range of audio frequencies where a second tone disrupts the perception of the first tone through auditory masking. Critical bands are also linked to auditory masking effects, which reduce the detectability of a sound signal when it coexists with a second signal of greater intensity within the same critical band. The implications of masking phenomena are extensive, encompassing a nuanced interplay between loudness and intensity.

When listening to just two harmonics, the critical point in note recognition was placed between 1500 and 2000 Hz. To explain this pattern, it is necessary to remember that the fundamental identification based on at least two harmonics can take place only if both harmonics can be heard. If the two harmonics are too close to each other, they are analyzed at very near points along the basilar membrane. If the distance is lesser than 1 mm, the masking phenomenon occurs, and one of the two tones is not perceived^{22,27}.

The arrangement of the points of maximum cochlear oscillation does not follow a linear relationship with frequency. Instead, it is based on a logarithmic ratio in base 2; this means the distance between two points of maximum oscillation is constant, 4 mm, in relation to the doubling of the frequency²¹. As a result, the distance between the points of maximum oscillation of the basilar membrane induced by two tones with a constant difference in frequency, as occurs in our experiment, is lesser in the high-frequency region. Therefore, in the high-frequency range, it is easier for one of the two tones to be masked by the other. Consequently, it becomes more difficult to identify the fundamental.

In normal musical listening it is always possible to identify the fundamental even if not perceived while in our study this event does not always occur. The low performance reported in our experiment can be explained by the adoption of synthetic tones in which all the harmonics have the same intensity, situation far from normality. Moreover, in real music listening for a normally hearing subject, it is never possible that only two or four harmonics are perceived. Since the masking effect is mainly due to the background noise which is characterized by a higher acoustic pressure up to about 500 Hz²⁴, all the harmonics above 500 Hz, even for the notes whose fundamental is below this frequency (C5, i.e. the 52nd note of the piano keyboard) can be clearly distinguished, leading to the identification of the fundamental frequency of those perceived harmonics.

Conclusion

Our study confirms that it is possible to identify a note solely based on the presence of harmonics near the fundamental frequency, and identification success is higher if f_0 is under 2000 Hz. Moreover, in presence of more harmonics we have demonstrated a higher rate of correct identifications.

Our results could have implications for models and computational algorithms for pitch determination. A better understanding of the mechanisms humans use for fundamental note identification should lead to improved computer listening capabilities for the same tasks.

Data availability

All data pertaining to this systematic review are available from the corresponding author upon reasonable request.

Received: 6 February 2024; Accepted: 5 February 2025

Published online: 24 April 2025

References

- Benson, D. J. Music. A Mathematical Offering. Cambridge University Press. Preprint at https://www.logosfoundation.org/kursus/music_math.pdf (2008).
- Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S. & Robles, L. Basilar-membrane responses to tones at the base of the chinchilla cochlea. *J. Acoust. Soc. Am.* **101**(4), 2151–2163. <https://doi.org/10.1121/1.418265> (1997).
- Chan, W. X., Lee, S. H., Kim, N., Shin, C. S. & Yoon, Y. J. Mechanical model of an arched basilar membrane in the gerbil cochlea. *Hear Res.* **345**, 1–9. <https://doi.org/10.1016/j.heares.2016.12.003> (2017).
- Jabeen, T., Holt, J. C., Becker, J. R. & Nam, J. H. Interactions between passive and active vibrations in the organ of corti in vitro. *Biophys J.* **119**(2), 314–325. <https://doi.org/10.1016/j.bpj.2020.06.011> (2020).
- Laudanski, J., Zheng, Y., Brette, R. A structural theory of pitch (1,2,3). *eNeuro.* **1**(1). <https://doi.org/10.1523/ENEURO.0033-14.2014> (2014).
- Duifhuis, H., Willems, L. F. & Sluyter, R. J. Measurement of pitch in speech: an implementation of Goldstein's theory of pitch perception. *J. Acoust. Soc. Am.* **71**(6), 1568–1580. <https://doi.org/10.1121/1.387811> (1982).
- Bennet, W. R., Morrison, A. C. H. *The Science of Musical Sound*. Springer Cham (2008).
- Suzuki, H. Vibration and sound radiation of a piano sound-board. *J. Acoust. Soc. Am.* **80**, 1573–1582 (1986).
- Dai, H. On the relative influence of individual harmonics on pitch judgment. *J. Acoust. Soc. Am.* **107**, 953–959 (2000).
- Moore, B. C. J., Glasberg, B. R. & Peters, R. W. Relative dominance of individual partials in determining the pitch of complex tones. *J. Acoust. Soc. Am.* **77**, 1853–1860 (1985).
- McPherson, M. J. & McDermott, J. H. Relative pitch representations and invariance to timbre. *Cognition.* **232**, 105327. <https://doi.org/10.1016/j.cognition.2022.105327> (2023).
- Schuck, O. & Young, R. Observations on the vibrations of piano strings. *J. Acoust. Soc. Am.* **15**(1), 1–11 (1943).
- Jaatinen, J. & Pätynen, J. Effect of inharmonicity on pitch perception and subjective tuning of piano tones. *J. Acoust. Soc. Am.* **152**(2), 1146. <https://doi.org/10.1121/10.0013572> (2022).
- Giordano, N. Sound production by a vibrating piano soundboard: experiment. *J. Acoust. Soc. Am.* **103**, 2128–2133 (1996).
- Jaatinen, J., Pätynen, J. & Lokki, T. Uncertainty in tuning evaluation with low-register complex tones of orchestra instruments. *Acta Acustica.* **5**(49), 1–13 (2021).
- Terhardt, E., Stoll, G. & Seewann, M. Pitch of complex signals according to virtual-pitch theory: Tests, examples, and predictions. *J. Acoust. Soc. Am.* **71**(3), 671–678 (1982).
- Terhardt, E., Stoll, G. & Seewann, M. Algorithm for extraction of pitch and pitch salience from complex tonal signals. *J. Acoust. Soc. Am.* **71**(3), 679–688 (1982).
- Suzuki, Y. & Takeshima, H. Equal-loudness-level contours for pure tones. *J. Acoust. Soc. Am.* **116**(2), 918–933. <https://doi.org/10.1121/1.1763601> (2004).
- Kim, J. Analysis of factors affecting output levels and frequencies of MP3 players. *Korean J. Audiol.* **17**(2), 59–64. <https://doi.org/10.7874/kja.2013.17.2.59> (2013).
- Deutsch, D., Henthorn, T. & Dolson, M. Absolute pitch. *Music Percept.* **21**, 339–356 (2004).
- Pierce, J. R. *The Science of Musical Sound*. Scientific American Books New York. Preprint at <https://archive.org/details/scienceofmusical0000pier> (1983).
- Scheffers, M. T. Simulation of auditory analysis of pitch: an elaboration on the DWS pitch meter. *J. Acoust. Soc. Am.* **74**(6), 1716–1725. <https://doi.org/10.1121/1.390280> (1983).
- Preisler, A. The influence of spectral composition of complex tones and of musical experience on the perceptibility of virtual pitch. *Percept Psychophys.* **54**(5), 589–603. <https://doi.org/10.3758/bf03211783> (1993).
- Houtsama, A. J. M. & Smyrznyski, J. Pitch identification and discrimination for complex tones with many harmonics. *J. Acoust. Soc. Am.* **87**, 304–310 (1990).
- Shackleton, T. M. & Carlyon, R. P. The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. *J. Acoust. Soc. Am.* **95**(6), 3529–3540. <https://doi.org/10.1121/1.409970> (1994).
- Graves, J. E. & Oxenham, A. J. Pitch discrimination with mixtures of three concurrent harmonic complexes. *J. Acoust. Soc. Am.* **145**(4), 2072. <https://doi.org/10.1121/1.5096639> (2019).
- Greenwood, D. D. Auditory masking and the critical band. *J. Acoust. Soc. Am.* **33**, 484–502 (1961).

Acknowledgements

None.

Author contributions

Design of the work: R. A., S. L., A. Al.; writing: R. A., A., U.; data collection: A. Ap., L. M., A. Am., M. A.; final approval: A. A.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Competing interests

The authors declare no competing interests.

Financial interests

The authors declare they have no financial interests.

Informed consent

The review did not involve animals. Informed consent was collected from all participants of the study.

Additional information

Correspondence and requests for materials should be addressed to A.U.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025