 Vol. 46, No. 3, pp. 459–469 (2021), doi: 10.24425/aoa.2021.138138

Research Paper

The Lower Limit of Pitch Perception for Pure Tones and Low-Frequency Complex Sounds

Carlos JURADO[(1)*], Marcelo LARREA[(1)], Brian C.J. MOORE[(2)]

[(1)] *Escuela de Ingeniería en Sonido y Acústica, Universidad de Las Américas*
Avenue Granados and Colimes, EC170125, Ecuador
[*]Corresponding Author e-mail: carlos.jurado@udla.edu.ec

[(2)] *Cambridge Hearing Group, Department of Experimental Psychology, University of Cambridge*
Downing Street, Cambridge CB2 3EB, England

(*received January 5, 2021; accepted May 9, 2021*)

The lower limit of pitch (LLP) perception was explored for pure tones, sinusoidally amplitude-modulated (SAM) tones with a carrier frequency of 125 Hz, and trains of 125-Hz tone pips, using an adaptive procedure to estimate the lowest repetition rate for which a tonal/humming quality was heard. The LLP was similar for the three stimulus types, averaging 19 Hz. There were marked individual differences, which were correlated to some extent across stimulus types. The pure-tone stimuli contained a single resolved harmonic, whereas the SAM tones and tone-pip trains contained only unresolved components, whose frequencies did not necessarily form a harmonic series. The similarity of the LLP across stimulus types suggests that the LLP is determined by the repetition period of the sound for pure tones, and the envelope repetition period for complex stimuli. The results are consistent with the idea that the LLP is determined by a periodicity analysis in the auditory system, and that the longest time interval between waveform or envelope peaks for which this analysis can be performed is approximately 53 ms.

Keywords: pitch; lower limit; periodicity analysis.

1. Introduction

Vibrating objects often produce periodic complex sounds whose spectra are composed of a fundamental frequency (F0) and several higher-order harmonics. The perceived pitch of a complex sound usually corresponds to its F0, even when the component at F0 is physically absent (SEEBECK, 1841). Pitch can be perceived for a wide range of F0s, although for high F0s the pitch percept is dominated by the fundamental component, rather than the higher harmonics (PLOMP, 1967). The lowest F0 for which a pitch can be heard has been explored using both direct judgments of pitch (RITSMA, 1962; GUTTMAN, PRUZANSKY, 1962) and discrimination tasks, such as F0 discrimination (KRUMBHOLZ *et al.*, 2000) and melody discrimination (PRESSNITZER *et al.*, 2001). The lower limit for pitch (LLP) has been found to correspond to F0s of 19 to 30 Hz. Most previous studies have used periodic complex tones as stimuli, with spectral components that seldom cover the frequency range below 200 Hz. Here, the LLP was explored using an adaptive task estimating the lowest repetition rate for which a tonal/humming quality was heard for three types of stimuli: pure tones, sinusoidally amplitude-modulated (SAM) tones with a carrier frequency of 125 Hz, and trains of 125-Hz tone pips.

Pitch perception plays a key role in understanding the environment (in speaker recognition: ATAL, 1972; SHEN, SOUZA, 2017; in auditory source segregation: DE CHEVEIGNÉ, 1997; OXENHAM, 2008), and perhaps unsurprisingly a pitch corresponding to F0 also occurs for non-human species (see e.g. WALKER *et al.*, 2009; HOESCHELE, 2017). The rich literature on human pitch perception has revealed the complexities of the phenomenon, which is influenced by factors such as the frequency range of the harmonics (RITSMA, 1962; KRUMBHOLZ *et al.*, 2000), the relative amplitude and resolvability of the harmonics (e.g. Shackleton, Carlyon, 1994; Moore, Gockel, 2011), the relative phases of the harmonics (e.g. Patterson, 1987; Pressnitzer *et al.*, 2001), the sound's temporal envelope and temporal-fine structure (e.g. Moore, 2008, 2019; Santurette, Dau, 2011) and the peak factor of the waveform (e.g. Pressnitzer *et al.*, 2001; Jackson, Moore, 2013). These factors are consistent with the idea that the auditory system performs a spectro-temporal analysis to determine pitch (Meddis, O'Mard, 1997; Shamma, Dutta, 2019).

The present paper is concerned with the LLP, the lowest repetition rate of sound that can convey a sensation of pitch. Complex sounds with very low repetition rates, such as 5 Hz, are perceived as strongly fluctuating, with a quality sometimes called fluctuation strength (Fastl, 1983) or flutter (Krumbholz *et al.*, 2000), and do not have a clear pitch (Warren, Bashford, 1981). If the repetition rate is gradually increased, the perceptibility of the individual fluctuations tends to decrease, the percept becomes more "smooth", and eventually a tonal quality, a pitch, is perceived. Ritsma (1962) explored the "existence region" for pitch using SAM tones with a wide range of F0s and carrier frequencies. He showed that the lowest F0 that evoked a pitch depended on the carrier frequency; higher-frequency carriers required a higher F0. Hereafter in this paper, the term LLP is used to refer to the lowest F0 at which a pitch can be heard for a given frequency region of the harmonics. The term overall LLP (OLLP) is used to refer to the lowest F0 at which a pitch can be heard when the frequency region of the harmonics is chosen to give the lowest LLP across the range of hearing.

Ritsma (1962) estimated an OLLP of about 50 Hz, based on the results for the lowest carrier frequency tested, which was 300 Hz. Guttman and Pruzansky (1962) estimated a lower OLLP of 19 Hz, using short periodic tone pulses whose spectral energy was reported to be audible for frequencies down to 130 Hz. In agreement with the latter authors, Warren and Bashford (1981) estimated the OLLP to be 20 Hz, using repeating segments of broadband noise. Pressnitzer *et al.* (2001) evaluated the LLP with a melody-discrimination task, using harmonic complexes filtered with different high-pass cutoff frequencies. For their lowest high-pass cutoff frequency of 200 Hz they observed an OLLP of 30 Hz when the harmonic phases were chosen to give a single clear waveform peak per period. They also used an alternating-phase condition for which the envelope repetition rate was 2F0. For this, the mean data showed an average OLLP corresponding to an F0 of about 24 Hz, with one of their three subjects showing an OLLP < 20 Hz (see Fig. 2 *ibidem*).

Other studies have evaluated pitch perception through F0-discrimination thresholds (F0DLs; Cullen, Long, 1986; Shackleton, Carlyon, 1994; Krumbholz *et al.*, 2000; Mehta, Oxenham, 2020). These studies have shown that when the F0 is decreased below the LLP, the F0DL increases sharply, although plateaus in the F0DL are reached for F0s well above and well below the LLP. The LLP inferred from the transition region decreases with decreasing center frequency of the stimuli (Krumbholz *et al.*, 2000). The F0DLs of Krumbholz *et al.* (2000) suggested an OLLP of about 30 Hz for a harmonic complex with a lower cut-off frequency of 200 Hz. Mehta and Oxenham (2020) determined F0DLs for harmonic complexes with a wide F0 range (30–2000 Hz). They randomly varied the rank (N) of the lowest harmonic present, to discourage F0 discrimination based on the frequency change of the lowest harmonic. For the lowest F0 used, the F0DL was rather large (about 8%), and decreased (improved) slightly with increasing N, in contrast to the results for higher F0s, for which the F0DLs increased (worsened) with increasing N. The large F0DLs and different effect of N for the F0 of 30 Hz may have been a consequence of all of the harmonics being unresolved (Moore, Gockel, 2011), even for the lowest value of N, which was 4.5 on average. However, it seems clear that a pitch was heard for F0 = 30 Hz, indicating that the OLLP is at or below 30 Hz.

Jackson and Moore (2013) determined F0DLs for a group of harmonics (group B) embedded in a group of fixed non-overlapping harmonics (group A) with the same mean F0. They used F0s of 50, 100, and 200 Hz, and the rank (N) of the lowest harmonic in group B was varied from 1 to 15. When all of the harmonics had the same level, F0DLs for F0 = 50 Hz increased as N was increased from 5 to 7, in agreement with most of the data described above. However, in one of their conditions, the level of each component was set to give a loudness level of 40 phon (see Fig. 5 in (Jackson, Moore, 2013)). In this condition, performance was best when $N = 1$, which was probably also their only condition with a resolved harmonic. Performance worsened as N was increased up to about 5 and then reached a plateau. However, performance was much worse when the harmonics had equal loudness levels than when they had equal levels. Jackson and Moore suggested that increasing the relative levels of the lowest harmonics (to make their loudness the same as for the higher harmonics) had the effect of partially masking potentially useful information carried by the higher harmonics. Overall, these results suggest that the sensation level and/or loudness of the lowest harmonics plays an important role in F0 discrimination for low F0s. Due to the steep frequency dependence of sound transmission through the middle ear to the cochlea for frequencies below 100 Hz (Marquardt *et al.*, 2007; Jurado, Marquardt, 2016), the relative and absolute levels of the individual components may also play an important role in determining the OLLP.

In the present study, the OLLP was estimated using pure tones with frequencies from 3 to 40 Hz. To the knowledge of the authors, pure tones with frequencies well down into the infrasound range have not previously been used to estimate the LLP. The use of pure tones was considered advantageous as it avoids possible effects of the relative levels of the components in complex tones (as observed by Jackson and Moore, 2013). Also, a pure tone may be regarded as an extreme example of a sound that contains a resolved harmonic. To evaluate whether the LLP is similar for pure tones and for complex low-frequency stimuli, SAM tones (with modulation frequency f_m) and tone-pip trains (with repetition rate f_{rep}), both with a 125-Hz carrier, were also used as stimuli. These were chosen to have a carrier frequency (f_c) below what has been typically used in previous studies determining the OLLP, which might lead to lower values of the OLLP.

2. Method

2.1. Stimuli

To minimize potential effects of temporal integration when using pure tones whose frequencies (f) could fall well in the infrasound range (see Jurado *et al.*, 2020), while not making the experiment excessively long, all stimuli had an overall duration of 3000 ms. Figure 1 shows example waveforms of the three stimulus types and their spectra. The value of f, f_m or f_{rep} was varied over the range 3–40 Hz, based on pilot tests indicating that the OLLP was within this range. All stimuli were presented monaurally to the subject's right ear.

2.1.1. Pure tones

The levels of the pure tones were set to give a loudness level of 65 phon, based on ISO-226 (2003) for $f \geq 20$ Hz and on Møller and Pedersen (2004) for $f < 20$ Hz (both interpolated). This loudness level was chosen as it ensured audibility for all subjects, taking into account the relatively large individual differences in sensitivity that have been observed in the infrasound range (Kühler *et al.*, 2015; Jurado *et al.*, 2020). The pure tones had 0.2-s cosine-squared ramps at the start and end and a 2.6-s steady portion. The value of f was varied according to the procedure described in Subsec. 2.3. To maintain the desired loudness level, the sound pressure level (SPL) was updated each time f was changed.

2.1.2. SAM tones

The SAM tones had a modulation depth of 100% and f_c = 125 Hz. No onset/offset ramps were applied apart from those inherent in the SAM. The value of f_m was varied in the procedure. The peak level of the SAM tones was set to 82.2 dB pSPL, which corresponds to the peak level of a 125-Hz pure tone at a 65-phon loudness level according to ISO-226 (2003; interpolated).

2.1.3. Tone-pip trains

Each tone pip was a 24-ms Hanning-windowed sinusoid. No onset/offset ramps were applied apart from those inherent in the Hanning window. The value of f_c was 125 Hz. The value of f_{rep} was varied in the

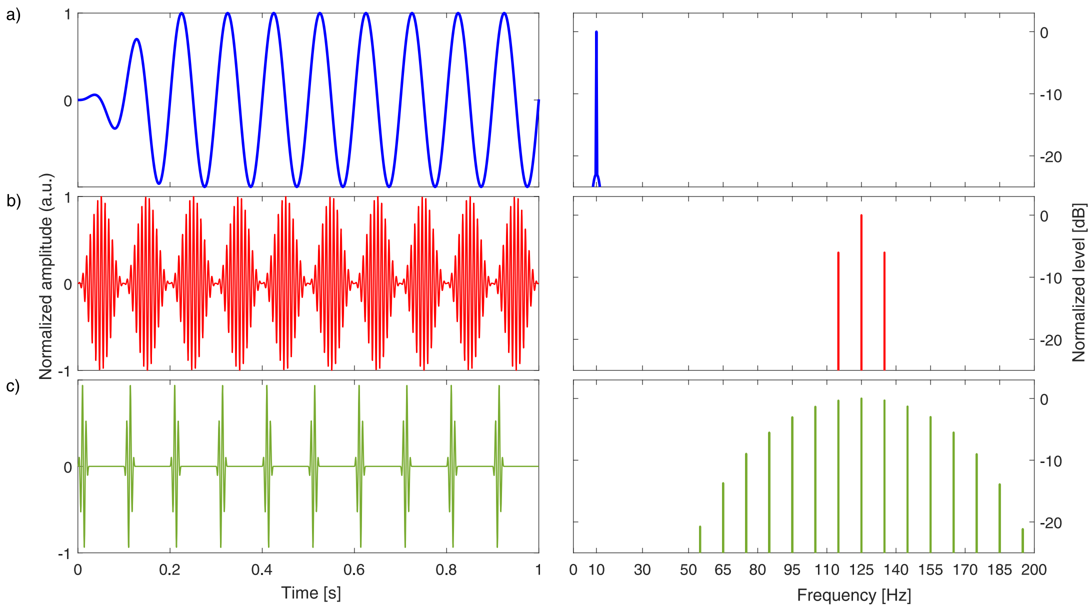

Fig. 1. Examples of the first 1-s of the waveforms of the stimuli (left) and their spectra (right). The stimulus types were: a) pure tones; b) SAM tones with a 125-Hz carrier; c) 125-Hz tone-pip trains. In the right panel, the spectral component with the highest level was set to 0 dB for ease in visualization. In this example, all stimuli had a repetition rate of 10 Hz.

procedure. The peak level of each tone pip was 82.2 dB pSPL, the same as for the SAM tones.

It should be noted that, unlike the complex sounds used in most previous studies of the LLP, the SAM tones and tone-pip trains did not usually have spectra corresponding to a simple harmonic series, since the values of f_m and f_{rep} were not constrained to be integer sub-multiples of f_c. It was anticipated that the lack of harmonicity would not have a material effect, since the perception of pitch for complex sounds with very low-frequency components appears to be determined mainly by the envelope repetition rate rather than by the harmonic structure or the temporal fine structure (Krumbholz *et al.*, 2000; Pressnitzer *et al.*, 2001; Moore *et al.*, 2009).

2.2. Apparatus

2.2.1. Set up

Digital signals were generated in MATLAB (The MathWorks, Inc., Natick, MA), with a sampling rate of 48 kHz and a precision of 24 bits, and were converted from digital to analog form using a Fireface-802 soundcard (RME Audio, Haimhausen, Germany). One of the soundcard's line outputs was connected to the miniature receiver of an ER10C measurement system (Etymotic Research, Elk Grove Village, IL), which was used to generate the pip-train and SAM stimuli. To play back the pure tones, the headphone output of the soundcard was used to drive a DA270-8 10-inch aluminum-cone subwoofer (Dayton Audio, OH). To enhance its low-frequency response, the loudspeaker cone was enclosed tightly by an acrylic cover, so that air-volume changes produced by the loudspeaker displacement could be efficiently transformed into pressure variations (Kinsler *et al.*, 1999, Chapter 5). The acrylic cover had a small hole where a silicon tube was inserted (0.8 m in length, ~0.7 mm inner diameter). The other end of the tube terminated in the subject's right ear canal via a tightly fitted pierced earplug (Etymotic ER10C-14A, B or C depending on ear-canal size). The same earplug, with its miniature tubes for the ER10C microphone and receivers, was used for in-situ calibration of both sound sources (described in Subsec. 2.2.3). Subjects were tested in an audiometric cabin, within a double-walled sound-isolated room of the Acoustics Laboratory at Universidad de Las Américas.

2.2.2. Harmonic distortion

The levels of the 2nd, 3rd and 4th harmonics produced by distortion were measured for the sinusoidal stimuli with $f \leq 40$ Hz, and with the level set to give a loudness level of 65 phon, as used in the experiment (harmonics above the 4th had much lower relative levels). These measurements were performed with a 2.5-Hz resolution. For $f \geq 7.5$ Hz, harmonic levels were below the hearing threshold (at 7.5 Hz the levels of the 2nd, 3rd and 4th harmonics reached 55.8, 61.8 and 39.1 dB SPL, respectively, which are below the detection threshold according to Møller and Pedersen, 2004, for frequencies < 20 Hz and ISO-226, 2003, for frequencies above that). However, below 7.5 Hz, the relative levels of the harmonics increased so that if presented at 65 phon, the harmonics (alone) would have been audible. For $f \leq 7.5$ Hz, the tones were presented at a level corresponding to a loudness level of 40 phon. This made it possible to maintain the harmonic levels below the detection threshold. The difference in loudness level for frequencies below and above 7.5 Hz meant that there was an abrupt change in level when the frequency was varied from below to above 7.5 Hz, or *vice versa*. However, frequencies below 7.5 Hz were seldom reached during the adaptive procedure (Subsec. 2.3), since they were well below the OLLPs.

2.2.3. Sound calibration

Calibration consisted of two parts: (a) a reference equipment calibration and (b) *in situ* calibrations. In (a), a GRAS (Holte, Denmark) 46AZ 0.5-inch microphone set (flat response down to 1 Hz), connected to an input of the RME soundcard, was placed at one end of a 1.3 cm^3 cavity. The probe of the ER10C system with an earplug was tightly fitted at the other end. Using white noise (20-s long) and frequency domain deconvolution, the transfer function of the ER10C microphone and that of one of the ER10C receivers was obtained. The white noise signal was divided into 50 segments of 0.4-second duration (2.5-Hz resolution), which were averaged to improve the signal-to-noise ratio (SNR) of the recording. After compensating for its response, the ER10C microphone was used to measure the transfer function of the subwoofer sound source and ER10C receiver. To define the responses on an absolute SPL scale, a reference sound signal (1000 Hz at 94 dB SPL) from a sound calibrator (CESVA CB006) was recorded with the GRAS microphone.

In (b), a similar procedure was used to measure the transfer functions of both sound sources with the ER10C microphone, but now with the probe placed in the subject's ear canal. To avoid deleterious effects of physiological noise (e.g. from breathing strongly or moving abruptly) on the in-situ measurements, a time-domain artifact rejection routine was implemented. Here, segments of the recording that had a power more than 3 median-scaled absolute deviations from the median power, were discarded (Burke, 1998). The remaining segments were weighted averaged to improve the SNR (Hoke *et al.*, 1984). If more than 20% of the segments had to be discarded, the measurement was repeated. Based on the difference between the complex spectra of the *in-situ* and reference calibrations, frequency-dependent multiplying factors were applied to the electrical signals so that the desired SPLs at specified frequencies were reached in the ear canal. Levels were confirmed with a CESVA SC310 sound-level meter, with its 1/2 inch microphone placed in the 1.3 cm^3 cavity. These were within 1.2 dB of the expected levels. More detail of these calibration procedures can be found in Marquardt *et al.* (2007) and Jurado *et al.* (2017).

The *in-situ* calibrations were repeated after about every second run of the procedure, to check that the probe was correctly fitted.

2.3. Procedure

All subjects received 2 hours of training (at most 3 days prior to the main experiment). Before commencing, each subject confirmed that they could hear a 5-Hz tone with a loudness level of 40 phon and a 40-Hz tone with a loudness level of 65 phon, presented in sequence, each 4-s long. As these stimuli also served as examples of "extreme cases", this initial presentation was also done when testing with the 125-Hz SAM and 125-Hz pip-train stimuli. The three stimulus types were presented in random order in a single session, lasting up to about 2.5 hours in total, including regular breaks to maintain focus.

An adaptive 1-up 1-down procedure was used. The subject was instructed to focus on the tonal or "humming-like" quality of the sound, and to press the left of two buttons if the sinusoid was perceived to have a tonal/hum-like quality (which decreased the frequency or repetition rate for the next presentation), or to press the right button otherwise (which increased the frequency for the next presentation). The procedure started with the highest frequency of 40 Hz, based on pilot tests showing that this frequency was well above the OLLP for all stimuli. The subject was instructed to scan the "transition region", thereby crossing the OLLP in both directions. This was done by first reducing the frequency until the tonal/humming quality was definitely lost and then increasing the frequency until a distinct tone-like quality was again heard. Later, the subject was asked to continue scanning the transition region but within a finer frequency range, in this manner slightly crossing their OLLP in both directions.

The initial step size of 6 Hz was decreased to 3 Hz after two reversals and further decreased to 1 Hz after two more reversals, where it remained. The procedure stopped after 12 reversals were obtained. The OLLP was defined as the average frequency at the last 6 reversals. Two runs were obtained for each stimulus type, and if these differed by more than an amount Δ [Hz], a third run was obtained. The OLLP frequency for that condition was defined as the average of the two or three estimates. The value of Δ was based on the absolute differences between the 1st- and 2nd-run OLLP estimates averaged across all conditions and subjects up to that point in the experiment (at the start, data from pilot tests from three participants were used). The distribution of these differences was used to determine outlier values, indicating insufficient reproducibility based only on two runs. The value of Δ was taken as the third quartile plus 1.5 times the interquartile range (Tukey, 1977). This estimate was updated as the experiment progressed. By the end of the experiment, Δ was 3.3 Hz.

If within the task subjects were uncertain as to what to listen for, they could press a middle button to listen to the 40-Hz stimulus (of the corresponding stimulus type), that served as an example of a "tonal/humming" sound with a pitch. Following this, the procedure continued as usual. This feature was included after pilot tests, where it was found that it helped to refresh the subject's memory as to what sound quality was sought, especially after having listened to several stimuli around the OLLP. Participants were instructed to use this feature only when really necessary, to avoid turning the procedure into a comparison with the 40-Hz stimulus. The percentage of times per run (i.e. the number of middle-button presses divided by the number of trials for a run) for which this feature was used was recorded for each subject. Across all subjects and stimulus types, this percentage was on average 6%, with a minimum of 0% (did not require the feature) and a maximum of 22%.

The procedure was implemented with custom-made scripts in MATLAB, and a graphical-user interface was used to monitor responses. The *in-situ* sound calibration levels were also monitored (see Subsec. 2.2.3).

2.4. Subjects

Eight subjects (1 female, 7 male), aged 25–43 yr (mean = 34 yr), participated. Their hearing thresholds were less than 15 dB hearing level (HL) over the range 125–4000 Hz, as assessed using standard audiometry (British Society of Audiology, 2018). To ensure normal sensitivity to low-frequency sounds, subjects underwent tympanometry; all had normal middle-ear pressure (between −20 and +20 daPA). Also the detection threshold for a 20-Hz pure tone with a 1000-ms duration was measured, using a two-alternative forced-choice adaptive procedure. All thresholds were ≤16 dB HL relative to the threshold specified in ISO-226, 2003, corrected by +3 dB to account for monaural listening. The experiment was approved by the Research Ethical Committee of Universidad de Las Américas.

3. Results

The individual and mean OLLPs are shown in Fig. 2. The average OLLPs (with across-subject SDs

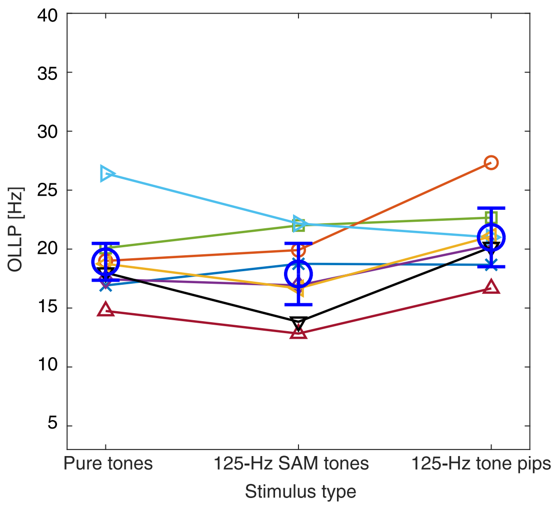

Fig. 2. Individual OLLP values (small symbols) and mean values (large circles) for pure tones, 125-Hz SAM tones and 125-Hz tone pips. The ordinate spans the frequency range used in the procedure. Error bars show the SDs of the repetition rates at the last six turnpoints, averaged across subjects and all runs for that condition.

in parentheses) were 18.9 (3.4), 17.9 (3.5), and 21.0 (3.1) Hz for pure tones, 125-Hz SAM tones, and 125-Hz pip trains, respectively. A one-way analysis of variance was performed on the OLLPs, with stimulus type as factor. There was no significant effect of stimulus type ($F_{2,21} = 1.79$, $p = 0.19$). The average LLP across all conditions was 19.3 Hz. To provide a measure of how much subjects changed the stimulus frequency around the LLP in the adaptive procedure, the error bars in Fig. 2 show the SD of the frequencies at the last 6 turnpoints, averaged across all runs and subjects. This "uncertainty range" was about ±2 Hz, a reasonably small proportion of the frequency span evaluated in the procedure and smaller than the across-subject SDs.

The OLLPs for the three stimulus types were positively correlated across subjects, with Pearson coefficients $r \geq 0.36$. However, only the correlation between OLLPs for the pure tones and SAM tones was significant ($r = 0.77$, $p < 0.05$).

4. Discussion

4.1. Comparison with previous work

The OLLPs found here for the complex stimuli are similar to that obtained by Guttman and Pruzansky (1962) for short periodic tone pulses whose spectral energy was reported to be audible for frequencies down to 130 Hz. The OLLP found here is also similar to that reported by Warren and Bashford (1981) for periodic broadband noise segments. Since the stimuli of Warren and Bashford were periodic, and had spectral components that formed a harmonic series, while our SAM and tone-pip trains and the stimuli of Guttman and Pruzansky had spectra that usually did not form a harmonic series, it seems likely that the OLLP for complex sounds is not determined by whether or not the stimuli have a harmonic structure; rather, the envelope repetition rate is the determining factor.

The OLLP found here for the complex stimuli is lower than the OLLP of about 30 Hz found by Krumbholz *et al.* (2000) using an F0 discrimination task and by Pressnitzer *et al.* (2001) using a melody-discrimination task. The discrepancy may be related to the different tasks and criteria used. The discrepancy may also be related to the fact that the stimuli used by Krumbholz *et al.* (2000) and by Pressnitzer *et al.* (2001) were highpass filtered, and the lowest cut off frequency used by them was 200 Hz. Our stimuli, and those of Guttman and Pruzansky (1962) and Warren and Bashford (1981), contained frequency components extending down to lower frequencies. It seems likely that the lowest OLLP is obtained when the stimuli contain very low-frequency components, below 200 Hz.

4.2. The role of resolvability

The pure-tone stimuli used here may be considered as containing a single resolved harmonic. We consider next whether the other stimuli used here contained any resolved components. The components in a complex sound need to be separated by about 1.25ERB$_N$ for them to be "heard out" or resolved with 75% accuracy (Plomp, 1964; Moore, Ohgushi, 1993; Moore *et al.*, 2006), where ERB$_N$ is the average value of the equivalent rectangular bandwidth of the auditory filter for listeners with normal hearing (Glasberg, Moore, 1990). If the components are separated by less than 1ERB$_N$ they may be regarded as completely unresolved. In the region of the OLLP, the lowest harmonic in the SAM stimuli would have had a frequency close to $(125 - 18)$ Hz $=$ 107 Hz. The value of ERB$_N$ for a center frequency of 107 Hz is about 36 Hz (Glasberg, Moore, 1990; Jurado, Moore, 2010). Hence, the components were separated by about 0.5ERB$_N$ and would have been completely unresolved. The tone-pip trains had spectra extending down to about 50 Hz. However, even for these stimuli, it is likely that all components were unresolved. Considering only cochlear processes (including the helicotrema shunt mechanism below about 40 Hz, but excluding the middle-ear "fixed" high-pass filter), the ERB of the auditory filter centered at 50 Hz was estimated by Jurado and Moore (2010) and Jurado *et al.* (2011) to be 26.5 Hz (this is also the narrowest bandwidth of any auditory filter, so it will be called ERB$_{MIN}$ in the following). Hence, in the region of the OLLP all components would have been separated by less than 0.79ERB$_N$, and again would have been completely unresolved. The OLLP was very similar for the three types of stimuli, suggesting that the OLLP does not depend on whether or not resolved harmonics are present. Rather, the OLLP seems to be determined by the repetition rate of the waveform for the pure tones and of the envelopes for the complex stimuli.

4.3. Possible factors determining the OLLP

4.3.1. Measurement of time delays by the auditory system

Given that the OLLP seems to be determined by the repetition rate of the sounds, it seems plausible that the OLLP is limited at least partly by the longest time interval that can be measured accurately by the auditory system. This was argued previously by Krumbholz *et al.* (2000) and Pressnitzer *et al.* (2001). For our data, this limiting time interval is about 53 ms. Yrttiaho *et al.* (2008) provided physiological evidence for such a limit. They measured auditory cortical responses to vowels using magnetoencephalography, for F0 values between 9 and 113 Hz. The auditory-evoked field (AEF) N1m response latency increased monotonically with decreasing F0 down to 19 Hz (see Fig. 4, right-hand panels, in (Yrttiaho *et al.*, 2008)). Below 19 Hz, this trend broke down and the response waveform resembled the shape of cascaded "transient N1m-like" responses. They interpreted this result as reflecting the limited capability of the auditory system to extract periodicity from the sound and suggested that the transition point reflects the OLLP. Our average OLLP of 19 Hz fits very well with the transition in N1m responses that they found.

It should be noted that the AEFs recorded by Yrttiaho *et al.* (2008) tended to increase in amplitude with decreasing F0. This trend is also apparent in the frequency-following responses (FFRs) described by Tichko and Skoe (2017; using triangular waves with frequencies down to F0 = 16.4 Hz) and has also been observed for FFRs to infrasonic and low-frequency pure tones, as used here (Jurado, Marquardt, 2020). It seems likely that response latency is a neural correlate of pitch but response amplitude is not; rather response amplitude probably depends on the degree of synchronization across various neural generators in the auditory pathway (Dau, 2003; Tichko, Skoe, 2017; Jurado, Marquardt, 2020).

Moore (1982) proposed that there was a limit to the range of inter-spike intervals in the auditory nerve that could be measured by the auditory system. This limit was assumed to be inversely proportional to the characteristic frequency (CF). Such a mechanism can account for the finding that the LLP varies with the frequency region of the harmonics. It can also account for why F0 discrimination is generally better when low-rank harmonics are present than when only high-rank harmonics are present (Bernstein, Oxenham, 2005). De Cheveigné and Pressnitzer (2006) proposed a mechanism by which a CF-dependent limit in the range of measurable time intervals might arise. They assumed that the durations of the impulse responses of the cochlear filters were inversely proportional to the filter bandwidths (as is the case for e.g. Gammatone filters; Carney, Yin, 1988). They also assumed that the phase at the output of each auditory filter could be adjusted by an amount that was proportional to frequency, and that a weighted sum of these shifted outputs could be performed. This gave what they called "synthetic delays". The delays could be used to perform periodicity analysis via a mechanism such as autocorrelation. The maximum synthetic delay is inversely proportional to the filter bandwidth, which can account for the center-frequency dependence of the LLP. The longest possible delay is limited by the duration of the impulse response of the auditory filter with the lowest center frequency, and this determines the OLLP. According to this theory, based on the OLLP measured here, the impulse response of the "bottom" auditory filter should last about 53 ms. This seems reasonable given that the lowest auditory filter has a center frequency of about 50 Hz and an ERB of 26.5 Hz (the ERB_{MIN}; from the "composite-cases" of Jurado and Moore, 2010, and Jurado *et al.*, 2011, i.e. bandwidth estimates reflecting cochlear processes).

4.3.2. Limitations on template formation

The internal representation of complex sounds with F0 < ERB_{MIN} is limited in two ways relative to that for sounds with higher F0s. Firstly, since the harmonics are separated by less than ERB_{MIN}, the excitation pattern will not display clear peaks and dips corresponding to the frequencies of the individual harmonics; the place representation of the spectrum will be degraded. Secondly, the patterns of phase locking will always reflect the interaction of two or more harmonics, rather than the frequencies of individual harmonics; the temporal representation of the spectrum will also be degraded. Several models of pitch predict that a complex tone will have a clear pitch when there are place and temporal representations of the frequencies of individual harmonics (Moore, 1982; Meddis, O'Mard, 1997; Bernstein, Oxenham, 2005; Shamma, Dutta, 2019). When such representations are lacking, the pitch may be extracted from the temporal fine structure of the response to unresolved harmonics if their harmonic number is below about 14–16 (Moore, Moore, 2003; Moore *et al.*, 2006a; 2009), or from the repetition rate of the temporal envelope, but this gives a less clear pitch.

It may be the case that the pitch of complex sounds is usually extracted using templates that are built up via exposure to sounds for which there are place and temporal representations of the frequencies of individual components (Terhardt, 1974; Gerson, Goldstein, 1978; Moore, 1982; Meddis, O'Mard, 1997; Bernstein, Oxenham, 2005; Shamma, Dutta, 2019). Once formed, these templates can also extract pitch from unresolved harmonics, albeit less precisely (Shamma, Dutta, 2019). For complex sounds with F0 < $\mathrm{ERB_{MIN}}$, the templates may not form properly, because of the degraded spectro-temporal representation of the harmonics. This is analogous to a neural network that has not been trained adequately (see e.g. Han, Wang, 2014; Drugman *et al.*, 2018, for examples of machine-learning approaches that rely on well-defined spectral information for pitch extraction). Studies of infants are consistent with the idea that the formation of templates based on spectro-temporal information is involved in pitch perception. Specifically, the cortical representation of the missing fundamental resembles that for adults once the infant reaches 3–4 months of age (He, Trainor, 2009), and this is about the same age as when frequency resolution reaches adult levels (Spetner, Olsho, 1990).

4.3.3. Minimum overlap with the frequency range of information-bearing modulations

There may be evolutionary factors that have influenced the OLLP. Most human and animal communication sounds are characterized by relatively slow envelope fluctuations (for examples of the latter, see Singh and Theunissen, 2003; Koumura *et al.*, 2019). For speech, periodic envelope fluctuations associated with the F0 convey information about the rate of vocal fold vibration, whereas the mostly non-periodic low-rate envelope fluctuations convey information about vocal-tract shape and movement. Most energy in the modulation spectrum of speech falls at frequencies below 20–30 Hz (Plomp, 1983; Kanedera *et al.*, 1999; Elliott, Theunissen, 2009; Varnet *et al.*, 2017). Also, very slow amplitude modulations (up to 20–40 Hz) convey the most important information for intelligibility (Drullman *et al.*, 1994; Kanedera *et al.*, 1999; Elliott, Theunissen, 2009). Usually, the range of F0s in human voices does not overlap with the modulation rates conveying information about vocal-tract shape and movement, making it easy for the auditory system to dissociate the two. Pitch perception may have evolved in such a way that pitch is not perceived for F0s that fall in the range of the most important species-specific modulation frequencies. Consistent with this idea, Joly *et al.* (2014) presented evidence that the OLLP for Rhesus Macaques is about 30 Hz, and, according to data on the same species collected by Fukushima *et al.* (2015), this lies just above the most prominent range of modulation frequencies present in their vocal calls (see Fig. 4 in (Fukushima *et al.*, 2015)).

4.4. Comparison of OLLPs and the lowest F0s of musical instruments

It is of interest to compare our OLLPs with the F0s of the lowest notes that can be played by different musical instruments. Figure 3 shows examples of these (many of these data were taken from: "Logos Foundation – Instrument frequencies and ranges," 2016) together with the mean ± 1 SD of the OLLP values. Only a few rare instruments are capable of pro-

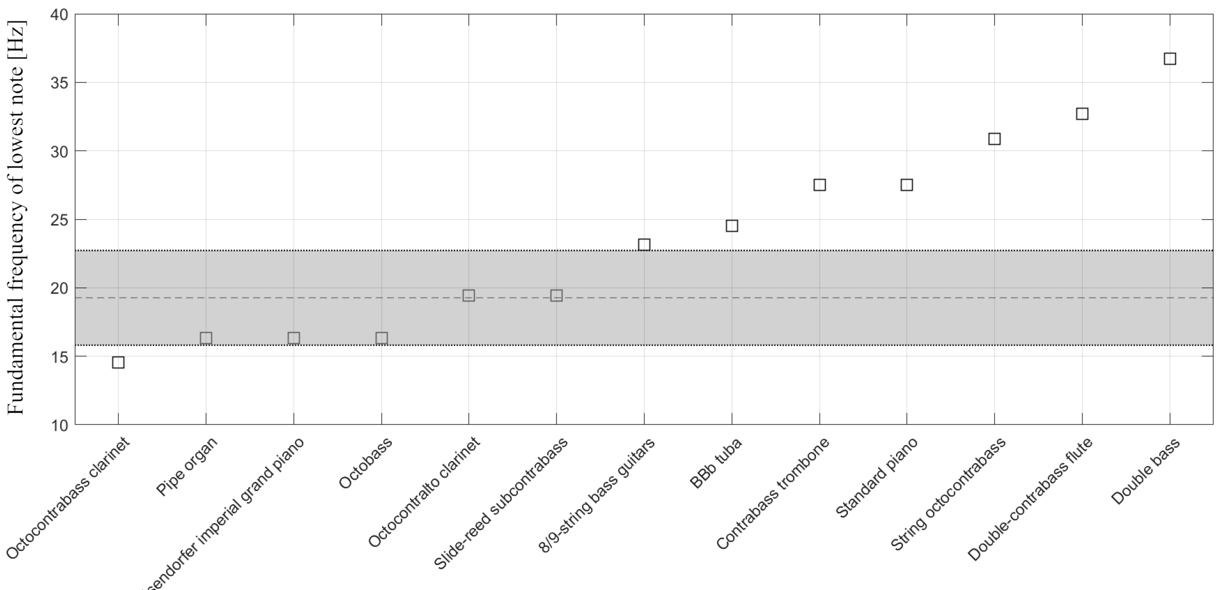

Fig. 3. Lowest-note F0 (Hz) of various musical instruments (squares). The shaded area shows ± 1 SD of the individual LLP values about the mean (dashed line).

ducing F0s below the range of the OLLPs measured here. It is not clear whether the makers of instruments such as the Octocontrabass clarinet thought that the lowest notes would evoke a pitch or whether they simply wanted to make an instrument that would produce extremely low notes, regardless of how they were perceived. In any case, it seems clear that most instruments that are in common use have lower F0 limits that are above the OLLP.

5. Conclusions

The overall lowest repetition rate for which sound can still convey a sensation of pitch was evaluated for pure tones, 125-Hz SAM tones and 125-Hz tone-pip trains. The OLLP had a mean value of 19 Hz and was similar for the three stimulus types. The results suggest that the OLLP is determined by the waveform repetition period for the pure tones and the envelope repetition period for the 125-Hz carrier stimuli. This is consistent with the idea that the OLLP is determined by the longest time interval that can be measured accurately by the auditory system, and that this interval is about 53 ms.

Given that the LLP for complex stimuli decreases as the frequencies of the lowest components in the stimuli decrease (Ritsma, 1962; Krumbholz *et al.*, 2000; Pressnitzer *et al.*, 2001), the common OLLP observed for our stimuli, which included pure tones with frequencies down to the infrasound range, suggests that the OLLP found here corresponds to the lowest LLP across the human hearing range.

Acknowledgments

We would like to thank all participants involved for their effort. Carlos Jurado also thanks Juan Vizuete for his help with the implementation of the setup.

References

1. Atal B.S. (1972), Automatic speaker recognition based on pitch contours, *The Journal of the Acoustical Society of America*, **52**(6B): 1687–1697, doi: 10.1121/1.1913303.
2. Bernstein J.G., Oxenham A.J. (2005), An autocorrelation model with place dependence to account for the effect of harmonic number on fundamental frequency discrimination, *The Journal of the Acoustical Society of America*, **117**(6): 3816–3831, doi: 10.1121/1.1904268.
3. British Society of Audiology (2018), *Recommended Procedure: Pure-tone air-conduction and bone conduction threshold audiometry with and without masking*, British Society of Audiology, Reading, UK.
4. Burke S. (1998), Missing values, outliers, robust statistics & non-parametric methods, *Valid Analytical Measurement Bulletin*, **19**: 22–27.
5. Carney L.H., Yin T.C.T. (1988), Temporal coding of resonances by low-frequency auditory nerve fibers: Single-fiber responses and a population model, *Journal of Neurophysiology*, **60**(5): 1653–1677, doi: 10.1152/jn.1988.60.5.1653.
6. de Cheveigné A. (1997), Concurrent vowel identification III: A neural model of harmonic interference cancellation, *The Journal of the Acoustical Society of America*, **101**(5): 2857–2865, doi: 10.1121/1.419480.
7. de Cheveigné A., Pressnitzer D. (2006), The case of the missing delay lines: Synthetic delays obtained by cross-channel phase interaction, *The Journal of the Acoustical Society of America*, **119**(6): 3908–3918, doi: 10.1121/1.2195291.
8. Cullen J.K., Long G.R. (1986), Rate discrimination of high-pass-filtered pulse trains, *The Journal of the Acoustical Society of America*, **79**(1): 114–119, doi: 10.1121/1.393762.
9. Dau T. (2003), The importance of cochlear processing for the formation of auditory brainstem and frequency following responses, *The Journal of the Acoustical Society of America*, **113**(2): 936–950, doi: 10.1121/1.1534833.
10. Drugman T., Huybrechts G., Klimkov, V., Moinet A. (2018), Traditional machine learning for pitch detection, *IEEE Signal Processing Letters*, **25**(11): 1745–1749, doi: 10.1109/LSP.2018.2874155.
11. Drullman R., Festen J.M., Plomp R. (1994), Effect of reducing slow temporal modulations on speech reception, *The Journal of the Acoustical Society of America*, **95**(5): 2670–2680, doi: 10.1121/1.409836.
12. Elliott T.M., Theunissen F.E. (2009), The modulation transfer function for speech intelligibility, *PLOS Computational Biology*, **5**: e1000302, doi: 10.1371/journal.pcbi.1000302.
13. Fastl H. (1983), Fluctuation strength of modulated tones and broadband noise, [in:] *Hearing – Physiological Bases and Psychophysics*, Klinke R., Hartmann R. [Eds], pp. 282–288, Springer, Berlin, Heidelberg, doi: 10.1007/978-3-642-69257-4_41.
14. Fukushima M., Doyle A.M., Mullarkey M.P., Mishkin M., Averbeck B.B. (2015), Distributed acoustic cues for caller identity in macaque vocalization, *Royal Society Open Science*, **2**(12): 150432, doi: 10.1098/rsos.150432.
15. Gerson A., Goldstein J.L. (1978), Evidence for a general template in central optimal processing for pitch of complex tones, *The Journal of the Acoustical Society of America*, **63**(2): 498–510, doi: 10.1121/1.381750.
16. Glasberg B.R., Moore B.C.J. (1990), Derivation of auditory filter shapes from notched-noise data, *Hearing Research*, **47**(1–2): 103–138, doi: 10.1016/0378-5955(90)90170-t.
17. Guttman N., Pruzansky S. (1962), Lower limits of pitch and musical pitch, *Journal of Speech, Language,*

and Hearing Research, **5**(3): 207–214, doi: 10.1044/jshr.0503.207.

18. Han K., Wang D. (2014), Neural networks for supervised pitch tracking in noise, [in:] *Proceedings of the International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, Florence, Italy, pp. 1488–1492, doi: 10.1109/ICASSP.2014.6853845.
19. He C., Trainor L.J. (2009), Finding the pitch of the missing fundamental in infants, *Journal of Neuroscience*, **29**(24): 7718–7722, doi: 10.1523/JNEURO SCI.0157-09.2009.
20. Hoeschele M. (2017), Animal pitch perception: melodies and harmonies, *Comparative Cognition and Behavior Reviews*, **12**: 5–18, doi: 10.3819/CCBR.2017.120002.
21. Hoke M., Ross B., Wickesberg R., Lütkenhöner B. (1984), Weighted averaging – theory and application to electric response audiometry, *Electroencephalography and Clinical Neurophysiology*, **57**(5): 484–489, doi: 10.1016/0013-4694(84)90078-6.
22. ISO-226 (2003), *Acoustics – normal equal-loudness contours*, International Organization for Standardization, Geneva, Switzerland.
23. Jackson H.M., Moore B.C.J. (2013), The dominant region for the pitch of complex tones with low fundamental frequencies, *The Journal of the Acoustical Society of America*, **134**(2): 1193–1204, doi: 10.1121/1.4812754.
24. Joly O., Baumann S., Poirier C., Patterson R.D., Thiele A., Griffiths T.D. (2014), A perceptual pitch boundary in a non-human primate, *Frontiers in Psychology*, **5**, Article 998, doi: 10.3389/fpsyg.2014.00998.
25. Jurado C., Gallegos P., Gordillo D., Moore B.C.J. (2017), The detailed shapes of equal-loudness-level contours at low frequencies, *The Journal of the Acoustical Society of America*, **142**(6): 3821–3832, doi: 10.1121/1.5018428.
26. Jurado C., Larrea M., Patel H., Marquardt T. (2020), Dependency of threshold and loudness on sound duration at low and infrasonic frequencies, *The Journal of the Acoustical Society of America*, **148**(2): 1030–1038, doi: 10.1121/10.0001760 .
27. Jurado C., Marquardt T. (2016), The effect of the helicotrema on low-frequency loudness perception, *The Journal of the Acoustical Society of America*, **140**(5): 3799–3809, doi: 10.1121/1.4967295.
28. Jurado C., Marquardt T. (2020), Brain's frequency following responses to low-frequency and infrasound, *Archives of Acoustics*, **45**(2): 313–319, doi: 10.24425/aoa.2020.133151.
29. Jurado C., Moore B.C.J. (2010), Frequency selectivity for frequencies below 100 Hz: Comparisons with mid-frequencies, *The Journal of the Acoustical Society of America*, **128**(6): 3585–3596, doi: 10.1121/1.3504657.
30. Jurado C., Pedersen C.S., Moore B.C.J. (2011), Psychophysical tuning curves for frequencies below 100 Hz, *The Journal of the Acoustical Society of America*, **129**(5): 3166–3180, doi: 10.1121/1.3560535.
31. Kanedera N., Arai T., Hermansky H., Pavel M. (1999), On the relative importance of various components of the modulation spectrum for automatic speech recognition, *Speech Communication*, **28**(1): 43–55, doi: 10.1016/S0167-6393(99)00002-3.
32. Kinsler L.E., Frey A.R., Coppens A.B., Sanders J.V. (1999), *Fundamentals of Acoustics*, 4th ed., New York: Wiley-VCH.
33. Koumura T., Terashima H., Furukawa S. (2019), Cascaded tuning to amplitude modulation for natural sound recognition, *Journal of Neuroscience*, **39**(28): 5517–5533, doi: 10.1523/JNEUROSCI.2914-18.2019.
34. Krumbholz K., Patterson R.D., Pressnitzer D. (2000), The lower limit of pitch as determined by rate discrimination, *The Journal of the Acoustical Society of America*, **108**(3): 1170–1180, doi: 10.1121/1.1287843.
35. Kühler R., Fedtke T., Hensel J. (2015), Infrasonic and low-frequency insert earphone hearing threshold, *The Journal of the Acoustical Society of America*, **137**(4): EL347–EL353, doi: 10.1121/1.4916795.
36. Logos Foundation (2016), *Instrument frequencies and ranges*, https://www.logosfoundation.org/kursus/frequency_table.html (date last viewed: 05-Oct-20).
37. Marquardt T., Hensel J., Mrowinski D., Scholz G. (2007), Low-frequency characteristics of human and guinea pig cochleae, *The Journal of the Acoustical Society of America*, **121**(6): 3628–3638, doi: 10.1121/1.2722506.
38. Meddis R., O'Mard L. (1997), A unitary model of pitch perception, *The Journal of the Acoustical Society of America*, **102**(3): 1811–1820, doi: 10.1121/1.420088.
39. Mehta A.H., Oxenham A.J. (2020), Effect of lowest harmonic rank on fundamental-frequency difference limens varies with fundamental frequency, *The Journal of the Acoustical Society of America*, **147**(4): 2314–2322, doi: 10.1121/10.0001092.
40. Mřller H., Pedersen C.S. (2004), Hearing at low and infrasonic frequencies, *Noise and Health*, **6**(23): 37–57.
41. Moore B.C.J. (1982), *An Introduction to the Psychology of Hearing*, 2nd ed., London: Academic Press.
42. Moore B.C.J. (2008), The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people, *Journal of the Association for Research in Otolaryngology*, **9**(4): 399–406, doi: 10.1007/s10162-008-0143-x.
43. Moore B.C.J. (2019), The roles of temporal envelope and fine structure information in auditory perception, *Acoustical Science and Technology*, **40**(2): 61–83, doi: 10.1250/ast.40.61 .
44. Moore B.C.J., Glasberg B.R., Flanagan H.J., Adams J. (2006), Frequency discrimination of complex tones; assessing the role of component resolvability and temporal fine structure, *The Journal of the*

Acoustical Society of America, **119**(1): 480–490, doi: 10.1121/1.2139070.

45. Moore B.C.J., Glasberg B.R., Low K.E., Cope T., Cope W. (2006), Effects of level and frequency on the audibility of partials in inharmonic complex tones, *The Journal of the Acoustical Society of America*, **120**(2): 934–944, doi: 10.1121/1.2216906.

46. Moore B.C.J., Gockel H.E. (2011), Resolvability of components in complex tones and implications for theories of pitch perception, *Hearing Research*, **276**(1–2): 88–97, doi: 10.1016/j.heares.2011.01.003.

47. Moore B.C.J., Hopkins K., Cuthbertson S. (2009), Discrimination of complex tones with unresolved components using temporal fine structure information, *The Journal of the Acoustical Society of America*, **125**(5): 3214–3222, doi: 10.1121/1.3106135.

48. Moore B.C.J., Ohgushi K. (1993), Audibility of partials in inharmonic complex tones, *The Journal of the Acoustical Society of America*, **93**(1): 452–461, doi: 10.1121/1.405625.

49. Moore G.A., Moore B.C.J. (2003), Perception of the low pitch of frequency-shifted complexes, *The Journal of the Acoustical Society of America*, **113**(2): 977–985, doi: 10.1121/1.1536631.

50. Oxenham A.J. (2008), Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants, *Trends in Amplification*, **12**(4): 316–331, doi: 10.1177/1084713808325881.

51. Patterson R.D. (1987), A pulse ribbon model of monaural phase perception, *The Journal of the Acoustical Society of America*, **82**(5): 1560–1586, doi: 10.1121/1.395146.

52. Plomp R. (1964), The ear as a frequency analyzer, *The Journal of the Acoustical Society of America*, **36**(9): 1628–1636, doi: 10.1121/1.1919256.

53. Plomp R. (1967), Pitch of complex tones, *The Journal of the Acoustical Society of America*, **41**(6): 1526–1533, doi: 10.1121/1.1910515.

54. Plomp R. (1983), The role of modulation in hearing, [in:] *Hearing – Physiological Bases and Psychophysics*, Klinke R., Hartmann R. [Eds], Springer, Berlin, pp. 270–276.

55. Pressnitzer D., Patterson R.D., Krumbholz K. (2001), The lower limit of melodic pitch, *The Journal of the Acoustical Society of America*, **109**(5): 2074–2084, doi: 10.1121/1.1359797.

56. Ritsma R.J. (1962), Existence region of the tonal residue – I, *The Journal of the Acoustical Society of America*, **34**(9A): 1224–1229, doi: 10.1121/1.1918307.

57. Santurette S., Dau T. (2011), The role of temporal fine structure information for the low pitch of high-frequency complex tones, *The Journal of the Acoustical Society of America*, **129**(1): 282–292, doi: 10.1121/1.3518718.

58. Seebeck A. (1841), Observations on some conditions of tone formation [in German: Beobachtungen über einige Bedingungen der Entstehung von Tönen], *Annalen der Physik*, **129**(7): 417–436, doi: 10.1002/andp.18411290702.

59. Shackleton T.M., Carlyon R.P. (1994), The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination, *The Journal of the Acoustical Society of America*, **95**(6): 3529–3540, doi: 10.1121/1.409970.

60. Shamma S., Dutta K. (2019), Spectro-temporal templates unify the pitch percepts of resolved and unresolved harmonics, *The Journal of the Acoustical Society of America*, **145**(2): 615–629, doi: 10.1121/1.5088504.

61. Singh N.C., Theunissen F.E. (2003), Modulation spectra of natural sounds and ethological theories of auditory processing, *The Journal of the Acoustical Society of America*, **114**(6): 3394–3411, doi: 10.1121/1.1624067.

62. Spetner N.B., Olsho L.W. (1990), Auditory frequency resolution in human infancy, *Child Development*, **61**(3): 632–652, doi: 10.1111/j.1467-8624.1990.tb02808.x.

63. Terhardt E. (1974), Pitch, consonance, and harmony, *The Journal of the Acoustical Society of America*, **55**(5): 1061–1069, doi: 10.1121/1.1914648.

64. Tichko P., Skoe E. (2017), Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators, *Hearing Research*, **348**: 1–15, doi: 10.1016/j.heares.2017.01.014.

65. Tukey J.W. (1977), *Exploratory Data Analysis*, Reading, Mass: Addison-Wesley Pub.

66. Varnet L., Ortiz-Barajas M.C., Erra R.G., Gervain J., Lorenzi C. (2017), A cross-linguistic study of speech modulation spectra, *The Journal of the Acoustical Society of America*, **142**(4): 1976–1989, doi: 10.1121/1.5006179.

67. Walker K.M.M., Schnupp J.W.H., Hart-Schnupp S.M.B., King A.J., Bizley J.K. (2009), Pitch discrimination by ferrets for simple and complex sounds, *The Journal of the Acoustical Society of America*, **126**(3): 1321–1335, doi: 10.1121/1.3179676.

68. Warren R.M., Bashford J.A. (1981), Perception of acoustic iterance: Pitch and infrapitch, *Perception and Psychophysics*, **29**(4): 395–402, doi: 10.3758/BF03207350.

69. Yrttiaho S., Tiitinen H., May P.J.C., Leino S., Alku P. (2008), Cortical sensitivity to periodicity of speech sounds, *The Journal of the Acoustical Society of America*, **123**(4): 2191–2199, doi: 10.1121/1.2888489.