version 251013

An efficient continuous-tone procedure for the study of frequency discrimination

Jason Tzu-Hsien Lien[1,5], Abhinav Uppal[2], Rupesh Chillale[3,9], Nahaleh Fatemi[4], Torsten Marquardt[5], Alain de Cheveigné[6,7,8]

1. SOUND Lab, Department of Clinical Neurosciences, University of Cambridge, UK.
2. Shu Chien-Gene Lay Department of Bioengineering, University of California, San Diego, USA.
3. Institute for Systems Research & Department of Electrical and Computer Engineering, University of Maryland, College Park, Maryland, USA.
4. Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, USA.
5. University College London, Ear Institute, United Kingdom.
6. Laboratoire des Systèmes Perceptifs, Unité Mixte de Recherche 8248, Centre National de la Recherche Scientifique, Paris, France.
7. Département d'Etudes Cognitives, Ecole Normale Supérieure, Paris Sciences et Lettres University, France.
8. University College London, Department of Psychology, United Kingdom.
9. School of Arts of Sciences, Ahmedabad University, Ahmedabad 380009, India.

Corresponding author: Alain de Cheveigné, alain.de.cheveigne@ens.fr

Abstract

This paper describes a method of gathering experimental data for the study of frequency discrimination. The method is efficient in terms of participant time, is easy to deploy on an online platform to gather behavioral responses, and allows simultaneous measurement of brain activity without contamination from onset responses. The participant hears a continuous tone within each block, with frequency steps at predetermined intervals, and is invited to respond “up” or “down” to these steps by pressing one of two keys. The magnitude of the steps varies according to a predetermined schedule (constant-stimuli, not adaptive) from large (easy) to small (hard), and the discrimination threshold is derived from a psychometric function fitted to the responses to this sequence. Comparison with earlier published procedures shows that the new method yields comparable thresholds, albeit less variable and slightly lower, and thus presumably closer to sensory limits.

I. Introduction

Estimation of thresholds of sensory detection or discrimination is an early goal of psychophysics (Fechner 1860). In particular, frequency discrimination thresholds, measured as a function of parameters such as frequency, amplitude, duration, or spectral envelope, are informative about mechanisms of pitch perception (e.g. Moore 1973; Micheyl et al. 2006; Demany et al. 2016). In a typical experiment, a participant is presented with a sequence of trials each made up of two, three, or four tones separated by silent gaps, and responds to each trial. For two tones, the question might be up vs down, for three tones it might be which of them differs from the other two, for four tones it might be which of the first or second pair contains tones that differ in pitch. A three- or four-tone procedure has benefits in terms of flexibility or insensitivity to bias (e.g. Semal and Demany 2006), but each trial is longer, and it may take more time to gather the required number of responses. Conversely, a *one-tone-per-trial* procedure may be more efficient than any of these and, indeed, Arzounian et al. (2017a) showed that a sliding two-alternative forced choice (2AFC) procedure that involved one tone per trial yielded discrimination thresholds for frequency change direction (up vs down) similar to those for a classic two-tone-per-trial procedure, in a shorter time. The present study explores the possibility of further improvements by removing the silent gap between tones. The experiment is divided into short blocks during which the stimulus consists of a *continuous tone*, frequency-modulated in steps. The participant is asked to indicate after each step if the pitch went up or down, an example of *continuous psychophysics* (Burge and Bonnen 2025).

Threshold estimation commonly involves an adaptive staircase procedure in which the stimulus for each trial depends upon on the participant's previous response(s). An adaptive procedure has two advantages: the stimulus parameter continuum is sampled most densely in the vicinity of the threshold, where responses are most informative and, with appropriate rules, the threshold itself can be derived algorithmically from the response data without the need to fit a psychometric function (e.g. Levitt, 1971). Considerable effort has been devoted to optimizing adaptive procedures (e.g. Watson and Pelli 1983; Kaernbach 1991; Paire et al. 2023). In contrast, in a "constant-stimuli" procedure the parameter values are predetermined, a sigmoidal function is fitted to the responses to model the participant's psychometric function, and the threshold is derived as the abscissa value at criterion performance (e.g. 75% correct). The drawback is a suboptimal sampling of the parameter space, and thus suboptimal use of experiment time, but there are several potentially useful features, in addition to simplicity, that are reviewed in the Discussion. The present study adopts a constant-stimuli design.

In an adaptive threshold measurement procedure, the difficulty of the task fluctuates within a block, hard trials alternating with easy trials, as the level of difficulty is adaptively adjusted based on the participant's success or failure. Pitch change judgments are subject to bias from preceding trials (Arzounian et al. 2017b, and references therein), and it is conceivable that bias from large frequency steps might affect judgments of subsequent smaller frequency steps, contributing to variability. In the present study, frequency step size decreases monotonically over a block, so a step is always preceded by a step of only slightly larger size. Bias due to context might still exist, but it is potentially smaller.

In summary, our methodology differs from that for most previous studies in the use of (1) a continuous tone within a block, rather than a sequence of pulsed tones, (2) a constant-stimuli method rather than an adaptive procedure, (3) trials ordered from easy-to-hard. The aim of this study was to compare, in the same participants, the new procedure with previous procedures. Anticipating the results, frequency discrimination thresholds for the new method were comparable, albeit slightly smaller and less variable, and there was no indication of a drawback.

The new procedure might be attractive in several ways. As mentioned, a one-step-per-trial procedure is relatively efficient in terms of responses per unit time (Arzounian et al. 2017a). A constant-stimuli procedure is easy to implement, in particular in an online setting, as the stimuli presented in a block are contained in a single file, and there are no response-adaptive adjustments to the stimulus. The easy-to-hard ordering of conditions within a block is subjectively pleasant and possibly less prone to fluctuations in criteria or motivation than the rapidly fluctuating difficulty of an adaptive procedure. This ordering may also facilitate procedural and perceptual learning (Liu et al. 2008; Wisniewski et al. 2017). Finally, a continuous tone allows recording of brain responses to frequency change without the salient onset response associated with each tone. We illustrate this last point with electroencephalography (EEG) data from one participant recorded for both types of procedure (continuous and pulsed) with identical parameters.

II. Methods

A. Behavior

The aim was to compare, in the same participants, the new constant-stimuli-continuous-tone procedure with an adaptive, pulsed-tone procedure that has previously been validated with respect to a standard two-tone-per-trial 2AFC procedure (Arzounian et al. 2017a). The experiments were carried out during the COVID pandemic, and some aspects of the methodology were affected by our inability to communicate face-to-face with participants.

1. Conditions

Three conditions were defined. In condition 1 (the new method) the stimulus was a continuous tone (Fig. 1A) with frequency steps of decreasing size (Fig. 1C). The participant was invited to respond "up" or "down" after each step (Fig. 1B) by pressing one of two keys on a keyboard. Condition 2 was identical to 1 except that the amplitude was modulated to form a sequence of tones with the same predetermined frequencies as in condition 1. Condition 3 was similar to condition 2, but the frequencies were adjusted adaptively using a staircase procedure. Frequency transitions were regularly spaced in time for conditions 1 and 2, but were slightly irregularly spaced for condition 3 (Fig. 1A), each new tone appearing one second after the response to the previous tone (Fig. 1B). By comparing conditions, we could probe the effects of using a continuous tone rather than a series of discrete tones (1 vs 2) and of using a fixed easy-to-hard schedule of frequency steps rather than an adaptive procedure (2 vs 3). Condition 3 was previously compared with a classic two interval 2AFC procedure (Arzounian et al. 2017b).

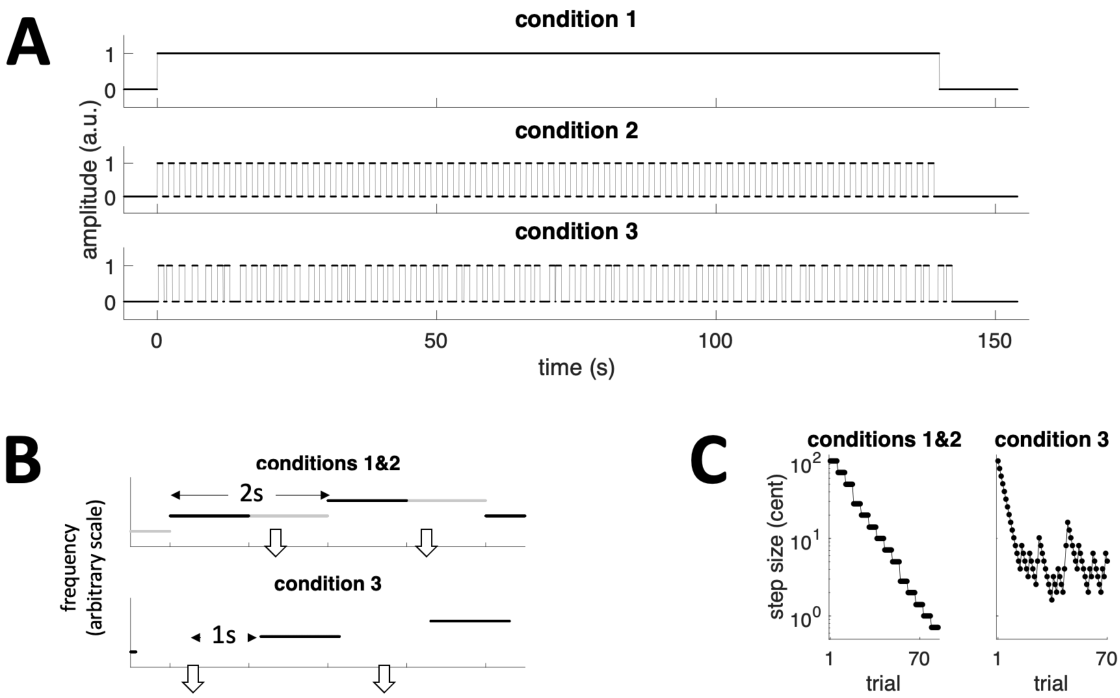

FIG. 1. Structure of the stimuli in the behavioral experiment. A: Amplitude as a function of time for the three conditions. Within each block of trials, the stimulus was a continuous tone for condition 1, a sequence of regularly spaced tones for condition 2, a sequence of tones with timing dependent on participant's responses for condition 3. B: For conditions 1 and 2, frequency steps were spaced at 2 s intervals; for condition 3, each new tone occurred 1 s after the response to the previous tone. The arrows symbolize the participant's response. C: Step size as a function of trial number. For conditions 1 and 2, the frequency step size decreased monotonically over the block. For condition 3, the step size fluctuated according to the adaptive procedure.

2. Stimuli

The experiment was divided into blocks of ~140 s duration. Within each block, the stimulus was a pure tone with variable frequency for condition 1, or a succession of tones each of duration 1 s for conditions 2 and 3, either regularly spaced at 2-s intervals for condition 2, or slightly irregularly spaced for condition 3, as spacing depended on the participant's response time (Fig. 1A). The stimulus frequency changed in steps with random sign, either regularly at intervals of 2 s (conditions 1 and 2) or 1 s after the participant's response to the previous frequency step (condition3) (Fig. 1B). For conditions 1 and 2, the magnitude of the frequency step decreased after every trial by a nominal factor of $1/\sqrt{2}$, starting at 100 cents (one semitone, 1/12 octave) and finishing at 0.71 cents. The sequence of step sizes was adjusted to include powers of 10 (values: 100, 71, 50, 35.5, 25, 17.25, 10, 7.1, etc.), hence a slight irregularity in the magnitude of the decrement (Fig. 1C, left). For condition 3 the step size was varied adaptively depending on whether the participant's response was correct (factor $2^{-1/3}$) or incorrect (factor 2) (Fig. 1C, right) to ensure good sampling of the psychometric curve in the region of 75% correct (Kaernbach 1991).

For condition 1, the stimulus for a block was synthesized as $s(t) = \sin(2\pi\varphi(t))$ where the phase $\varphi(t)$ increased continuously at a rate determined by the instantaneous frequency, avoiding a waveform discontinuity at each frequency transition. The phase function was further convolved with a rectangular window of duration 50 ms to smooth the frequency transitions themselves. The starting frequency was 1 kHz and subsequent frequencies were constrained to remain between 0.707 and 1.414 kHz by repeatedly drawing a random sequence until this condition was fulfilled. The resulting sequence converged to an asymptotic frequency within that range, the value of which depended on the randomly chosen step signs. Condition 2 used the same waveform as condition 1, but the signal was amplitude-modulated into “pulses” of duration 1 s separated by gaps of 1 s, with onsets and offsets shaped by a raised-cosine ramp of duration 50 ms. For condition 3, the stimulus for each trial was a tone of duration 1 s with 50-ms onset and offset ramps, synthesized on-the-fly with a frequency that depended on the response to the previous trial.

3. Procedure

Each session included 9 blocks in which conditions 1, 2, 3 were interleaved in that order. The three repetitions of each condition allowed checking for serial effects due to learning and/or fatigue. The first blocks of conditions 1 and 3 were preceded by a short training block that the participant could repeat at will. The entire session was controlled by a single experimental script that welcomed the participant, offered background information, gave instructions, presented the stimuli, and gathered the participant's responses. The participant could pause at will between blocks, or terminate the experiment and withdraw from the study at any time.

The participant was told to promptly press a key ('1' for up, '2' for down) on a keyboard after each frequency step according to whether the pitch went up or down. Feedback was provided by setting the color of a rectangle centered on a dark gray screen to dark green for a correct answer and dark red for an incorrect answer or miss. For conditions 1 and 2, the stimulus within a block did not depend on the participant's responses. For condition 3 the step size was adjusted adaptively as described in Section II.A.2. For conditions 1 and 2, the steps occurred at a fixed rate of one every two seconds. For condition 3 each step occurred one second after the response to the previous step. In condition 1, the participant had no information as to *when* a transition occurred, other than the percept of a pitch change; in conditions 2 and 3 this information was carried by the pulse timing. No participant complained that the fixed rate was difficult to follow; subjectively, the pace felt pleasantly brisk.

4. Participants

Participants (38) were recruited among acquaintances of the first author. Four were excluded for insufficient response numbers (indicating early interruption or extremely poor performance); of the remaining 34, twenty-nine reported playing an instrument; 13 of these played regularly. Experiments were conducted remotely via a dedicated web site. The participant downloaded the experimental script and a Matlab runtime (with technical assistance from the first author).

5. Data analysis

Response data for all conditions were treated by fitting a sigmoidal function to the correct response rate for each step size (logarithmically scaled):

$$C(x) = 0.5 + 0.5/(1 + e^{s(x_0 - x)})$$

where $x = \log_2|df|$, df is the relative change in frequency, x_0 is the threshold, and s is the slope. The response rate at each step size was calculated by counting 1 for a correct answer, 0 for an incorrect answer, and 0.5 for a missed response (this assimilates missing to guessing), and dividing the sum by the number of repeats. The fit was performed for each block and participant, using the Matlab function nlinfit(), yielding a threshold and a slope for each condition.

B. EEG

A potential appeal of the continuous-tone procedure is that it allows concomitant recording of brain responses (EEG or MEG) without contamination by prominent onset responses triggered by each tone. To illustrate this idea, we use data from one participant recorded at the Telluride Neuromorphic Engineering Workshop (2025) within the context of an unrelated experiment. EEG signals were recorded with a 32-channel BrainVision system at a sampling rate of 25 kHz (the main experiment involved recording auditory brainstem responses, hence the relatively high sampling rate). One channel was used to record the stimulus envelope to ensure synchronization and 31 channels were devoted to EEG. The data were down-sampled to 100 Hz, detrended by subtraction of the linear trend over segments of 2 s stitched together by overlap-add with 1-s overlap (de Cheveigné and Arzounian 2018), and cut into trials to form a matrix with dimensions time × channels × trials. The participant performed five blocks each for conditions 1 (continuous tone) and 2 (pulsed tone), interleaved; for each condition, the five matrices were concatenated into a single matrix that was then sorted according to trial number, resulting in two 3D matrices of size 200 samples × 31 channels × 332 trials, one for each condition. Data were high-pass filtered by fitting a sinusoid with period 200 samples and subtracting the fit (analogous to a "brickwall" filter with cutoff above 0.5 Hz) so as to attenuate relatively slow non-sensory components. Outlier trials (defined by a Euclidean distance from the average over trials greater than the Euclidean distance of the average from zero) were removed, and the data matrices for each condition were processed by the JD (joint decorrelation) algorithm (de Cheveigné and Parra, 2014) which produces a spatial filter optimized to extract the stimulus-locked response. The spatial filter was applied to the data for each condition, resulting in a 200 sample × N_c matrix for each, where N_c is the number of trials after outlier removal.

III. Results

A. Behavior

Data of 34 participants were available for analysis. The response on each trial was coded as 1 (correct), 0 (incorrect) or 0.5 (missed), a sigmoidal psychometric function was fitted as described in the Methods, and the threshold for each block was derived from its value at 0.75. Figure 2 shows the threshold for each condition for each participant averaged over three blocks (thin, color) and the geometric mean over participants (thick, black). A pairwise t-test indicated a lower threshold for condition 1 (new method) than for condition 3 (old method), p=0.0001, uncorrected. A goal of the study was to check that thresholds are *not higher for the new method than the old*, as a higher threshold might indicate that the new method introduced some new difficulty for the participant. This result confirms that they are indeed not higher, and the additional outcome that they are *significantly lower* adds weight to that conclusion, as it rules out lack of statistical power. This is the main behavioral result of this paper: the new, continuous-tone easy-to-hard constant-stimuli threshold estimation method does not yield higher thresholds than the pulsed-tone sliding adaptive method. That method was previously evaluated by Arzounian et al. (2017a) who found thresholds comparable to a classic two-tone-per-trial procedure. The difference between conditions 1 (continuous tone) and 2 (pulsed tone) was marginally significant (p=0.03, uncorrected), that between conditions 2 and 3 (adaptive procedure) was significant (p=0.005, uncorrected).

Averaging over conditions, thresholds were slightly lower for the third block (final) than for the second (p=0.009, uncorrected) or first (p<0.002, uncorrected), suggesting that the benefits of learning overcame any fatigue effects. Consistent with previous reports (Micheyl et al. 2006), thresholds differed widely between participants.

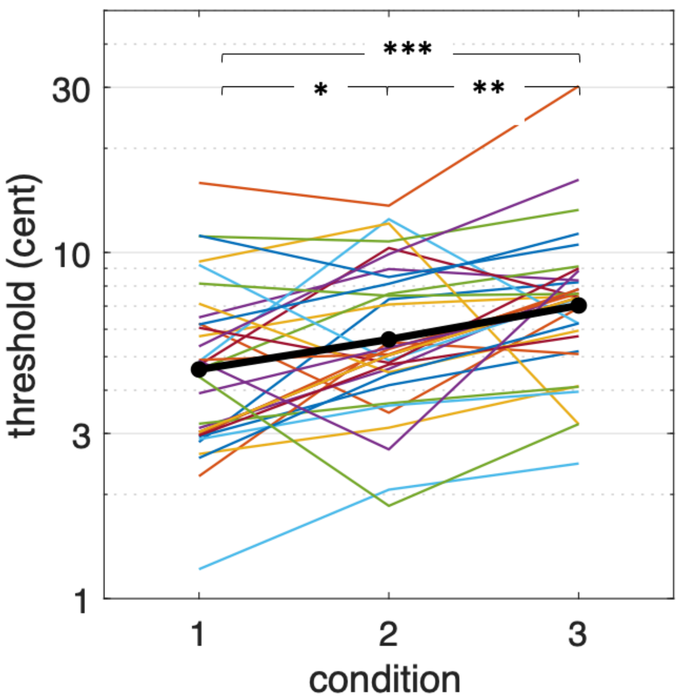

Fig. 2. Frequency discrimination threshold for each condition (1: continuous, 2: pulsed, 3: adaptive). Colored lines show results for individual participants. The black line shows the geometric mean.

The slope parameter of the fit to the psychometric function (not shown) was significantly greater (steeper) for conditions 1 and 2 than condition 3 (p<0.0001 for both, uncorrected) but only marginally different between conditions 1 and 2 (p=0.04, uncorrected). The steeper slope for conditions 1 and 2 might be related to the easy-to-hard ordering: as the step size becomes small and participants miss an increasing number of trials (Fig. 3C), they may stop responding (see Discussion).

A curious result is that latencies for condition 1 (continuous tone) tended to be smaller than for conditions 2 and 3 (pulsed tones) (Fig. 3A). One might have expected the opposite because conditions 2 and 3 offer an amplitude cue that indicates the instant of frequency change. For conditions 1 and 2, latency tended to increase over the duration of the block (Fig. 3B, blue and green), as expected as the frequency step became harder to detect (Fig. 1C, left), but the two curves were roughly parallel. For condition 3, the latency increased rapidly and then reached a plateau (Fig. 3B, red) as expected because the step size remained within a limited range (as was illustrated in Fig. 1C, right).

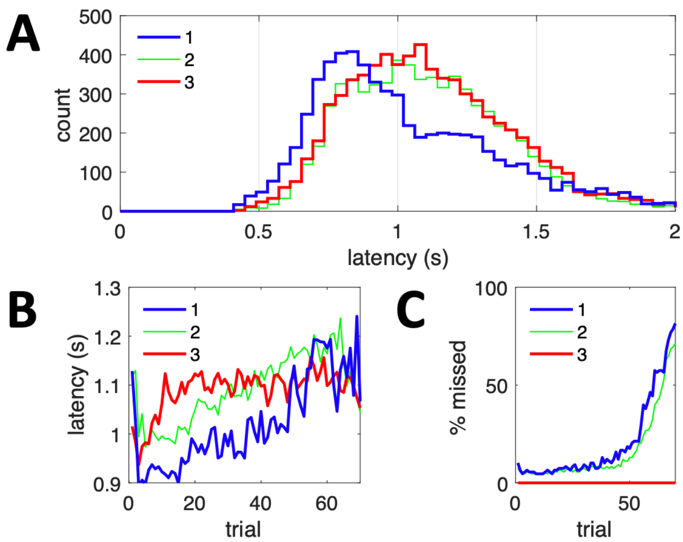

Fig. 3. A: Histograms of behavioral response latency for the three conditions (continuous, pulsed, adaptive), averaged over participants and blocks. B: Response latency as a function of trial number within a block. C: Percentage of trials missed as a function of trial number within a block. For condition 3 (adaptive) the percentage is necessarily zero (red).

B. EEG

EEG data were recorded from one participant performing 5 blocks of condition 1 (continuous) and 5 blocks of condition 2 (pulsed), interleaved. Data were preprocessed as described in the Methods. Figure 4 shows the time course of the most repeatable component of the JD analysis (see Methods) for condition 1 (A) and condition 2 (B), averaged over trials (left) or as a raster plot indexed by trial number (right). For condition 2 (pulsed), the response appears to be triggered by the onset, and to a lesser degree the offset, of each tone (Fig. 4B). This response does not differ systematically between early trials (large step) and late trials (small step) (Fig. 4B, right), and thus is unlikely to reflect the sensory or perceptual correlates of frequency change, as these should vanish as the step size becomes tiny. For condition 1 (continuous), the response to the frequency transition clearly depends on frequency step size, in amplitude and possibly also latency (Fig. 4A, right), suggesting that it reflects a sensory or perceptual correlate of frequency change. The point of this demonstration is that the continuous-tone paradigm allows such responses to be observed without contamination from tone onsets and offsets.

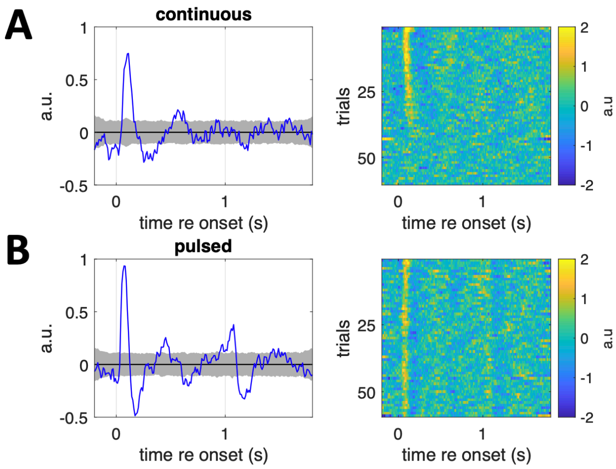

Fig. 4. A: Frequency step response averaged over trials (left) and raster plot of per-trial responses (right) for condition 1 (smoothed over groups of 5 trials). The values represent the output of an optimal spatial filter (see Methods) with arbitrary gain, hence the lack of units. The gray band (left) represents ±2 SD of the mean). B: As A but for condition 2.

IV. Discussion

The aim of this study was to validate a procedure for measuring frequency discrimination thresholds by comparison with a previously developed procedure, itself previously validated by comparison with a standard two-interval 2AFC procedure (Arzounian et al. 2017a). The results suggest that the new procedure yields discrimination thresholds comparable to those for the earlier methodss, without any obvious drawback.

The efficiency of a threshold-measurement procedure can be quantified as the inverse of the time required to reliably measure the threshold (Treutwein et al. 1995). An important factor influencing this is the rate at which responses can be gathered, which is the inverse of the duration of each trial. Figure 5 shows the time per trial for various procedures (see legend of Fig. 5 for assumptions). The new procedure (top bar) takes roughly half the time of a comparable adaptive two-tone-per-trial procedure (4th bar). As expected, the 4 tone per trial procedure is most expensive (its duration can of course be reduced by making tones and inter-tone intervals shorter; for example, the 4-tone-per-trial procedure of Semal and Demany (2006) required only 3.8 s per trial).

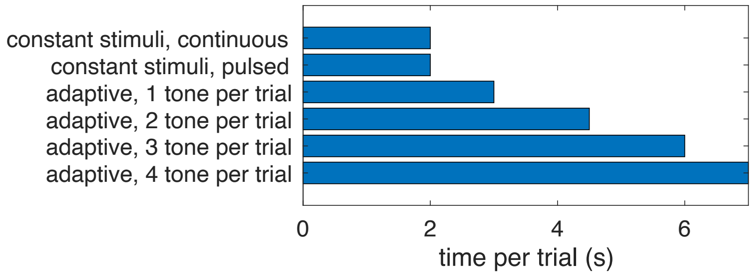

Fig. 5: Time per trial for procedures tested in conditions 1-3 (first three bars) and for classic procedures with two, three and four tones per trial (last three bars). For pulsed procedures (bars 2-6) the tone duration is 1 s, for multi-tone procedures (bars 4-6) the inter-tone interval is 0.5 s (for four tones, the interval between second and third tones is 1 s), for adaptive procedures (bars 3-6) the next trial starts 1 s after the response. The average participant response time is assumed to be 1 s.

Efficiency also depends on the choice of the levels (here frequency step sizes) that sample the psychometric function. Adaptive methods automatically ensure a denser sampling in the vicinity of the threshold where responses are most informative (e.g. Levitt 1972; Treutwein et al. 1995; Kristensen et al. 2025), and Bayesian methods further optimize this sampling (e.g. Watson and Pelli 1983), whereas constant-stimuli procedures lack such benefits. This study used a constant-stimuli procedure with a relatively wide range of frequency changes (0.71 to 100 cents) to accommodate the expected diversity of frequency discrimination thresholds (Micheyl et al. 2006). This is clearly suboptimal because a large proportion of trials led to ceiling or floor performance, and thus are not informative. This drawback can be alleviated by adjusting the range of levels from one block to the next (in which case the procedure is adaptive at the block level but not trial level). However, this was not attempted here. Here, the adaptive and non-adaptive procedures yielded results of comparable quality (in terms of value and variability) for similar experiment durations, suggesting that suboptimal sampling was offset by other qualities of the non-adaptive procedure.

In the new procedure, stimuli are presented in the order easy to hard (large to small frequency step), in contrast to the fluctuating difficulty inherent in trial-by-trial adaptation. Hypothetically, this might have several benefits. First, initial easy trials might favor the participant's understanding of the task, similar to the initially-descending staircase typically included in adaptive procedures, and promote procedural or sensory learning (Liu et al. 2008; Wisniewski et al. 2017). Second, the frequency step size progresses monotonically rather than fluctuating rapidly, avoiding repeated criterion changes that might add variance to the data. When the task is difficult and feedback is mostly negative, the participant may lose motivation, change criterion, or attend to a different dimension of the stimulus. With an easy-to-hard constant-stimuli procedure, this occurs only once, whereas with an adaptive procedure it occurs multiple times within a block. The participant may further entertain cognitive activity related to the adaptive rule (e.g., “I just made an error, therefore the next trials will be easier”), possibly at the expense of attention to the stimulus. A potential drawback of the easy-to-hard procedure is that the motivation to “push the limit” before giving up might differ between participants. However, a similar issue affects adaptive procedures.

A feature of the new procedure is that frequency changes are carried by a tone that is continuous. Perception of such frequency transitions, and of frequency modulation in general, has been investigated in many previous studies (Demany et al. 2009). A prime motivation here was to develop a procedure that allows concomitant recording of brain responses with EEG or magnetoencephalography (MEG), without contamination from onset responses (Fig. 4). Brain responses to frequency changes have been investigated in the literature, where they go under the name of *acoustic (or auditory) change comp*lex (ACC) (Liang et al. 2016; Zhang et al. 2021; Vonck et al. 2019, 2021; Guérit et al. 2023). These are often measured by repeatedly presenting a tone with a frequency transition positioned halfway (e.g. Liang et al. 2016), but the yield is increased (and contamination from the onset response reduced) if a continuous tone with multiple transitions is used instead of a series of separate tones each with a single transition (e.g., Guérit et al. 2023). More generally, there has been a recent uptick in interest in so-called “continuous psychophysics” in which responses are elicited at various time points within a continuous stimulus (Burge and Bonnen 2025).

Continuous stimulation increases the duration of stationary portions, which might lead to lower thresholds (Moore 1973), but a drawback is that the lack of a clear timing cue might cause the participant to miss transitions, leading to higher thresholds. For behavioral studies, this might be addressed by adding a faint cue (e.g. click, or amplitude fluctuation, or visual cue) at each transition. Otherwise, participants must rely on the regularity of the transition timing (here one per 2 s). Probing the detection of frequency change in participants who are unable to judge change direction (Semal and Demany 2006) would require an unpredictable timing schedule.

An interesting observation is that participants responded *faster* to a frequency change carried by a continuous tone than to the same change carried by a pulsed tone. The difference, on the order of 100 ms (compare blue and green traces in Figs. 3A and B), cannot be ascribed to the onset ramp applied to each tone (25 ms to half height). The reason for this effect is not clear. It might be the case that the momentary waveform aperiodicity introduced by an amplitude change in the pulsed condition interferes with the period estimation mechanism required to detect a frequency change. Alternatively, it might be the case that to ignore the perceptually salient tone onset requires an active process that takes time, as in the inter-feature masking phenomena described by Barascud et al (2014). This phenomenon may warrant further investigation, but is outside the scope of this paper.

Overall, it appears that the new procedure has few drawbacks compared to a classic procedure. Its non-adaptive design may ease implementation in an online setting, the easy-to-hard progression may promote learning and reduce variability, and the continuous-tone design facilitates recording of brain responses. Subjectively, the relatively fast pace (1 response every 2 s) is refreshing, as noted earlier by Arzounian et al. (2017a) for their sliding procedure.

V. Conclusion

This paper described a method of measuring frequency discrimination that is efficient in terms of participant time and easy to deploy on an online platform. Validation by comparison with earlier published procedures shows that the new method yields comparable thresholds, albeit slightly lower and thus presumably closer to sensory limits, and less variable. Electroencephalography (EEG) responses to frequency steps were uncontaminated by tone onset responses.

Acknowledgments

The first author is supported by a Taiwan Cambridge Scholarship. The last author is supported by grants ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*. EEG experiments were piloted at the Telluride Neuromorphic Engineering Workshop (2024 and 2025). Laurent Demany provided useful comments on an earlier version of this paper.

Conflict of interest

The authors report no conflict of interest.

Ethical approval

Data were collected under UCL ethics approval #0565/004.

Data availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

Arzounian D, de Kerangal, M, de Cheveigné A (2017a) A sliding two-alternative forced-choice paradigm for pitch discrimination, *J. Acoust. Soc. Am.* 142, 167-172, doi:10.1121/1.4992030.

Arzounian, D, de Kerangal, M., de Cheveigné A (2017b) Sequential dependencies in pitch perception, *J. Acoust. Soc. Am.* 142, 3047, doi: 10.1121/1.5009938.

Barascud, N., Griffiths, T. D., McAlpine, D., & Chait, M. (2014). "Change Deafness" Arising from Inter-feature Masking within a Single Auditory Object. *J. Cog. Neurosci.*, *26*(3), 514–528. https://doi.org/10.1162/jocn_a_00481

Burge, J., & Bonnen, K. (2025). Continuous psychophysics: Past, present, future. *Trends in Cognitive Sciences*, *29*(5), 481–493. https://doi.org/10.1016/j.tics.2025.01.005

de Cheveigné, A., & Arzounian, D. (2018). Robust detrending, rereferencing, outlier detection, and inpainting for multichannel data. *NeuroImage*, *172*, 903–912. https://doi.org/10.1016/j.neuroimage.2018.01.035

de Cheveigné, A., & Parra, L. C. (2014). Joint decorrelation, a versatile tool for multichannel data analysis. *NeuroImage*, *98*, 487–505. https://doi.org/10.1016/j.neuroimage.2014.05.068

Demany, L., Carlyon, R. P., & Semal, C. (2009). Continuous versus discrete frequency changes: Different detection mechanisms? *J. Acoust. Soc. Am.*, *125*(2), 1082–1090. https://doi.org/10.1121/1.3050271

Demany, L., Lucas, T., & Semal, C. (2016). Pitch priming in sequences of two sounds. *J. Acoust. Soc. Am.*, *140*(3), 2056–2063. https://doi.org/10.1121/1.4963093

Guérit, F., Harland, A. J., Richardson, M. L., Gransier, R., Middlebrooks, J. C., Wouters, J., & Carlyon, R. P. (2022). Electrophysiological and Psychophysical Measures of Temporal Pitch Sensitivity in Normal-hearing Listeners, *J. Assoc. Res. Otolaryngol.*, *24*(1), 47–65. https://doi.org/10.1007/s10162-022-00879-7

Fechner, G. T. (1860). Elemente der Psychophysik, Vol. 2. Wiesbaden: Breitkopf u. Härtel.

Kaernbach, C. (1991). “Simple adaptive testing with the weighted up-down method,” *Percept. Psychophys*. 49, 227–229.

Kristensen, S. B., Bødkergaard, K., & Bibby, B. M. (2025). Analysing the bias introduced by adaptive designs to estimates of psychometric functions. *J. Math. Psychol.*, *124*, 102899. https://doi.org/10.1016/j.jmp.2025.102899

Levitt, H. (1971). “Transformed up-down methods in psychoacoustics,” *J. Acoust. Soc. Am.* 49, 467–477.

Liang, C., Earl, B., Thompson, I., Whitaker, K., Cahn, S., Xiang, J., Fu, Q.-J., & Zhang, F. (2016). Musicians Are Better than Non-musicians in Frequency Change Detection: Behavioral and Electrophysiological Evidence. *Frontiers in Neuroscience*, *10*. https://doi.org/10.3389/fnins.2016.00464

Liu, E. H., Mercado, E., Church, B. A., & Orduña, I. (2008). The easy-to-hard effect in human (Homo sapiens) and rat (Rattus norvegicus) auditory identification. *J. Comp. Psychol.*, *122*(2), 132–145. https://doi.org/10.1037/0735-7036.122.2.132

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hear. Res.*, *219*(1–2), 36–47. https://doi.org/10.1016/j.heares.2006.05.004

Moore, B. C. J. (1973). Frequency difference limens for short-duration tones. *J. Acoust. Soc. Am.*, *54*(3), 610–619. https://doi.org/10.1121/1.1913640

Paire, A., Hillairet De Boisferon, A., & Paeye, C. (2022). Empirical validation of QUEST+ in PSE and JND estimations in visual discrimination tasks. *Behav. Res. Methods*, *55*(8), 3984–4001. https://doi.org/10.3758/s13428-022-02001-4

Semal, C., & Demany, L. (2006). Individual differences in the sensitivity to pitch direction. *J. Acoust. Soc. Am.*, *120*(6), 3907–3915. https://doi.org/10.1121/1.2357708

Treutwein, B. (1995). Adaptive Psychophysical Procedures. *Vis. Res.*, *35*, 2503–2522.

Vonck, B. M. D., Lammers, M. J. W., Van Der Waals, M., Van Zanten, G. A., & Versnel, H. (2019). Cortical Auditory Evoked Potentials in Response to Frequency Changes with Varied Magnitude, Rate, and Direction. , *J. Assoc. Res. Otolaryngol.*, *20*(5), 489–498. https://doi.org/10.1007/s10162-019-00726-2

Vonck, B. M. D., Lammers, M. J. W., Schaake, W. A. A., Van Zanten, G. A., Stokroos, R. J., & Versnel, H. (2021). Cortical potentials evoked by tone frequency changes compared to frequency discrimination and speech perception: Thresholds in normal-hearing and hearing-impaired subjects. *Hear. Res.*, *401*, 108154. https://doi.org/10.1016/j.heares.2020.108154

Wisniewski, M. G., Radell, M. L., Church, B. A., & Mercado, E. (2017). Benefits of fading in perceptual learning are driven by more than dimensional attention. *PLOS ONE*, *12*(7), e0180959. https://doi.org/10.1371/journal.pone.0180959.

Zhang, F., McGuire, K., Firestone, G., Dalrymple, K., Greinwald, J., & Fu, Q.-J. (2021). Cortical processing of location and frequency changes of sounds in normal hearing listeners. *Hear. Res.*, 400, 108110. https://doi.org/10.1016/j.heares.2020.108110.

Watson, A. B., & Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method. *Percept. Psychophys.*, *33*(2), 113–120. https://doi.org/10.3758/BF03202828