Biodiversity Journal, 2023,14 (2): 303–314 https://doi.org/10.31396/Biodiv.Jour.2023.14.2.303.314

Hidden, Relevant High-Frequency Bioacoustic Features Revealed by Extreme Amplification of Digital 16-bit PCM High-Resolution Field Recordings

Cesare Brizio

World Biodiversity Association, Museo Civico di Storia Naturale di Verona, Lungadige Porta Vittoria, 9 - 37129 Verona, Italy; e-mail: briziocesare@gmail.com

ABSTRACT Building on the author's previous usage of amplification above –0dB digital full scale as a tool to improve the comparability of high-resolution (250 kHz sampling frequency) and lower-resolution digital recordings, the effects of extreme amplification (+30 dB Digital Full Scale) on 250 kHz digital recordings are explored, showing how previously unobserved native, relevant song features in the frequency domain may emerge in the inaudible band thanks to the apparently unadvisable process of amplification above zero dB Digital Full Scale.

KEY WORDS Bioacoustics; Digital Recordings; Destructive Clipping; High-Resolution audio.

Received 12.01.2023; accepted 22.02.2023; published online 10.04.2023

INTRODUCTION

The process of analog/digital conversion inherent in digital audio recording has a discrete nature: the continuum of an analogic signal is fractioned in as many discrete units, as the sampling rate allows following the Whittaker–Nyquist–Shannon cardinal theorem of interpolation (see e.g. Nyquist, 1928 and Shannon, 1949). In the recording phase, each sample is separately assigned its specific value, directly proportional to the sound pressure collected by the microphone/recorder equipment during the sampling tempuscule and depending on the available bit depth. During the post-production phase, interventions such as amplification affect separately each sample, depending on its original state (Boulanger & Lazzarini, 2011).

The recent availability of 32–bit Float Audio, capable of capturing up to 1,528 dB (+770dB / –758dB) (Matsushita et al., 1989; Sound Devices, 2022) provides the ability to record at any distance, including the closest proximity, without incurring in destructive clipping. On the contrary, both high-resolution and low-resolution recordings taken at an audio bit depth of 16-bit may be affected by that adverse effect as soon as –0dB digital full scale, the threshold of wave distortion by equipment overdriving (from now on referred to as the "Clipping threshold"), is exceeded by excessive gain of the recording equipment (microphone or recorder) or by excessive proximity to the sound source (Smith, 2007). For that reason, when recording at 16–bit care must be taken to ensure that the recording takes place at an optimal distance, empirically set by thumb rules such as keeping the sound pressure (as observed by the Vu-meters of the recording equipment) at the reasonable level of around –15dBfs, with peaks not exceeding –9dBfs. Even though in the post-production phase some digital amplification may take place to improve the clarity of the recording, common sense dictates that there should be no reason to amplify a recording above the clipping threshold (a premise that will be challenged more below).

Whatever the cause, undesired effect or on-purpose interventions such as digital amplification, the clipping phenomenon does not equally affect every sample in a digital recording: while some may exceed the clipping threshold, others (or most) may escape that fate. Seemingly, the far-reaching consequences of that obvious consideration still did not fully translate in methodologies: this paper illustrates a possible practical application of the counter-intuitive practice of extreme digital amplification of 16-bit PCM audio recordings. In this paper, terms such as "extreme amplification" and "over-amplification" are interchangeably used to express the same concept.

To set the stage for explanations, it should be mentioned that the author (Brizio & Buzzetti, 2014; Brizio et al., 2020) effectively employed low cost USB microphones ("ultrasound microphones") capable of recording ultrasounds thanks to high sampling frequency (250 kHz), obtaining results consistent with those of 96 kHz digital recorders, and recognized or described *de novo* the song of several Orthopteran species on the basis of 250 kHz recordings. That way, novel information was derived from an analysis of the high frequency components of the songs, previously unreported in the scientific literature, revealing how an observation limited to audible frequencies, or to the 0–48 kHz range provided by 96 kHz recordings, may deliver incomplete or misleading results.

High-resolution (wide band) recordings, extensively encompassing the inaudible range, are obtained by special microphones and recording equipment, such as those illustrated in Materials and Methods, whose dynamic range and frequency response may differ substantially from that of the microphones used for the 44, 48 or 96 kHz recordings, still making up the bulk of the audio samples stored in publicly accessible or institutional audio repositories.

As a consequence of the often radical difference in the frequency response of the different equipment, insurmountable difficulties may arise when comparing with the unaided ear a 250 kHz recording (0–125 kHz range) with reference audio at lower resolution, and similar problems may affect also the comparison by digital means such as pressure/time envelopes and time/frequency spectrograms.

To address that issue, Brizio & Buzzetti (2014) proposed a specific comparison protocol (Fig. 1) based on the processing of a copy of the original recording: the copy is low-pass filtered preserving the audible frequencies, then incrementally amplified, even above the clipping threshold if needed, until the audible components emerge with enough clarity to allow a comparison with the low-resolution reference audio. Furthermore, under specific conditions, the 2014 protocol demands to take on purpose a destructively clipped recording for a subsequent low-pass filtering: in fact, due to the ultrasound microphone's lower sensitivity to the audible band, destructive clipping will most probably preserve the underrepresented, audible portion of the recording, that can be suitably amplified to allow comparison with low-resolution audio files.

While the 2014 amplification-based protocol used extreme amplification as a tool to improve comparability between high-resolution and low-resolution audio files, the author recently became aware that amplification above the clipping threshold may equally serve another useful purpose: more specifically, the author observed that frequency/sound pressure analyses performed on an over-amplified copy of unclipped original digital audio, even though not fully reliable in their entirety, can reveal native (non-artefact) song features present in the original recording, but unobserved until highlighted by the extreme amplification process.

MATERIAL AND METHODS

Findings are based on the analysis of the song of four orthopteran species listed in Table 1, recorded in Italy, but may apply to any high resolution 16–bit recording containing hidden high-frequency features. All the songs include relevant high frequency (above 40 kHz) components: two fall into the High-Q and two in the Low-Q categories.

Q is the dimensionless parameter that describes how underdamped an oscillator or resonator is. Two definitions of Q have prevailed in literature since its first appearance in 1914 (see Green, 1955). They become approximatively equivalent as Q becomes larger and damping decreases. In terms of frequency-to-bandwidth ratio, Q is defined as:

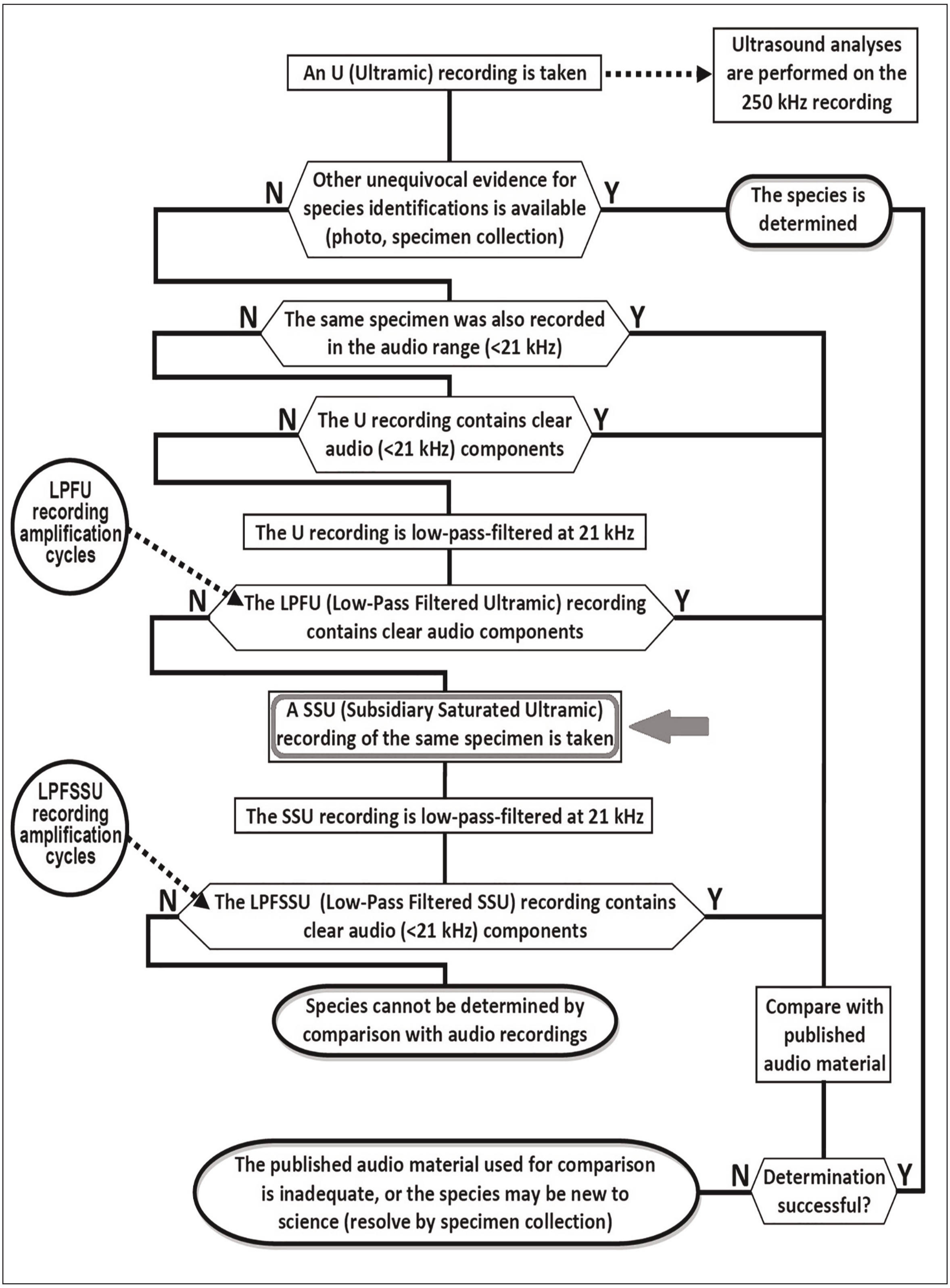

Figure 1. Suggested protocol to allow species identification form Ultramic 250 (sampling frequency: 250 kHz) recordings (flow chart) – from Brizio & Buzzetti, 2014. An arrow marks the protocol step requiring the obtainment of an on-purpose destructively-clipped ultrasound recording.

$$Q \stackrel{\text{def}}{=} \frac{f_r}{\Delta f} = \frac{\omega_r}{\Delta\omega}$$

where fr is the resonant frequency, Δf is the resonance width or full width at half maximum i.e. the bandwidth over which the power of vibration is greater than half the power at the resonant frequency, ωr = 2π fr is the angular resonant frequency, and Δω is the angular half-power bandwidth. The other common nearly equivalent definition for Q is the ratio of the energy stored in the oscillating resonator to the energy dissipated per cycle by damping processes:

$$Q \stackrel{\text{def}}{=} 2\pi \times \frac{\text{energy stored}}{\text{energy dissipated per cycle}} = 2\pi f_r \times \frac{\text{energy stored}}{\text{power loss}}$$

High-Q sound (Elsner & Popov, 1978; Montealegre-Z. & Morris, 1999) results in one or more (e.g. Gryllidae) isolated peaks of frequency, clearly distinguishable from the rest of the frequency emission and emerging as thin, well-defined horizontal lines in the time/frequency spectrograms. On the other hand, "wide band" or "low-Q" sound display wide and ill-defined horizontal spectral band, in which sometimes is possible to distinguish spectral subpeaks.

Ultrasound monophonic recordings at 250 kHz sampling frequency were obtained by using a Dodotronic Ultramic 250 microphone connected via USB cable to an Asus Eee PC 1225B notebook pc, using SeaWave software by CIBRA, under Windows 10 64–bit operating system. The song of *Melanogryllus desertus* was recorded with the same microphone using a tablet PC Asus K013 and the Bat Recorder software under Android operating system.

Sound description includes pressure envelopes (relative pressure in dB full scale), time/frequency spectrograms, and frequency/sound pressure analysis diagrams, generated on a Gigabyte Brix desktop computer by Adobe Audition 1.0 software running under Windows 10 64bit operating system. In the frequency/sound pressure visualization, the noise of the AD converter may appear in the noise floor profile as a wide, ill-defined hump centered at around 50 kHz and extending between 40 and 60 kHz.

Figures are based on Adobe Audition screenshots, post-produced with Adobe Photoshop Elements, as needed to ensure optimal contrast. MS-Paint was used to draw custom horizontal/vertical reference rulers. Visual improvements did not alter the data or the analysis results.

Quantitative measures such as peak frequencies were performed in Adobe Audition, directly on screen. The entire frequency range available in each recording was thoroughly analyzed, but to optimize informativeness, the frequency/pressure observation window was customized in each picture: sound pressure axes are limited at the top by the highest volume observed and at the bottom by the bottom noise at any frequency.

In each recording, the relative relevance of the different frequency bands, including those emerging from the amplification process described more under, is strongly affected by the distance between the singing specimen and the microphone. Issues out of the scope of this paper, such as the propagation distance of the inaudible components as a whole, or of specific narrow inaudible bands, compared to the propagation distance of audible frequencies are not explored here. From basic acoustic principles, it can be expected that – besides to the 1/r law of distance damping – the absorption of high frequencies, exhaustively covered e.g. by Vladišauskas & Jakevičius (2004), and by Bass et Al. (1984) rises with frequency, temperature and humidity – indicatively, in standard atmosphere it may amount to 5 dB/m for frequencies in the order of magnitude of 100 kHz. Table 1 includes a column, "Recording Distance", from which this datum can be appreciated. Operationally, knowing that the highest frequencies are more sensible to atmospheric dumping, while field-recording the author strived to get as close as possible to the singing specimens, but did not systematically take recordings at different distances for comparative purposes.

The frequency resolution of FFT-based analyses is directly proportional to FFT size (Welch et al., 2012; National Instruments Corporation, 2018). Concerning the expected size of the pictures in this article, an FFT size of 8192 byte was chosen as the best compromise between detail and smoothness of the picture. Frequency analyses were generated with the Blackman-Harris method (Blackman, 1958; Harris, 1978; Nuttall, 1981) by scanning a continuous interval of the audio samples as per Table 1.

It should be stressed that the method described herein under should be applied to high-quality, unclipped recordings, where no part of the original audio, or just a minimal part, should be at or above the clipping threshold: an heavily clipped recording may have irreversibly lost also those faint high-frequency components that the amplification phase is aimed at discovering, and that may be collected only at close recording distance, e.g. in recordings where the song hovers at no less than -15 dBfs with peaks up to -5dBfs or more.

For each of the species considered, one recording was selected and a copy was obtained by "save as" option under Adobe Audition. The copy was strongly amplified (+30dB) and two separate analyses, original and copy, were performed. Amplification was performed via the relevant menu option provided by Adobe Audition and, to that purpose, by any similar software. The theory behind the software-based PCM Audio amplification process is outside the scope of this paper, and is exhaustively covered in Boulanger & Lazzarini (2011).

Five illustrations, time-pressure envelope and time-frequency spectrogram of the two versions, plus a compared frequency-pressure analysis are provided for each species considered: a reference grid was not superposed on the figures because, considering their size, it would have hampered rather than facilitated their reading. The Results section describes the differences observed after the extreme amplification process.

RESULTS

Orthopterological comments relative to the bioacoustic significance of the findings illustrated here are outside the scope of this paper, and may be covered by a future publication. The aim of this contribution is to describe in general terms the previously unobserved features emerging from the amplification process, illustrated in the figures and summarized as follows:

- obviously a non-selective amplification process increases proportionally signal (the emitted song) and background noise;

- it should be noted that the different profile of the over-amplified frequency analyses, that replicate the "peaks" but not the "valleys" of the original analyses, is due to the concurrent amplification of the noise, whose floor raises by 30 dBfs, with an effect similar to that observed in the time/pressure envelopes; the invasiveness of noise in the over-amplified recording is due to the fact that, while the song is restricted to specific times and frequencies, noise is ubiquitous and often continuous, and its intensity may be only marginally lower or subequal to the intensity (sound pressure) of the previously hidden, now emerging high-frequency features.

Species	Recording Date	Q	Locality (Italy)	Recording Distance	Interval Analyzed for Frequency / Sound Pressure	Highest frequency components, original recording	Highest frequency components, over-amplified recording
Natula averni (Costa, 1855)	12 May 2022	High	Portixeddu, Fluminimaggiore, South Sardinia Province	30 cm	23″	60 kHz	95 kHz
Melanogryllus desertus (Pallas, 1771)	06 May 2020	High	Poggio Renatico, Ferrara Province	80 cm	49″	55 kHz	110 kHz
Leptophyes laticauda (Frivaldszky, 1868)	08 August 2012	Low	Madonna dell' Acero, Lizzano in Belvedere, Bologna Province	80 cm	60″	48 kHz	107 kHz
Ctenodecticus bolivari Targioni-Tozzetti, 1881	09 September 2019	Low	Capo Pecora, Arbus, South Sardinia Province	30 cm	20″	103 kHz	123 kHz

Table 1. Details about the audio files (Sampling Frequency: 250 kHz) used for the analyses.

- as expected after such a drastic amplification, the time/pressure envelope loses most of its significance. For example, reading time-related features from the resulting visualization is problematic;

- any audio sample brought above the clipping threshold by the amplification process is irreversibly altered, thus compromising the reliability of quantitative, absolute or relative, measures of sound pressure.

Despite those obvious drawbacks, the over-amplification process has also interested audio samples whose feeble volume has kept them well under the distortion threshold, and consequently reliable, both before and after the amplification; those components are now available for visualization and analyses.

Based on a potentially heavily-clipped recording, those analyses may complement, but not substitute, the analyses performed on the original, unclipped recording, and may help to discover unreported features such as those summarized below.

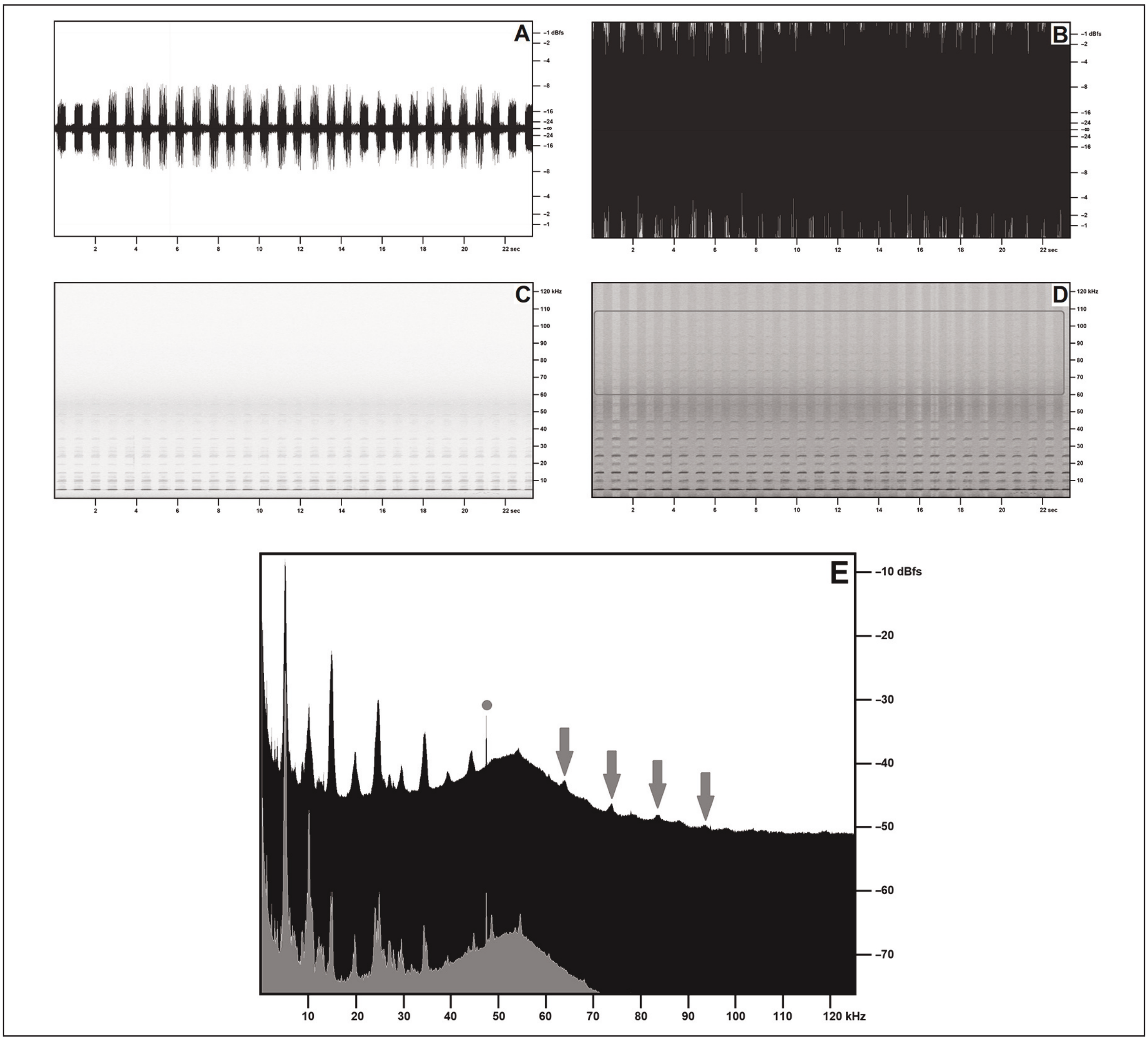

Figure 2. Song of *Natula averni* (Costa, 1855) – A, B: time/pressure envelope respectively before and after the +30dBfs amplification; C, D: time/frequency spectrogram respectively before and after the +30dBfs amplification. The bordered area in D shows the presence of the emerging harmonics not visible in C; E: comparison of frequency/pressure analyses before (light foreground) and after (dark background) the +30dBfs amplification. In E, a small circle marks a technogenic frequency spike unrelated with the song. The blunt hump approximately centered at 55 kHz is due to the noise of the A/D converter (see text). Four arrows mark as many frequency spikes representing the odd-numbered harmonic frequencies between the 13th and the 19th.

Considering that those emerging features include also non-bioacoustic components, robust consistency criteria are needed to select which features should be reliably attributed to the singing specimen. To avoid any ambiguity, it should be remembered that technogenic artefacts exist, that are modulated by the most intense parts of the signal: in that respect, those artefacts may be synchronous with the native features of the signal. Synchronicity alone cannot suffice, effective criteria to identify native components must take into account both the time and the frequency domains:

- in time domain analyses, consistency may be checked in the time/pressure envelope and in the time/frequency spectrogram - as it can be observed in Figs. 2–5 by comparing images A and B, as well as C and D. New elements emerging in full temporal coincidence with the echemes/syllables appearing in the non-amplified recording can be most parsimoniously explained as previously unobserved song

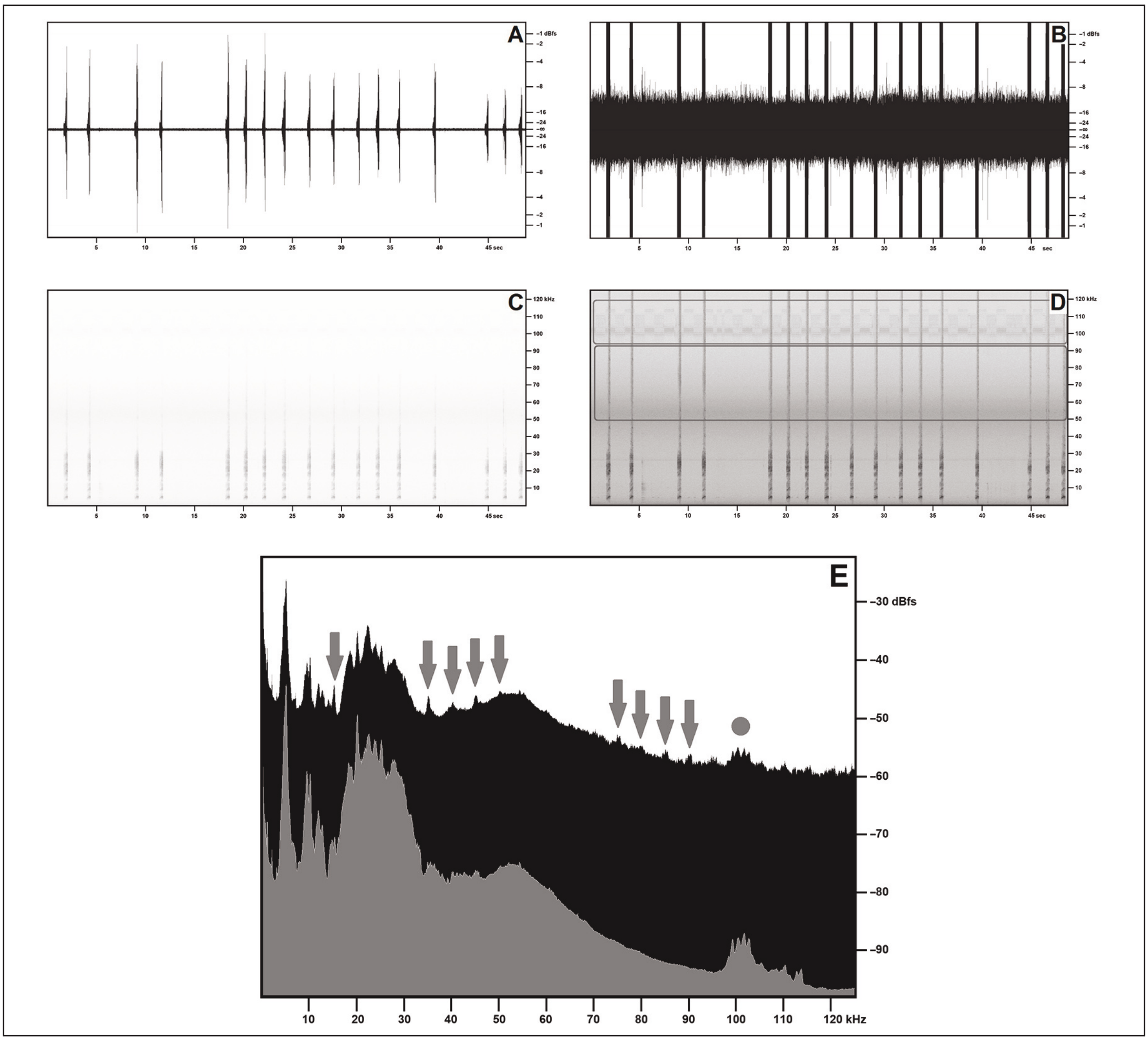

Figure 3. Song of *Melanogryllus desertus* (Pallas, 1771) – A, B: time/pressure envelope respectively before and after the +30dBfs amplification; C, D: time/frequency spectrogram respectively before and after the +30dBfs amplification. The lower bordered area in D shows the presence of the emerging harmonics not visible in C, the upper bordered area in D shows technogenic noise as a feeble checkered pattern, unrelated with the song; E: comparison of frequency/pressure analyses before (light foreground) and after (dark background) the +30dBfs amplification. In E, a small circle marks a multi-cusped technogenic frequency spike unrelated with the song. The blunt hump approximately centered at 55 kHz is due to the noise of the A/D converter (see text). Nine arrows mark as many frequency spikes evident after the amplification process, the rightmost of which represent the 14th to the 18th harmonic frequencies. With some doubt, a small spike at around 110 kHz may mark the 22th harmonic.

components only when, as in the case of high-Q songs, new regularly-spaced horizontal segments marking the harmonic frequencies can be made out in the spectrographic view (D);

- in frequency domain analyses, as exemplified by the analyses E in Figs. 2–5, consistency can be assumed when a typical song feature appearing in the leftmost part of the analysis reappears in its rightmost (higher frequencies) part. The previously unobserved high-frequency replica may consist of evenly-spaced spikes (high-Q) or of irregular humps and cusps that reproduce with a lower pressure excursion the typical pattern observed in the unamplified low-Q recording.

The over-amplified high-Q songs show previously unobserved, higher harmonic components that survive the consistency criteria illustrated above:

- in the case of *Natula averni* (Costa, 1855), illustrated in Fig. 2, the original recording showed a geometrical unitary harmonic progression of a narrow primary peak with a fundamental frequency of 4850 Hz. The third and following harmonics are two- or

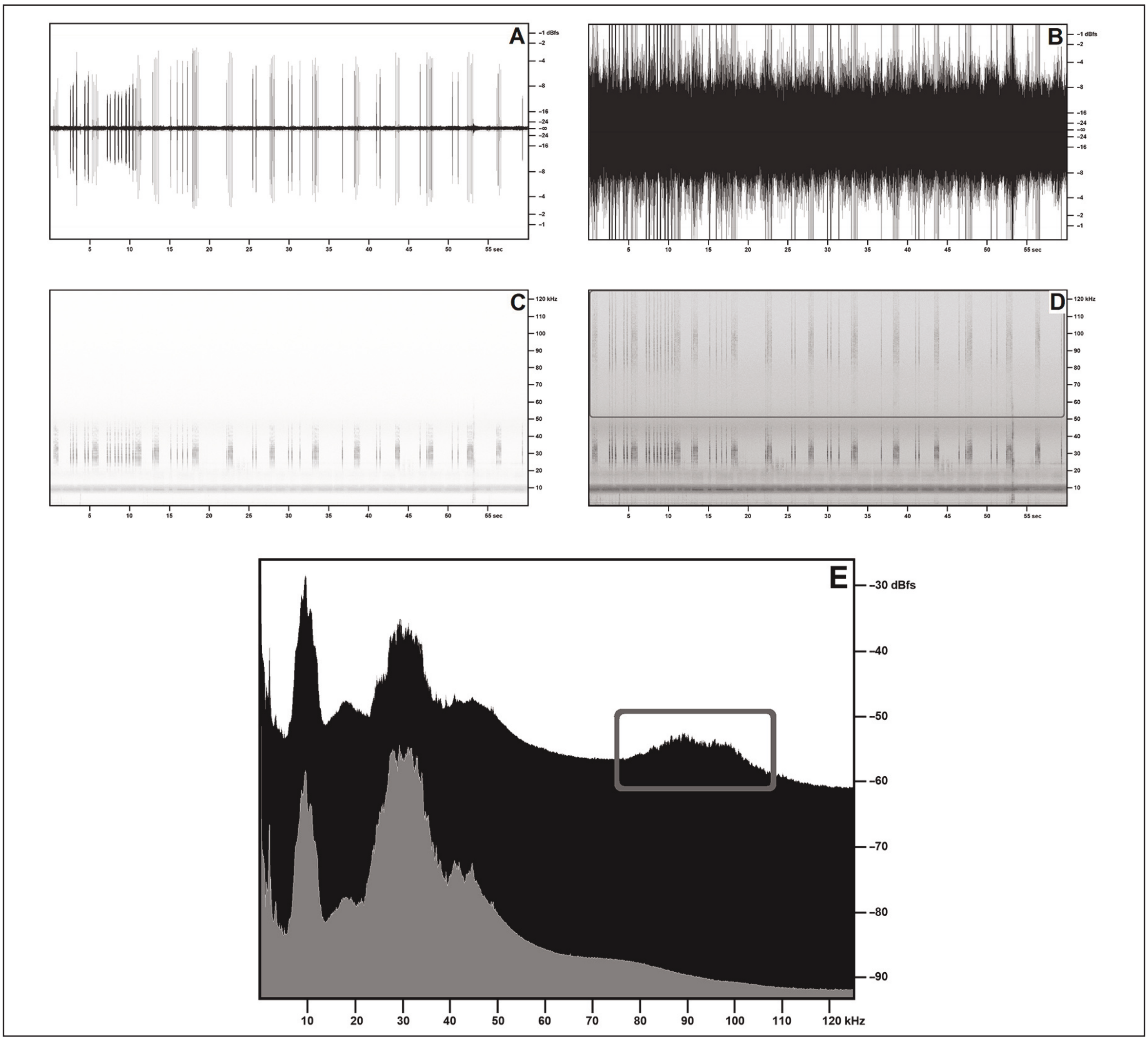

Figure 4. Song of *Leptophyes laticauda* (Frivaldszky, 1868) – A, B: time/pressure envelope respectively before and after the +30dBfs amplification; C, D: time/frequency spectrogram respectively before and after the +30dBfs amplification. In D, the bordered area encloses the emerging prosecution of song-related elements in the higher frequency range, not visible in C; E: comparison of frequency/pressure analyses before (light foreground) and after (dark background) the +30dBfs amplification. In E, the bordered area marks the most relevant emerging song elements also appearing in D.

three-cusped, and engage a slightly wider band. With reference to the center of a group of cusps, observed harmonics include (in round brackets, the arithmetical multiple of the observed fundamental) 9,918 Hz (9,800) – 14,800 Hz (14,550) – 19,620 Hz (19,400) – 24,350 Hz (24,250) – 29,440 Hz (29,100) – 34,170 Hz (33,950). Single cusps from at least three more harmonic frequencies are visualized, up to the eleventh instance at approximately 54.500Hz (53350). The hump from the inherent noise of the A/D converter is evident in this recording, as well as a technogenic frequency spike at 47,300 Hz. After over-amplification, the even-numbered harmonic look feebler than the odd-numbered harmonics, and new odd-numbered harmonics become evident at around 64 kHz (13th harmonic), 73.8 kHz (15th harmonic), 83.4 kHz (17th harmonic), 93.6 kHz (19th harmonic);

- in the case of *Melanogryllus desertus* (Pallas, 1771), whose song is harmonically more complex, the iterations of the multi-cusped fundamental cluster of frequencies is more difficult to follow. Two technogenic elements, the A/D converter and high

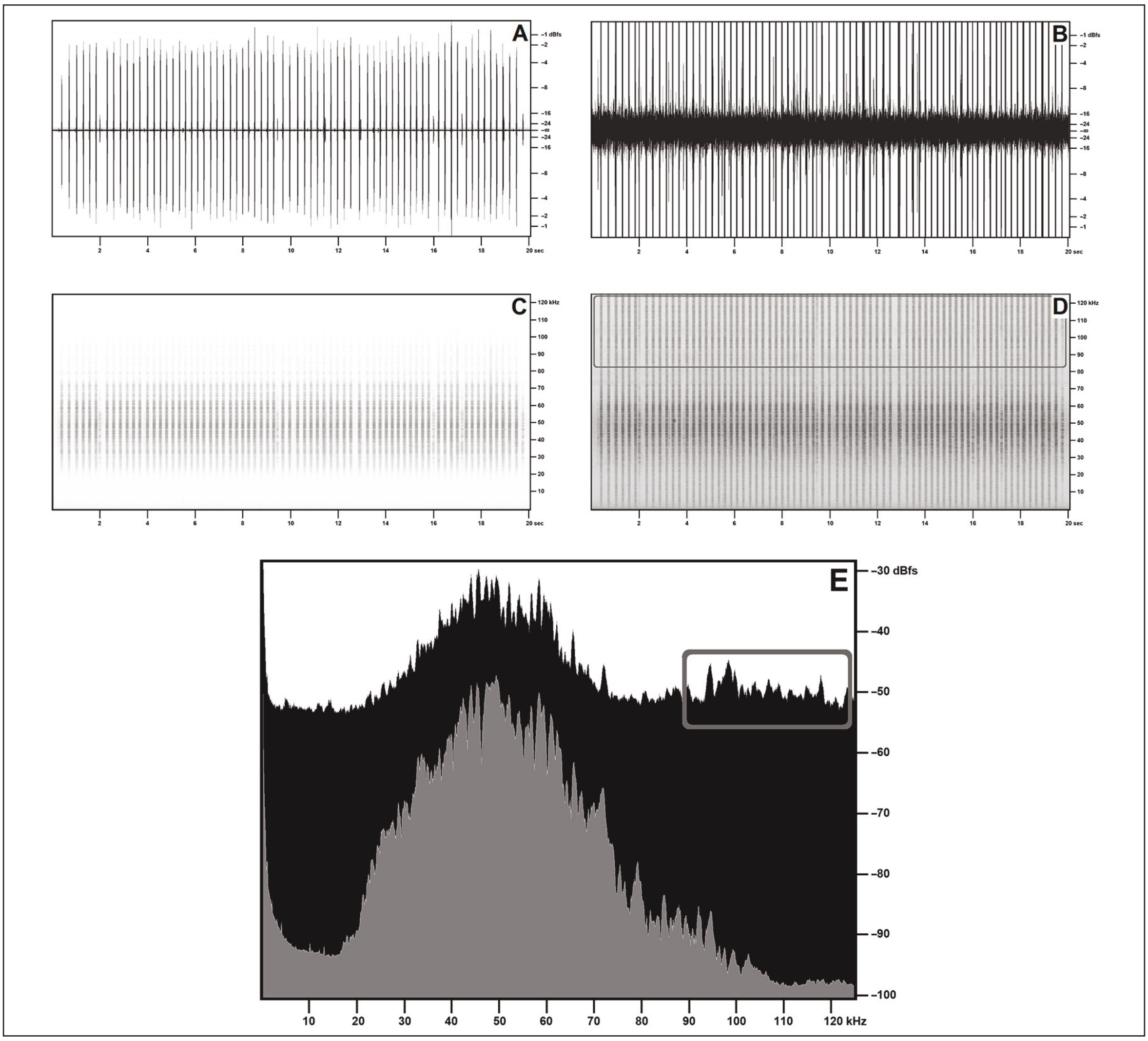

Figure 5. Song of *Ctenodecticus bolivari* Targioni-Tozzetti, 1881 – A, B: time/pressure envelope respectively before and after the +30dBfs amplification; C, D: time/frequency spectrogram respectively before and after the +30dBfs amplification. In D, the bordered area encloses the emerging prosecution of song-related elements in the higher frequency range, not visible in C; E: comparison of frequency/pressure analyses before (light foreground) and after (dark background) the +30dBfs amplification. In E, the bordered area marks the most relevant emerging song elements also appearing in D.

frequency noise from the nearby railroad, affect the original recording, that shows a well-defined fundamental spike at 4,974 Hz. As shown in Fig. 3, after over-amplification, multi-cusped features centered at around 75 kHz, 80 kHz, 85 kHz and 90 kHz become apparent, clearly corresponding with the 14^{th} to the 18^{th} harmonic frequencies. With some doubt, a small spike at around 110 kHz may mark the 22th harmonic

The extreme amplification revealed consistent, hidden features also in the low-Q songs. The wider and less defined bands appearing as prickly humps in the frequency/pressure analyses tend to reappear in one or more iterations above the fundamental band:

- in the case of *Leptophyes laticauda* (Frivaldszky, 1868), covered by Fig. 4, in the original recording the bands and sub-bands of song components progressively disappear above 50 kHz, while in the over-amplified recording a blunt, two-humped prickly iteration becomes apparent between 80 kHz and 110 kHz, particularly so in the time-frequency spectrogram. The hiatus between 50 kHz and 80 kHz may be explained by the noise of the A/D converter obliterating possible feeble song-related components, that disappear and are no more available for the over-amplification process;

- in the case of *Ctenodecticus bolivari* Targioni-Tozzetti, 1881 clear frequency bands absent from the original recording emerge above 100 kHz, as shown in Fig. 5.

DISCUSSION

Even though the extreme amplification is applied on a copy of the original recording, that should be archived in its original unaltered state for possible subsequent investigations, in line of principle, an objection should be raised about the destructive nature of an over-amplification process that implies the occurrence of irreversible clipping. Could we obtain the same results in real time from the original, unamplified recording? In other words, could we apply some form of real-time digital signal gain that – while leaving the original recording unaltered – provides the same visual results illustrated above?

In a general sense, the answer may look affirmative: it suffices to think to the “preview” mode provided by most digital audio management software before committing the selected digital transformation. A preview implies applying the required algorithm to a sample of the selected audio, and showing on screen or playing (usually as an audio loop) the resulting altered audio: that way, the user may appreciate the effect of the current parameters and fine-tune the settings as needed for the desired result. During the preview phase, preview data are stored in RAM or in a temporary file. The preview ceases to exist as soon as the transformation is applied, and the transformed file is saved, or the transformation is aborted, with loss of the now useless preview.

Similarly, without even applying any transformation to the data, some digital audio software including SeaPro (Pavan, 1998–2017) provide a “digital gain” slider that affects visualization and allow the manipulation of amplitude scales, as well as the remapping of the associated color scales – a set of parameters that determine the look of the time-frequency spectrogram. Even the adoption of higher- or lower-contrast color scale may affect decisively the way in which spectral data are displayed on screen. Screenshots similar to those in Figs. 2–5 may be saved anytime but, as long as numerical data in the file are not affected, any analysis performed on the original file will be unaltered, and the displayed image will be lost as soon at the visualization parameters are changed or the program is closed.

Finally, considering that most software that provides frequency /sound pressure analyses such as E in Figs. 2–5 allows zooming along the vertical (pressure) axis, one may conclude that the momentary application of digital gain or the assignment of a wider, more contrasting palette to the lower pressure range, combined with a zooming-in to the lower pressure range may be equivalent to actual, numerical digital amplification as described here.

In that case, those combined provisions applied to the original unaltered recording should show the emerging low-volume features with the same clarity and completeness that can be observed in B, D, E in Figs. 2–5.

Such a reasonable supposition was not verified in the case study reported here. In Figs. 3–5 one can easily appreciate how the frequency profile of the over-amplified recording in the background includes elements (peaks or cusp systems) that simply did not appear in the original (foreground) recording. A comparison between C and D in each figure confirms that also in the time-frequency spectrographic image the elements emerging in D are plainly missing in C regardless to the contrast applied to the image. From actual tests performed on the original recordings, the software cited in the Materials and Methods section proved incapable to derive from the original recordings illustrations similar to those obtained from their respective over-amplified versions: while all the just-in-time provisions cited above may improve the overall resolution of images derived from the original recording, and reveal some trace of the hidden feeble components that would emerge from the over-amplification process, two things deserve mention:

- the dynamic range of the RGB displays, with particular reference to the visualization of commonplace 24-bit RGB images, does not necessarily provide enough room to surrogate the over-amplification process described here by a simple remapping of a given pressure band to a wider, more diverse array of colors. Remembering that a substantial difference among two shades of gray or two colors is needed for the naked eye to perceive them as clearly distinct, the number of different colors or shades needed may be higher than those effectively available. Furthermore, whichever the number of bands, such an expedient would be effective only in case that the tiny differences between the feeble components and the background noise are great enough to engage different contiguous bands - which is undemonstrated;

- numerically speaking, the merit of the over-amplification process is exactly one, emphasizing signals, in such a way that, as an example, a gradient too small to engage two adjacent, different color bands in the color map of a given visualization, appears clearly after amplification because now it is big enough to engages an higher number of bands.

CONCLUSIONS

The ideal recording equipment, with particular reference to the microphone, is the one that induces the minimal level of noise in any recording, while ensuring the highest dynamic range. In actuality, any recording is affected by noise in many form, ambient and technogenic, intrinsic and extrinsic. Low-pressure components such as those associated with the highest frequencies, here exemplified as the “emerging components” after over-amplification, are comparable in intensity to the bottom noise but, as demonstrated above, may anyway get silently captured in a digital recording. They survive the competition with bottom noise, but at the price of invisibility. This study restores that lost visibility.

The better the equipment, the higher the relevance of low-pressure components and the frequency range from which useful information can be extracted. The good news that appears from this study is that - regardless to the equipment - any high-resolution recordings may provide more information than expected – it is just the matter of extracting it.

More specifically, the extreme (+30dB full digital scale) digital amplification of 16-bit high-resolution PCM recordings proved capable to extend the amount of information provided in the frequency domain by the original recording: any component of the original un-amplified recording too feeble to cross the visualization threshold in renderings such as time/frequency spectrograms and frequency/pressure analyses, after extreme amplification may become apparent in the same renderings, thus allowing to extend consequently the original song description.

The process of extreme amplification cannot always nor simply be surrogated just by altering or extending the amplitude/color maps used to display the spectrograms and frequency analyses: while the latter interventions may provide results vaguely similar to those of the process advocated here, there are structural, numerical reasons why simple “cosmetic”, just-in-time enhancement of images obtained from the original recording are less effective in revealing hidden components than their counterparts from over-amplified recordings.

Since it is based on already available audio material, this increase in informativeness can be obtained at a very low cost. It can be hypothesized that an extensive reconsideration of already published high-resolution 16-bit recordings based on the method proposed here could increase the exhaustiveness and the precision of bioacoustic reports such as song descriptions.

The increasing economical value of scientific research encourages an extensive, careful and informed usage of any kind of scientific record available, including exisiting and novel audio recordings. Science is better served by striving to extract the highest possible amount of information from any existing audio sample, including the 16-bit audio bit depth recordings making up the greatest majority of available audio reference samples. High-resolution 16-bit audio such as the 0–125 kHz recording obtained at a sampling frequency of 250 kHz will continue to be produced for many years to come, and by no means should be considered as an obsolescent nor deprecated technology. As long as 16-bit audio will be a scientifically viable medium, bioacousticians will face the inherent limitation of that audio format, including sensitivity to clipping distortion: at the same time, such a well-known limitation should not be overemphasized. Once a correctly-produced unclipped recording is available, there may be good reasons to explore the possible increase in usable information that may derive from an unconventional practice such as the voluntary over-amplification of a copy of the original recording.

ACKNOWLEDGEMENTS

I thank Gianni Pavan (Cibra - Centro Interdisciplinare di Bioacustica e Ricerche Ambientali, Dipartimento di Scienze della Terra e dell'Ambiente, Università di Pavia) for his insightful comments on an earlier version of the manuscript and for his useful suggestions.

REFERENCES

Bass H.E., Sutherland L.C., Piercy J. & Evans L., 1984. Absorption of sound by the atmosphere. In: Physical acoustics: Principles and methods, Vol. 17 (A85-28596 12–71). Orlando, FL, Academic Press, Inc., pp. 145–232.

Blackman R.B. & Tukey J.W., 1958. The measurement of power spectra from the point of view of communications engineering. Bell System Technical Journal, 37: 185–282. https://doi.org/10.1002/j.1538-7305.1958.tb03874.x

Boulanger R. & Lazzarini V. (Eds.), 2011. The Audio Programming Book. Cambridge, Massachusetts, The MIT Press, pp. 209–212.

Brizio C. & Buzzetti F.M., 2014. Ultrasound recordings of some Orthoptera from Sardinia (Italy). Biodiversity Journal, 5: 25–38.

Cesare Brizio, Filippo Maria Buzzetti & Gianni Pavan, 2020. Beyond the audible: wide band (0–125 kHz) field investigation on Italian Orthoptera (Insecta) songs. Biodiversity Journal, 11: 443–496. https://doi.org/10.31396/Biodiv.Jour.2020.11.2.443.496

Elsner N. & Popov A.V., 1978. Neuroethology of acoustic communication. Advances in Insect Physiology, 13: 229–355. https://doi.org/10.1016/S0065-2806(08)60267-2

Green E.I., 1955. The story of Q. American Scientist, 43: 584–594. https://www.jstor.org/stable/27826701?seq =1

Harris F.J, 1978. On the use of windows for harmonic analysis with the discrete Fourier transform. Proceedings of the IEEE, 66: 51–83. https://doi.org/10.1109/PROC.1978.10837

Matsushita Y., Jibiki T., Takahashi H. & Takamizawa T., 1989. A 32/24 bit digital audio signal processor. IEEE Transactions on Consumer Electronics, 35: 785–792. https://doi.org/10.1109/30.106896

Montealegre-Z. F. & Morris G.K., 1999. Songs and Systematics of Some Tettigoniidae from Colombia and Ecuador I. Pseudophyllinae (Orthoptera). Journal of Orthoptera Research, 8: 162–236.

National Instruments Corporation. White Paper 4278, 2018. The Fundamentals of FFT-Based Signal Analysis and Measurement in LabVIEW and LabWindows/CVI. Available online: http://www.ni.com/white-paper/4278/en/

Nuttall A.H., 1981. Some windows with very good sidelobe behavior. IEEE Transactions on Acoustics, Speech, and Signal Processing, 29: 84–91. https://doi.org/10.1109/TASSP.1981.1163506

Nyquist H., 1928. Certain topics in telegraph transmission theory. Transactions of the American Institute of Electrical Engineers, 47: 617–644. https://doi.org/10.1109/T-AIEE.1928.5055024

Pavan G., 1998–2017. SeaPro software. http://www-9.unipv.it/cibra/seapro.html

Shannon C.E., 1949. Communication in the presence of noise. Proceedings of the Institute of Radio Engineers, 37 : 10–21. https://doi.org/10.1109/jrproc.1949.232969

Smith J., 2007. Pulse Code Modulation (PCM) in Mathematics of the Discrete Fourier Transform (DFT) with Audio Applications, Second Edition, online book. https://ccrma.stanford.edu/~jos/mdft/mdft.html Accessed on 29 December 2022

Sound Devices 2022. 32-Bit Float Files Explained. https://www.sounddevices.com/32-bit-float-files-explained/ Accessed on 29 December 2022

Vladišauskas A. & Jakevičius L., 2004. Absorption of ultrasonic waves in air. Ultragarsas, 1: 46–49.

Welch T.B., Wright Cameron H.G. & Morrow M.G., 2012. Real-Time Digital Signal Processing from MATLAB to C with the TMS320C6x DSPs, 2nd ed.; CRC Press: Boca Raton, FL, USA.