JASA ARTICLE

Dependency of threshold and loudness on sound duration at low and infrasonic frequencies

Carlos Jurado,[1,a)] Marcelo Larrea,[1] Hiren Patel,[2] and Torsten Marquardt[2]
[1]*Escuela de Ingeniería en Sonido y Acústica, Universidad de Las Américas, Avenue Granados and Colimes, EC170125, Ecuador*
[2]*UCL Ear Institute, University College London, 332 Grays Inn Road, London, WC1X 8EE, United Kingdom*

ABSTRACT:
Many environmental sounds contain significant energy in the infrasonic and low-frequency (ISLF) ranges that have been associated with cases of annoyance and noise complaints. This study assessed the effect of sound duration on audibility and loudness of ISLF sounds. A first experiment evaluated detection thresholds for tones of 4, 16, and 32 Hz with durations up to 4000 ms. Furthermore, equal-loudness-level contours (ELCs) were obtained as function of duration up to 2000 ms. Tones of 1000 Hz were also included here. Results displayed the known pattern of general sound level decrease with increasing duration up to several hundred milliseconds. ELCs stabilized slightly earlier than thresholds, but after 1000 ms, levels remained roughly constant for both measures except for 4-Hz tones, where the decrease continued up to the longest durations tested. As 4-Hz cycles are perceptually resolved as separate pressure pulses, the authors hypothesized their duration dependence would resemble that of pulse trains. Hence, a second experiment evaluated pulse-train thresholds (1000-Hz carrier) for durations up to 4000 ms. For both pulse repetition rates of 4 and 32 Hz, threshold stabilized after 1000 ms as for tones $\geq$16 Hz, suggesting the continuing threshold decrease for a 4-Hz tone is specific to infrasound.
https://doi.org/10.1121/10.0001760

(Received 19 January 2020; revised 16 July 2020; accepted 25 July 2020; published online 26 August 2020)
[Editor: Joshua G. Bernstein] Pages: 1030–1038

I. INTRODUCTION

It has been long established that below a critical duration, shorter sounds require a higher intensity than longer ones to be detected, or to be of equal loudness (Garner, 1947; Small *et al.*, 1962; Watson and Gengel, 1969; Scharf, 1978; Gerken *et al.*, 1990). This phenomenon has been generally referred to as temporal integration. While the underlying physiological mechanisms are still not well understood, it has been acknowledged that both peripheral mechanisms and central auditory processes are involved (for reviews and discussions, see Moore, 2007; Verhey, 2010; Viemeister, 2014). Critical duration, the stimulus duration beyond which threshold and loudness do not significantly change anymore, has been reported to be generally less than 1000 ms (Scharf, 1978; Poulsen, 1981; Florentine *et al.*, 1996), and some studies have indicated an increase with decreasing frequency (Plomp and Bouman, 1959; Watson and Gengel, 1969; Gerken *et al.*, 1990). In this study, we investigated to what extent this trend extends towards very low frequencies, as to the knowledge of the authors, almost no temporal integration data exist for very low frequency sounds (including infrasonic frequencies, i.e., <20 Hz). An exception, although with very limited data, is the study of infrasound detection thresholds by Kühler *et al.* (2015), suggesting that threshold for a 10-Hz tone remains constant for durations >1000 ms. As to the review of infrasound hearing thresholds and equal-loudness-level contours (ELC) by Møller and Pedersen (2004), they selected primary studies that used stimulus durations 500 – 2000 ms (or longer) to avoid potential effects of temporal integration. Thus, the effect of duration on audibility and loudness for infrasonic and low-frequency (ISLF) sounds has remained uncertain. Considering the increase in complaints about environmental noise with spectral content below 100 Hz (Leventhall, 2004; Schmidt and Klokker, 2014; Baliatsas *et al.*, 2016), there is a need to understand temporal integration in this frequency range to better assess the impact of environmental ISLF sounds.

Two experiments were conducted for this study: experiment 1 assessed the auditory temporal integration for ISLF tones, both at threshold and suprathreshold levels and up to durations that well exceeded 1000 ms. In experiment 2, control threshold measurements were performed using 1-kHz pulse trains with repetition rates of 4 and 32 Hz, in an attempt to explain our findings for 4-Hz infrasound tones.

II. EXPERIMENT 1

A. Methods

1. Detection threshold as function of duration

Detection thresholds were obtained for tones with frequencies of 4, 16, and 32 Hz. A two-alternative-forced choice (AFC), 3-down 1-up adaptive procedure was used.

a)Electronic mail: carlos.jurado@udla.edu.ec, ORCID: 0000-0003-2799-8977.

Start levels were set 15 dB above previously published sensation thresholds (ISO-226, 2003, for frequencies >20 Hz; Møller and Pedersen, 2004, for infrasound). The step-size started at 3 dB, changed to 2 dB after one reversal, and decreased to 1 dB after another reversal. The procedure stopped after 11 reversals, and threshold was obtained by averaging the last eight. The 13 combinations of frequency and duration (see Sec. II A 3 for duration values) were measured in random order. Three further measurements for each of the 13 conditions were obtained in new random orders and threshold for each condition was defined as the median value of the four measurements.

2. Equal-loudness-level contours

Loudness matches were obtained for the same ISLF tones as used in the detection threshold measurements, using various durations (see below). Since published data on the duration dependence of ELCs are rare compared to those for thresholds, and the loudness-matching procedures differ, 1000-Hz tones were also included in the ELC measurements to allow comparison of a mid-frequency ELC with those obtained for the ISLFs in the same subjects.

For a given loudness match, the reference tone had the same frequency as the comparison tone but its duration was fixed at 500 ms and its level set to 40 phon. A fixed number of decibels, corresponding to the frequency specific dB-difference between threshold and 40-phon ELC (based on ISO-226, 2003, for frequencies >20 Hz, and Møller and Pedersen, 2004, for infrasound) were added to the individual's 500-ms duration threshold to reach this loudness level. As 1000-Hz thresholds were not obtained, the 1000-Hz reference tone was set to 40 dB sound pressure level (SPL) (equal to 40 phon according to ISO-226, 2003).

A 2-AFC task was used in combination with a maximum-likelihood tracking (MLT) procedure. These tracking procedures determine the adaptive value for the next stimulus presentation by an estimation of the listener's psychometric function (PMF), which is updated iteratively (see, e.g., Green, 1990). In our experiment, subjects had to indicate which of the two tones (reference or variable), presented in random sequence, was louder. Maximum comfortable levels, obtained beforehand per frequency for each subject, limited the maximum stimulus level during these adaptive tracks.

One frequency was measured at a time, with the presentations of all tracks for a given frequency (i.e., of the various durations) randomly interleaved within the measurement of the entire ELC for this frequency. Based on previous responses, the level of the variable tone was set randomly at either the 80% or 20% point of the most likely PMF (cumulative Gaussian). Presenting levels slightly below and above the currently estimated point of subjective equality (PSE) (i.e., the 50% point on our PMF) has been previously proposed to improve the estimate of its slope (e.g., Brand and Kollmeier, 2002; Shen and Richards, 2012). Takeshima *et al.* (2001) gave additional reasons to set the stimulus level not exactly at the PSE: (1) The wider range of presentation levels avoids the context effect that may occur when many stimuli are presented repeatedly within a narrow range. The occasional presentation of stimuli with a clear loudness difference "re-calibrates" the listener's judgment. (2) Subjects become tired of responding to stimuli that have converged to equal loudness. Presentation at the 20% and 80% values, corresponding to a 1:5 judgment-bias ratio, maintains a subtle loudness difference, which ensures that the subjects remain motivated right to the end of the track. (3) The wider range of presentation levels makes the procedure more robust against lapses of attention (Green, 1995; Shen and Richards, 2012). Our own experience with MLT confirm these advantages.

To avoid long runs at either the 20% or the 80% level, the randomness was constrained to quadruplets of presentations, which consisted of the four combinations of the two possible presentation orders and presentation levels. Each track terminated after 16 presentations and a final PMF was determined from it. If the slope of a final PMF was shallower than 15% per dB (indicating low response consistency) this measurement was repeated until four PMFs steeper than 15% per dB were obtained for each combination of frequency and duration. The majority of subjects had to repeat ~10% of their measurements; subject S2 (asterisks in Figs. 1 and 2) had to repeat 25%. From each of the four PMFs, the stimulus level giving the 50% value was determined. The median of the four measurements per condition gave the stimulus level that is estimated to have produced the same loudness as the 500-ms reference tone.

3. Stimuli duration

Stimuli had onset and offset cosine ramps, whose durations depended on the tone's frequency. For frequencies 4, 16, and 32 Hz, each ramp had a duration of one cycle, and for 1000 Hz five cycles. It should be noted that the stimulus duration definition has varied across the literature, including, e.g., the interval between stimulus onset and offset, the stimulus segments that are above 50% or 90% of full amplitude, and energy based definitions (see considerations in, e.g., Gerken *et al.*, 1990; Heil *et al.*, 2017). Because the one-cycle ramps of our ISLF tones constitute a considerable part of the overall stimulus duration, especially for the shortest stimuli, half of the onset and offset ramp durations were included in the nominal stimulus duration values. In other words, stated durations throughout this report can be interpreted as "half-height" durations.

Table I gives an overview of the nominal stimulus durations tested for the various tone frequencies. In total, there were 13 and 19 conditions for the threshold and loudness measurements, respectively.

4. Apparatus

Stimuli were generated in MATLAB (MathWorks, Inc., Natick, MA) and digital-to-analog converted at a 48-kHz sample rate using an RME Fireface UC (Audio AG,

TABLE I. Total stimulus durations and ramp durations (in brackets) used in experiment 1 (*1000 Hz was only measured for loudness; **this duration was only used for the 4-Hz threshold measurements). The bottom row shows the inter-stimulus intervals (ISI) used for the various stimulus durations in the two-interval presentations of the threshold measurements. For the ELC tests, the ISI was 1000 ms for all conditions.

	Nominal stimulus duration (ms)							
Frequency (Hz)	32	64	125	250	500	1000	2000	4000
1000*	36(5)	69(5)	130(5)	255(5)	505(5)	1005(5)	2005(5)	
32			157(32)	282(32)	532(32)	1032(32)	2032(32)	
16				313(63)	563(63)	1063(63)	2063(63)	
4					750(250)	1250(250)	2250(250)	4250(250)**
ISI (ms)	500	500	400	300	100	100	100	100

Haimhausen, Germany) audio interface, connected to the controlling PC. The signal was amplified (BEAK Type BAA 120, Frankenblick, Germany) and low-pass filtered before driving a custom-built sound source. The low-pass filter in combination with a voltage divider was applied to limit the maximum loudness level to be less than 105 phon (based on extrapolation of data in ISO-226, 2003 for frequencies >20 Hz, and Møller and Pedersen, 2004, for infrasound). The sound source consisted of a 15-in. subwoofer speaker (15P80Nd, Beyma, Valencia, Spain), hermetically sealed in a wooden box. It was acoustically connected to the subject via an 8-m long polyethylene tube (inner diameter: 14 mm), itself fitted to a narrower silicon tube (inner diameter: 2.5 mm), which connected to a foam earplug (ER3–14A, Etymotic Research, Elk Grove Village, IL), placed inside the subject's ear canal. This source is identical to that used by Kühler *et al.* (2015) and allows for the generation of pure ISLF tones down to 2.5 Hz at a 60-phon level (according to Møller and Pedersen, 2004), with higher harmonics below the hearing threshold (see Fig. 2 in Kühler *et al.*, 2015).

The 1000-Hz tones were produced by an Etymotic ER4 earphone that was driven directly by the audio interface. Its output was channeled via a 20-mm-long plastic tube (1 mm of inner diameter), pierced through the foam of the earplug into the ear canal. A miniature microphone (FG-23453, Knowles Electronics, LLC, Itasca, IL) was similarly connected with the ear canal for all *in situ* calibrations. Measurements took place in a triple-walled sound booth. The timing of the stimulus presentations was visualized on a computer screen inside the booth and responses collected using a numerical keypad.

5. Calibration and probe-checks

The frequency response of the miniature microphone was measured using a 1/2 in. B&K microphone (type 4193) with a low-frequency adaptor (UC0211) in a 1.3 cm^3 cavity under white noise excitation. With its response compensated, this microphone was then used to measure the transfer functions of the sound sources in the ear canal during the measurement sessions. The electrical signal to the sound sources was adjusted so that the desired stimulus level was met. More details of this method can be found in Marquardt *et al.* (2007). Such *in situ* calibration took place 2 min after each earplug placement (to let the foam expand) as well as approximately every 20 min during and also at the end of the measurement sessions. Measurements were repeated if the SPL had changed by more than 1 dB.

6. Data analysis

For comparison with literature results, mean threshold and loudness data were fitted with an equation based on the leaky integrator model by Plomp and Bouman (1959):

$$I/I_0 = 1/\left(1 - e^{-(t/\tau)^n}\right), \tag{1}$$

where I_0 is the intensity for large durations (assumed constant beyond the critical duration) and τ is the integration time constant. The slope for durations well below the critical duration (corner of a two-line fit to the data, e.g., see Small *et al.*, 1962), which is linear when plotted on a log −log scale, is controlled by the exponent n. This slope will be called *initial* slope. In its original form, n equals one, and the initial slope is 10 dB/decade, the theoretical value of an energy integrator as assumed in classic works (e.g., Zwislocki, 1960). The slope parameter n was introduced when shown that initial slopes in the data were usually less than 10 dB/decade (see review by Gerken *et al.*, 1990). Although more sophisticated and physiologically relevant models have been proposed (see, e.g., Poulsen, 1981; Glasberg and Moore, 2002; Verhey, 2010; Hots *et al.*, 2014; Heil *et al.*, 2017), this simple two-parameter equation was chosen as it facilitates comparison with previous studies, commonly just reporting a time constant and a slope value. As some studies report their slopes in dB/doubling (i.e., decrease in sound level with a factor of 2 increase in duration rather than 10), note that 1 dB/doubling is equivalent to ∼3.3 dB per decade. A Levenberg-Marquardt least-squares minimization algorithm was used to find the optimal values of n and τ (Marquardt, 1963).

7. Subjects

Five adult subjects, a female and four males, participated in experiment 1. They had normal hearing according to standard pure-tone audiometry (British Society of Audiology, 2011) and normal middle-ear function according to standard tympanometry (British Society of Audiology, 2013). The UCL research ethics committee approved the experiment.

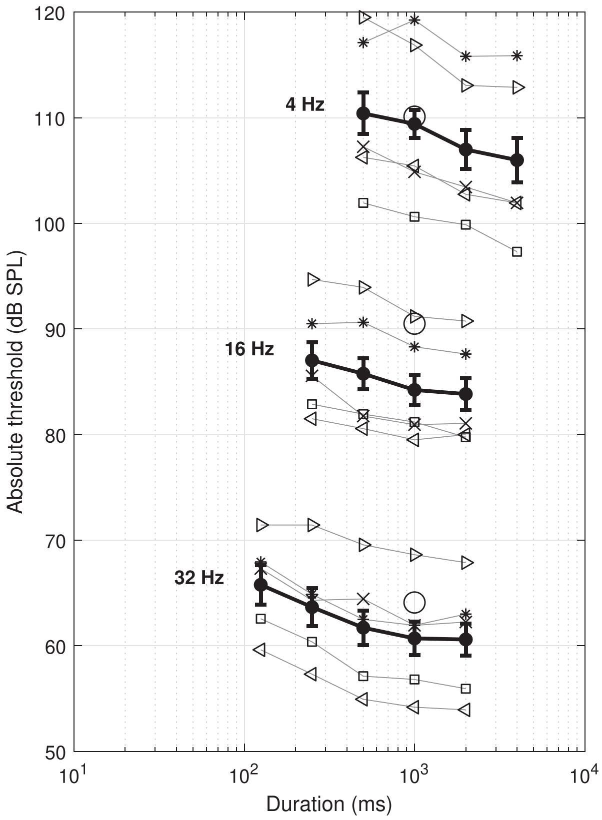

FIG. 1. Detection thresholds as function of stimulus duration. Individual (thin grey lines with individual markers) and average (bold lines) data are shown. The error-bars indicate the average of individual standard deviations across the repeated measurements. For comparison, 1000-ms data obtained by Watanabe and Møller (1990) are shown with unfilled circles (as these data were measured binaurally in a pressure chamber, 3 dB were added to account for binaural advantage as recommended by Møller and Pedersen, 2004).

B. Results

An overview of all detection-threshold data is shown in Fig. 1, with individual curves shown by thin lines and average curves in bold. Marked individual differences in overall levels were observed, with standard deviations of 5.6, 5.0, and 7.0 dB for 32, 16, and 4 Hz, respectively (obtained for the 2000-ms duration stimuli). These are similar to the ∼5 dB standard deviations for frequencies below 100 Hz reported by Møller and Pedersen (2004), excepting 4 Hz which was slightly larger in our study. The latter is, nevertheless, qualitatively in line with results by Kühler *et al.* (2015), who observed an increasing spread in thresholds for tones below 10 Hz. Note that the error bars in Fig. 1 do not represent the across-subject variability, but the average standard deviations of the within-subject measurements (and are therefore an indicator of reproducibility). The average detection thresholds for the 1000-ms tones obtained in our study present overall a reasonable agreement with 1000-ms data from Watanabe and Møller (1990) (circles in Fig. 1). In spite of this general frequency-dependence, the well-known trend of thresholds to decrease with increasing duration was clearly observed for all frequencies tested. A similar pattern of results was observed in the ELCs shown in Fig. 2, where individual data were vertically displaced for better visual clarity. Best seen in the average curves (bold lines in Figs. 1 and 2), both threshold and ELCs tended to level off beyond 1000 ms, except the 4-Hz curves, which seem to continue decreasing for even longer durations.

We tested the significance of this observation by comparing the level differences between the relevant data points (i.e., slopes). The longest duration tested for 16 and 32 Hz was 2000 ms and defined the upper duration for the slope comparison of both, threshold curves and ELCs. The lower duration had to be a point where the 16- and 32-Hz curves had already flattened. For thresholds this was the case after 1000 ms, while for the ELCs already after 500 ms (as confirmed also by time constant estimates; see Fig. 4). A one-way analysis of variance (ANOVA; with tone frequency as factor) indeed revealed a significant effect of frequency

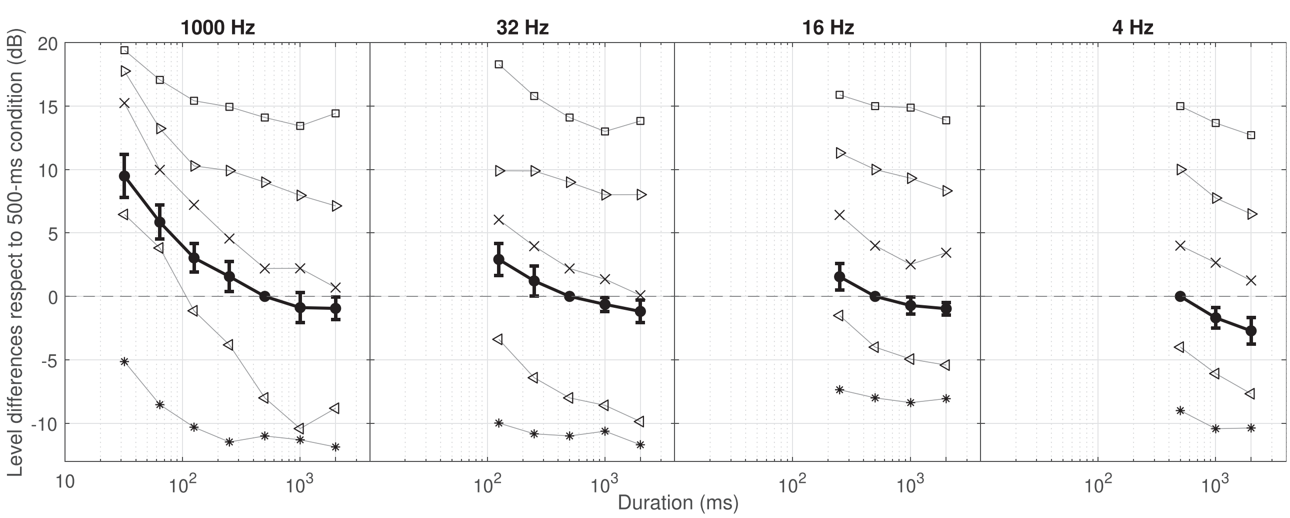

FIG. 2. Level differences of tones with various durations relative to the 40-phon reference tone with 500-ms duration (no error bars) that produced equal loudness. As in Fig. 1, the error-bars on the average curves (bold lines) indicate the average of individual standard deviations across the repeated measurements. Because information regarding the duration dependence lies in the shape of the curves only, individual curves (thin grey lines with same individual markers as those assigned in Fig. 1) were vertically offset by arbitrary amounts for visual clarity.

($F_{2,12} = 8.3$, $p < 0.01$) for the difference between the 1000- and 2000-ms duration thresholds. *Post hoc* comparisons (Hochberg and Tamhane, 1987) showed that 4 Hz presented significantly larger dB differences than both 32 and 16 Hz, while these did not differ between each other. For the ELCs, the one-way analysis of variance (ANOVA) on the individual dB differences also revealed a significant effect of tone frequency ($F_{3,16} = 5.5$, $p < 0.01$). *Post hoc* comparisons showed that the dB differences were significantly larger for 4 Hz than for all other frequencies, while these did not differ among each other. Notably, the dB differences observed for the mid-frequency tone of 1000 Hz did not differ from those of the 16- and 32-Hz low-frequency tones.

It is a possibility that the continuing decrease of the 4-Hz curves indicates a cycle-based time scale of the duration dependence at these low frequencies, rather than a scale based on absolute time. To see on which scale the curves better align, we plotted the data normalized to the 500-ms condition on the time scale [Figs. 3(A) and 3(B)] as well as normalized to the threshold/PSE for an eight-cycle duration, on a cycle-based duration scale [Figs. 3(C) and 3(D)]. We consider that plotting the curves close to each other, achieved by these normalizations, facilitates their shape comparison, which was the focus here. Although the slope of the 4-Hz curves does agree fairly well with the initial slopes of the 16- and 32-Hz curves when plotted on a cycle scale, our impression was that the 16-, 32-, and 1000-Hz curves [the latter available only for the ELCs, shown in panels (B) and (D)] line better up on the time scale, and the 4-Hz infrasound curves are the exception. We therefore continued the analysis on the basis that absolute time describes temporal integration rather consistently over the conventional frequency range of human hearing (i.e., $\geq$16 Hz), and the way that time is integrated in the infrasound range operates perhaps differently. This analysis involved the quantification of the initial slopes and time constants using Eq. (1). We preferred this relatively simple, but classic formula of temporal integration (based on Plomp and Bouman, 1959) because we considered that the limited duration range of our data does not allow the fitting of models with higher complexity (e.g., those of Poulsen, 1981 or Hots *et al.*, 2014). Furthermore, most of the previously published studies to which we can compare our results reported only these two parameters. The fitted curves, as well as derived parameters for our data set, are shown in Fig. 4. Quantitatively, there were relatively small deviations between fits and data; root-mean-squared deviations were in all cases <0.5 dB. Derived initial slopes for the 4- and 16-Hz threshold curves were slightly shallower than that of 32 Hz. For loudness, initial slopes were slightly shallower for 32 Hz than for 4, 16, and 1000 Hz. Thus, no systematic frequency dependence in the initial slopes was observed. Similarly, there was also no systematic difference in initial slope between the threshold and loudness data.

Time constants (vertical lines in Fig. 4) were generally shorter for loudness than threshold. For both measures, there was a slight tendency for τ to increase with decreasing frequency, which would be in line with other reports (Plomp and Bouman, 1959; Watson and Gengel, 1969; Gerken *et al.*, 1990). However, at least for tones $\geq$16 Hz, this trend was not extensive, and due to the limited number of data points available for the fits we are unable confirm such trend with certainty. As the average thresholds for 4 Hz did not stabilize even up to 4000 ms, the fit gave an unrealistic τ-value of 36 s, which makes application of the leaky-integrator model to the 4-Hz data highly questionable.

C. Discussion

While the temporal integration for the lower audio-frequency tones of 16 and 32 Hz was found to be qualitatively similar to that reported for the mid audio-frequency range, the 4-Hz infrasound tone appeared to follow a different rule.

Before discussing possible explanations for the 4-Hz findings, we first compare the data obtained for frequencies $\geq$16 Hz with previous studies, that to our knowledge have only examined systematically the temporal integration for frequencies >100 Hz. Note that the number of data points available to fit the temporal integration function was restricted by the minimum stimulus durations that could be tested, due to the long periodicity of low-frequency tones. The uncertainty of our estimate is therefore expected to increase as the stimulus frequency was lowered. Nevertheless, the initial slopes obtained for threshold as function of duration (7.6 and 10.0 dB/decade) are well within those reported in the literature (e.g., for frequencies $\leq$1000 Hz, Gerken *et al.*, 1990 report values from several studies that range from $\sim$3 to 12 dB/decade). For loudness, derived initial slopes (7.4–10.9 dB/decade) were broadly similar to those of thresholds, as has been observed previously (Scharf, 1978). For the 1000-Hz ELC, where shorter durations could be measured, our slope in the range 32–125 ms was similar to the $\sim$10 dB/decade slope obtained by Poulsen (1981) for this frequency in a similar duration range.

The range of the time constant, τ, obtained for the 32- and 16-Hz thresholds (362–522 ms) is also roughly in line with literature values reported for frequencies below 1000 Hz. For example, the review by Gerken *et al.* (1990) reports an average of 588 ms across different studies. We also observed systematically shorter τ values for loudness than for threshold, as has been previously observed (Scharf, 1978; Poulsen, 1981; Moore, 2007). The slight general tendency of τ to increase with decreasing frequency, as observed here for both threshold and loudness, has also been reported previously (Plomp and Bouman, 1959; Watson and Gengel, 1969; Gerken *et al.*, 1990). This tendency might explain why our τ values of the loudness-curve fits for 16 and 32 Hz (306 and 321 ms, respectively) were on the upper end of those reported previously, whereas our τ value of 252 ms for 1000 Hz falls in the middle (e.g., 110–320 ms for 1000 Hz at various levels by Poulsen, 1981).

Now, we come to the peculiar findings with 4-Hz infrasound tones. Infrasonic sinusoids do not elicit a pitch sensation (which appears to cease below about 30 Hz, according to Krumbholz *et al.*, 2000) and below approximately 10 Hz

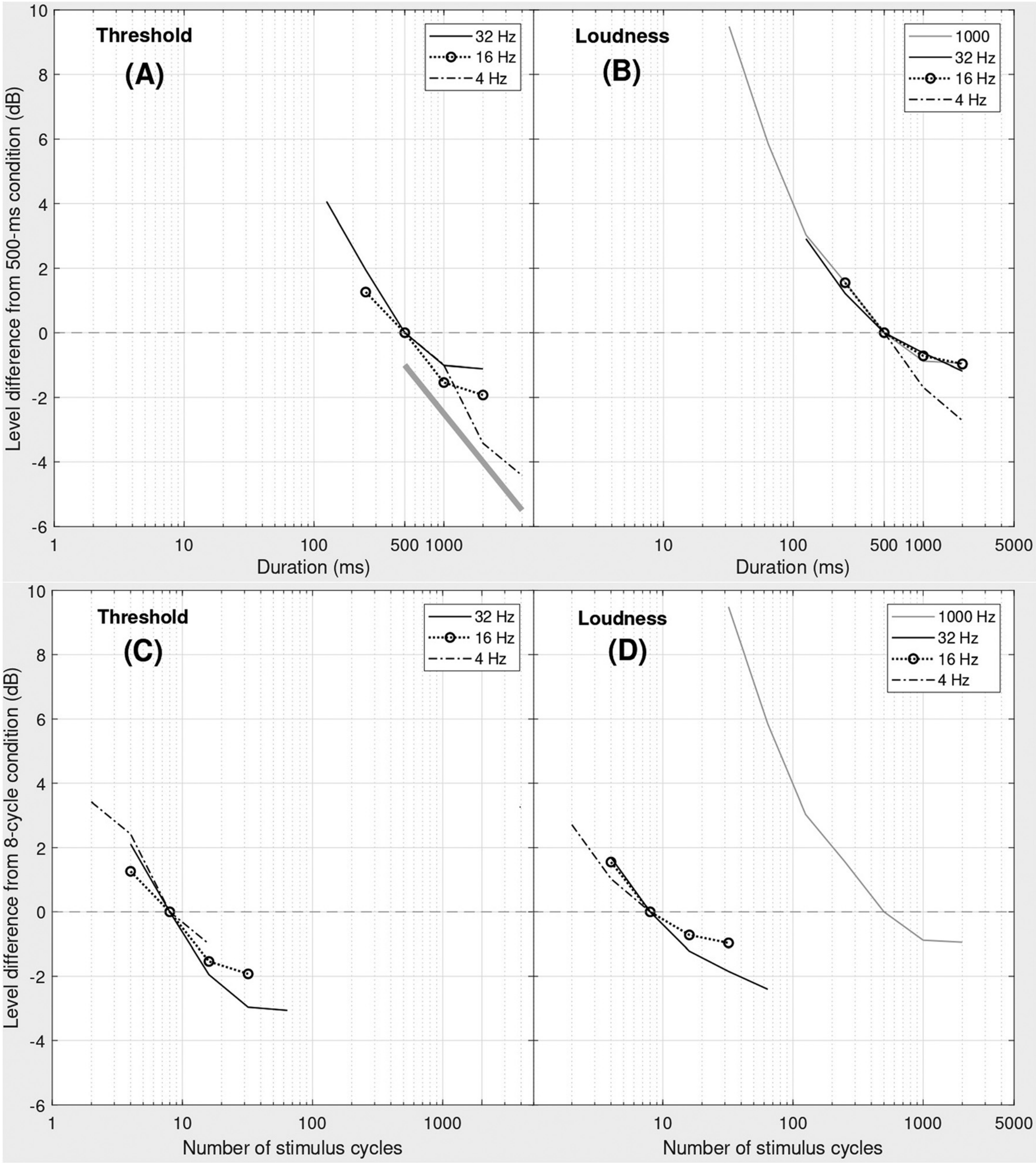

FIG. 3. Shape comparison between the across-subject average curves of Figs. 1 and 2. Data in panels (A) and (B) were normalized to the 500-ms values. The thick-grey line in panel A has a slope of −1.5 dB per doubling in the number of pulses (equivalent to 5 dB/decade), according to the detection probability theory for multiple events (see discussion). In panels (C) and (D), data are shown on a cycle scale and were normalized by the levels at the eight-cycle duration [with 1000 Hz in panel (D) normalized to the 500-ms values as no eight-cycle data were measured here].

are not perceived as continuous, i.e., their single cycles can be perceptually resolved as separate pressure pulses (Møller and Pedersen, 2004). These are important qualities that set infrasound apart from sound in the audio-frequency range. With this in mind, we postulated that threshold dependence on duration for tones below 10 Hz could be likened to that for a train of pressure pulses. It has been shown that detection thresholds for trains of clicks or tone pulses steadily decrease with the number pulses by roughly 2 dB per doubling in number (6.6 dB/decade, see, e.g., Carlyon *et al.*, 1990; Heil *et al.*, 2017). The observed decrease has been well explained by probabilistic auditory models that consider how multiple-event information can be optimally used for signal detection. For example, the "multiple looks" interpretation by Viemeister and Wakefield (1991) predicts a 1.5 dB threshold decrease per doubling in the number of observations, assuming independent looks and optimal information use (see also Green and Swets, 1988; Verhey, 2010). Indeed, our 4-Hz threshold average data decreased approximately with this slope [shown as thick grey line in Fig. 3(A)] and share the characteristic of tone-pulse threshold data by Heil *et al.* (2017). Their Fig. 7B shows how thresholds, obtained for 8-Hz trains of 3125-Hz tone bursts, steadily decrease at a rate of roughly –2 dB per doubling in

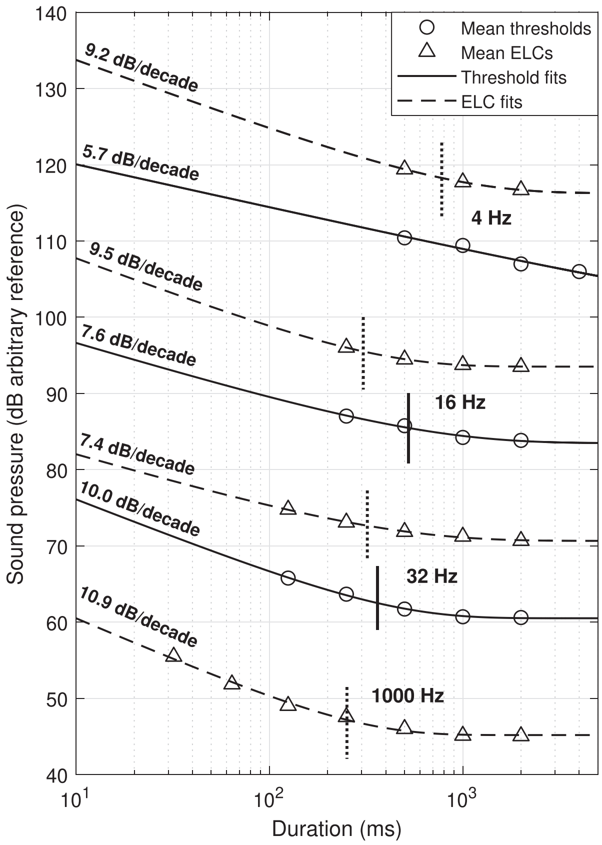

FIG. 4. Average threshold (circles) and loudness data (triangles) as function of stimulus duration, fitted with Eq. (1) describing a leaky-integrator (threshold: solid lines; loudness: dashed lines; fit function is described in Sec. II A 6). Threshold data are given in dB SPL, while loudness data have been vertically shifted to be close above the respective threshold function (with 1000 Hz positioned arbitrarily). Derived values for the initially linear slopes are given above each fit. Note that as frequency decreases, slope estimates should be regarded as increasingly vague approximations, due to restriction in the minimum stimulus durations that could be tested. The integration time constant (τ) of each fit is shown by a short vertical line (not shown for the 4-Hz threshold fit, as τ was here >5000 ms).

train duration. However, we were unable to find similar studies with threshold curves exceeding a 1000-ms duration, like our data. Thus, we decided to conduct a second experiment with trains of 1-kHz tone bursts that had durations up to 4000 ms. We focused here only on threshold measurements because the mentioned probabilistic models can only account for detectability. The continuing decrease of the 4-Hz ELC can potentially be explained by the (informal) reports from our listeners, noting that the trains of single pressure pulses were becoming increasingly annoying with duration. Loudness and annoyance have indeed been reported to be highly related, especially for infrasound (Andresen and Møller, 1984).

III. EXPERIMENT 2

A. Methods

Detection thresholds for pulse trains with repetition rates (f_r) of 4 and 32 Hz were measured as function of the total pulse-train duration. Each pulse consisted of a 10-ms Hanning-windowed tone, with a frequency of 1000 Hz. Threshold was determined with a 2-AFC procedure similar to experiment 1. Two sessions were run on different days, each of them consisting of two threshold measurements for each combination of repetition rate and duration. If these two measurements differed by more than 3 dB, an additional measurement was obtained still within the same session (across all subjects and both sessions, this led to 13% of additional measurements). For a given condition, no subject required an additional measurement in more than one session. Threshold was defined as the median of all four to five measurements.

Detection thresholds were measured for pulse-train durations of 250, 500, 1000, 2000, and 4000 ms; for $f_r = 32$ Hz a duration of 125 ms was added. No overall ramps as in experiment 1 were used (e.g., the pulse train with $f_r = 4$ Hz and 1000-ms duration contained four full-amplitude tone pulses) because here, the windowing of each of the tone pulses reduced the spectral spatter.

The stimuli were produced by one of the miniature speakers of an Etymotic ER10C system that was driven directly by an RME Fireface 802 audio interface. Its microphone was used for the *in situ* calibrations.

Experiment 2 was run at, and under the ethical approval of Universidad de Las Américas, in a double-walled sound booth; six subjects participated, including one that participated in both experiments. All had normal hearing as defined in experiment 1. Responses were collected via a pushbutton with LEDs synchronized with the stimulus intervals.

B. Results

Figure 5 shows the dependency of pulse-train detection thresholds on stimulus duration obtained for $f_r = 4$ and 32 Hz together with the respective thresholds for pure tones obtained in experiment 1 (dashed grey lines) for comparison. These data show that the 4-Hz pulse-train thresholds leveled off beyond the 500-ms duration. Although thresholds kept decreasing slightly beyond 500 ms with approximately –1.5 dB/decade, the dB differences between 500 and 4000 ms were significantly smaller ($F_{1,9} = 11.9$, $p < 0.01$) than those observed for the 4-Hz infrasound-tone threshold curves (that had a more than three times larger slope).

To further compare with results from experiment 1, Eq. (1) was also fitted to the pulse-train data. The initial slopes of mean thresholds below 500 ms were –7.7 and –10.2 dB/decade for $f_r = 4$ and 32 Hz, respectively (initial slopes for the corresponding pure tones were −5.7 dB/decade and −10.0 dB/decade, respectively). After 1000 ms, the 32-Hz pulse-train thresholds showed a slight increase, which was not significant ($F_{1,10} = 0.23$, $p = 0.64$). Nonetheless, the similarity in the initial slope between the 32-Hz pure-tone and pulse-train threshold functions was remarkable, as was the similarity in their time constant values (τ: 362 and 308 ms, respectively). While the initial slopes of the 4-Hz data differed to some extent (note, however, that the

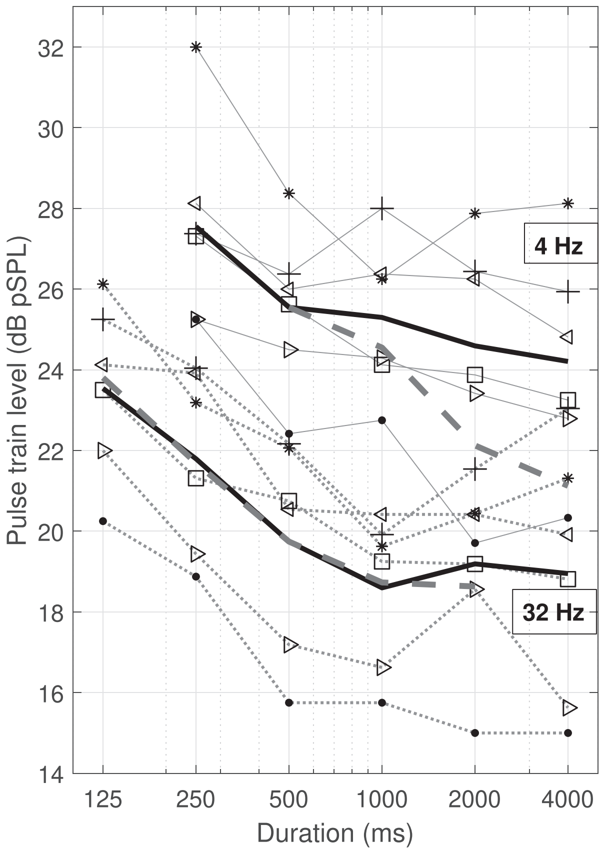

FIG. 5. Individual (thin lines with individual markers) and mean (bold solid lines) pulse-train thresholds as function of stimulus duration, obtained for the pulse repetition rates of 4 and 32 Hz. For comparison, mean thresholds for the 4- and 32-Hz ISLF tones are also shown (bold dashed lines; vertically shifted so that they match the respective average pulse-train thresholds at 500 ms). Note that markers do not refer to the same individuals as in Figs. 1 and 2, excepting one subject (squares) that participated in both experiments.

minimum duration was shorter for the pulse-train stimulus), the main difference was that the 4-Hz pulse train data levelled off with a time constant of 408 ms that was comparable to those found for the 32-Hz pulse train and tones ≥16 Hz.

IV. DISCUSSION

We set out to measure the dependency of threshold and loudness on duration for ISLF tones, to establish whether the temporal integration of these long-periodicity stimuli works on larger time scales than in the already well-studied frequency range above 100 Hz. A main outcome from this experiment was that temporal integration for tones ≥16 Hz does not substantially differ from that reported for higher frequencies (see Sec. II C). However, the 4-Hz tone threshold and ELC revealed significantly longer temporal integration for infrasound. In an attempt to explain this as a potential effect of probabilistic detection of multiple resolved pressure pulses (i.e., 4-Hz cycles), we measured thresholds for pulse trains with repetition rates of 4 and 32 Hz in a subsequent experiment.

The average threshold difference between the 4- and 32-Hz pulse trains of just 5.1 dB was not in line with the 9-dB increase in stimulus power, but was instead in agreement with the theoretical 4.5 dB difference (3 × 1.5 dB/doubling) predicted for the 2^3 times larger number of events in the 32-Hz pulse train compared to the 4-Hz pulse train. This result was also consistent with the rate-dependence of thresholds for constant-duration pulse trains reported by Heil *et al.* (2017; see rate-dependence for the 794.56-ms duration in their Fig. 7E), who used 3125-Hz tone pulses. However, the thresholds for neither pulse rate continued to decrease beyond 1000 ms with the theoretical slope of –5 dB/decade (–1.5 dB/doubling), unlike those of the 4-Hz infrasound tones. We conclude therefore that the multiple-event model is unlikely to account for the continuing threshold decrease observed with this infrasound stimulus.

At present time, reasons behind the continuing decrease in the 4-Hz tone threshold and ELC, as well as up to which duration they would continue to decrease, can only be speculated upon. One might argue that infrasound is detected by the vestibular rather than the auditory system. But for three reasons, we believe it is unlikely that the vestibular system is involved in airborne infrasound detection: (1) In our long experience with infrasound psychophysical measurements, no subject has reported dizziness or imbalance, neither as a threshold detection cue nor in response to higher levels used for the ELC. (2) Jurado and Marquardt (2019) showed that 4-Hz tones even at ∼ 90-phon loudness levels do not evoke any significant (saccular) vestibular-evoked myogenic potentials (VEMPs). (3) The fact that pink noise (250 – 4000 Hz) is able to mask 5- and 12-Hz tones (Burke *et al.*, 2019), is a clear indication that infrasound is detected by the cochlea.

A further reason could be the differing perceptual quality of the 4-Hz tone. All subjects agreed that the non-tonal character of infrasound made it harder to detect. Cues were easily masked, or confused by periods of internal physiological noise. Thus, a simple explanation could be that a period of "physiological quietness," during which infrasound can be detected, is more likely the longer the presentation interval is. On the other hand, many subjects also reported that, when the adaptive procedure reached threshold levels, they became often only aware of the 4-Hz "pressure pulses" during the late part of the presentation (sometimes even by the cessation of the stimulus), while the 16- and 32-Hz tones could be usually recognized already at the onset by their "humming" sound. Thus, it is possible that some detection cues might build up over time, and this might similarly underlie the observed decrease in the 4-Hz ELC. Salt *et al.* (2013) measured in guinea pigs electrical potentials evoked by (airborne) tones between 5 and 1000 Hz, and observed especially strong responses to 5-Hz infrasound tones. These potentials might underlie the perception of infrasound. Their data show, however, no build-up over time (their Fig. 1).

V. CONCLUSION

Down to 16 Hz, the pattern of threshold and loudness dependence on tone duration was broadly consistent with that reported previously for frequencies in the lower-audio frequency range (∼100–1000 Hz), with critical durations being much less than 1000 ms. At 4 Hz, however, ELCs and thresholds continued to decrease for durations up to at least 2000 ms. It might therefore be necessary to choose even longer stimulus durations when determining the loudness and detectability of infrasound. But this will require the development and testing of methods using only one presentation interval, as otherwise the durations of such experiments would become excessive, especially if also frequencies between 4 and 16 Hz are to be tested.

Although the auditory system temporally resolves the 4-Hz cycles so that probabilistic detection models for multiple events could potentially explain their detection, thresholds for resolved trains of short 1-kHz tone pulses with a rate of 4 Hz were not found to decrease for durations beyond 1000 ms.

Our finding might explain the differences amongst reports on detection-threshold curves and ELCs as a function of frequency that extend into the infrasound range, particularly in the degree of flattening of the threshold curves and ELCs below 15 Hz (see Møller and Pedersen, 2004, for a review). Inconsistencies may here be partly produced by differences in stimulus durations.

ACKNOWLEDGMENTS

The authors thank Associate Editor Joshua G. Bernstein and two anonymous reviewers for very helpful comments and suggestions on an earlier version of this manuscript. C.J. would also like to thank Christopher Guamán and Juan Vizuete for their help in the data collection process.

Andresen, J., and Møller, H. (**1984**). "Equal annoyance contours for infrasonic frequencies," J. Low Freq. Noise Vib. **3**, 1–9.
Baliatsas, C., van Kamp, I., van Poll, R., and Yzermans, J. (**2016**). "Health effects from low-frequency noise and infrasound in the general population: Is it time to listen? A systematic review of observational studies," Sci. Total Environ. **557–558**, 163–169.
Brand, T., and Kollmeier, B. (**2002**). "Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests," J. Acoust. Soc. Am. **111**, 2801–2810.
British Society of Audiology (**2011**). *Pure-Tone Air-Conduction and Bone-Conduction Threshold Audiometry with and without Masking: Recommended Procedure* (Reading, UK).
British Society of Audiology (**2013**). "Recommended Procedure. Tympanometry," http://www.thebsa.org.uk/wp-content/uploads/2014/04/BSA_RP_Tymp_Final_21Aug13_Final.pdf (Last viewed December 10, 2019).
Burke, E., Hensel, J., Fedtke, T., Uppenkamp, S., and Koch, C. (**2019**). "Detection thresholds for combined infrasound and audio-frequency stimuli," Acta Acust. Acust. **105**, 1173–1182.
Carlyon, R. P., Buus, S., and Florentine, M. (**1990**). "Temporal integration of trains of tone pulses by normal and by cochlearly impaired listeners," J. Acoust. Soc. Am. **87**, 260–268.
Florentine, M., Buus, S., and Poulsen, T. (**1996**). "Temporal integration of loudness as a function of level," J. Acoust. Soc. Am. **99**, 1633–1644.
Garner, W. R. (**1947**). "The effect of frequency spectrum on temporal integration of energy in the ear," J. Acoust. Soc. Am. **19**, 808–815.
Gerken, G. M., Bhat, V. K. H., and Hutchison-Clutter, M. (**1990**). "Auditory temporal integration and the power function model," J. Acoust. Soc. Am. **88**, 767–778.
Glasberg, B. R., and Moore, B. C. J. (**2002**). "A model of loudness applicable to time-varying sounds," J. Audio Eng. Soc. **50**, 331–342.
Green, D. M. (**1990**). "Stimulus selection in adaptive psychophysical procedures," J. Acoust. Soc. Am. **87**, 2662–2674.
Green, D. M. (**1995**). "Maximum-likelihood procedures and the inattentive observer," J. Acoust. Soc. Am. **97**, 3749–3760.
Green, D. M., and Swets, J. A. (**1988**). *Signal Detection Theory and Psychophysics* (Peninsula Publishing, Los Altos, CA), p. 241.
Heil, P., Matysiak, A., and Neubauer, H. (**2017**). "A probabilistic Poisson-based model accounts for an extensive set of absolute auditory threshold measurements," Hear. Res. **353**, 135–161.
Hochberg, Y., and Tamhane, A. C. (**1987**). *Multiple Comparison Procedures* (Wiley, Hoboken, NJ), 450 pp.
Hots, J., Rennies, J., and Verhey, J. L. (**2014**). "Modeling temporal integration of loudness," Acta Acust. Acust. **100**, 184–187.
ISO-226 (**2003**). "Acoustics—normal equal-loudness contours" (International Organization for Standardization, Geneva, Switzerland).
Jurado, C., and Marquardt, T. (**2019**). "On the effectiveness of airborne infrasound in eliciting vestibular-evoked myogenic responses," J. Low Freq. Noise, Vib. Active Control **39**, 3–16.
Krumbholz, K., Patterson, R. D., and Pressnitzer, D. (**2000**). "The lower limit of pitch as determined by rate discrimination," J. Acoust. Soc. Am. **108**, 1170–1180.
Kühler, R., Fedtke, T., and Hensel, J. (**2015**). "Infrasonic and low-frequency insert earphone hearing threshold," J. Acoust. Soc. Am. **137**, EL347–EL353.
Leventhall, G. (**2004**). "Low frequency noise and annoyance," Noise Heal. **6**, 59–72.
Marquardt, D. (**1963**). "An algorithm for least- squares estimation of nonlinear parameters," SIAM J. Appl. Math. **11**, 431–441.
Marquardt, T., Hensel, J., Mrowinski, D., and Scholz, G. (**2007**). "Low-frequency characteristics of human and guinea pig cochleae," J. Acoust. Soc. Am. **121**, 3628–3638.
Møller, H., and Pedersen, C. S. (**2004**). "Hearing at low and infrasonic frequencies," Noise Heal. **6**, 37–57.
Moore, B. C. J. (**2007**). "Temporal resolution and temporal integration," in *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues*, edited by B. C. J. Moore (Wiley, Sussex), pp. 117–141.
Plomp, R., and Bouman, M. A. (**1959**). "Relation between hearing threshold and duration for tone pulses," J. Acoust. Soc. Am. **31**, 749–758.
Poulsen, T. (**1981**). "Loudness of tone pulses in a free field," J. Acoust. Soc. Am. **69**, 1786–1790.
Salt, A. N., Lichtenhan, J. T., Gill, R. M., and Hartsock, J. J. (**2013**). "Large endolymphatic potentials from low-frequency and infrasonic tones in the guinea pig," J. Acoust. Soc. Am. **133**, 1561–1571.
Scharf, B. (**1978**). "Loudness," in *Handbook of Perception: IV. Hearing*, edited by E. C. Catrerette and M. P. Friedman (Academic, New York), pp. 187–242.
Schmidt, J. H., and Klokker, M. (**2014**). "Health effects related to wind turbine noise exposure: A systematic review," PLoS One **9**, e114183.
Shen, Y., and Richards, V. M. (**2012**). "A maximum-likelihood procedure for estimating psychometric functions: Thresholds, slopes, and lapses of attention," J. Acoust. Soc. Am. **132**, 957–967.
Small, A. M., Brandt, J. F., and Cox, P. G. (**1962**). "Loudness as a function of signal duration," J. Acoust. Soc. Am. **34**, 513–514.
Takeshima, H., Suzuki, Y., Fujii, H., Kumagai, M., Ashihara, K., Fujimori, T., and Sone, T. (**2001**). "Equal-loudness contours measured by the randomized maximum likelihood sequential procedure," Acta Acust. Acust. **87**, 389–399.
Verhey, J. L. (**2010**). "Temporal resolution and temporal integration," in *Oxford Handbook of Auditory Science: Hearing*, edited by C. J. Plack and D. R. Moore (Oxford University Press, Oxford), pp. 105–121.
Viemeister, N. F. (**2014**). "Temporal processing: Observations on the psychophysics and modeling of temporal integration and temporal resolution," in *Perspectives on Auditory Research*, edited by A. N. Popper and R. R. Fay (Springer, New York), pp. 601–610.
Viemeister, N. F., and Wakefield, G. H. (**1991**). "Temporal integration and multiple looks," J. Acoust. Soc. Am. **90**, 858–865.
Watanabe, T., and Møller, H. (**1990**). "Low frequency hearing thresholds in pressure field and in free field," J. Low Freq. Noise Vib. **9**, 106–115.
Watson, C. S., and Gengel, R. W. (**1969**). "Signal duration and signal frequency in relation to auditory sensitivity," J. Acoust. Soc. Am. **46**, 989–997.
Zwislocki, J. (**1960**). "Theory of temporal auditory summation," J. Acoust. Soc. Am. **32**, 1046–1060.