On the relation between pitch and level

Yi Zheng[1,2,3], Romain Brette[1,2]
[1]Sorbonne Universités, UPMC Univ Paris 06, INSERM, CNRS, Institut de la Vision, 17 rue Moreau, 75012 Paris, France.
[2]Institut d'Etudes de la Cognition, Ecole Normale Supérieure, Paris, France
[3]Beijing Advanced Innovation Center for Future Education, Beijing Normal University, Beijing, China.

Corresponding author : Romain Brette
Institut de la Vision, 17, rue Moreau, 75012 Paris, France
romain.brette@inserm.fr
+33 1 53 46 25 36

Abstract

Pitch is the perceptual dimension along which musical notes are ordered from low to high. It is often described as the perceptual correlate of the periodicity of the sound's waveform. Previous reports have shown that pitch can depend slightly on sound level. We wanted to verify that these observations reflect genuine changes in perceived pitch, and were not due to procedural factors or confusion between dimensions of pitch and level. We first conducted a systematic pitch matching task and confirmed that the pitch of low frequency pure tones, but not complex tones, decreases by an amount equivalent to a change in frequency of more than half a semitone when level increases. We then showed that the structure of pitch shifts is anti-symmetric and transitive, as expected for changes in pitch. We also observed shifts in the same direction (although smaller) in an interval matching task. Finally, we observed that musicians are more precise in pitch matching tasks than non-musicians but show the same average shifts with level. These combined experiments confirm that the pitch of low frequency pure tones depends weakly but systematically on level. These observations pose a challenge to current theories of pitch.

Highlights

- Pitch of low frequency pure tones decreases with increasing sound level.
- Pitch of harmonic complex tones does not change with sound level.
- Musical training has no effect on sound-level effects
- The relationship between pitch and level poses a challenge to current theories of pitch

Key words: Pitch, sound level, pure tone

1. Introduction

Pitch is the perceptual dimension along which tones are ordered from low to high on a musical scale. For musical tones, the main physical attribute that determines pitch is the repetition rate or fundamental frequency of the sound's waveform (Oxenham, 2012). Accordingly, theories of pitch perception, which can be broadly categorized as emphasizing either temporal cues or cochlear place of activation, have focused on how the auditory system might extract fundamental frequency (de Cheveigne, 2005). Recent psychophysical work has focused on distinguishing between these theories by assessing the perception of relatively complex pitch-evoking sounds, for example dichotic sounds (Bernstein and Oxenham, 2003), transposed tones (Oxenham et al., 2011) or mistuned harmonics (Hartmann and Doty, 1996). In this study, we wanted to address a more basic question: to what extent is the pitch of musical tones the perceptual correlate of fundamental frequency (for complex tones) or frequency (for pure tones)? In other words, is there a one-to-one mapping between pitch and (fundamental) frequency?

A number of earlier studies suggest that this is actually not exactly the case, specifically that the pitch of a pure tone can depend weakly on its level, a finding that is not straightforward to explain with standard theories of pitch (Licklider, 1951; Terhardt, 1974a). According to Stevens' rule (Stevens, 1935), the pitch of low-frequency pure tones (<500 Hz) decreases with increasing level while the pitch of high-frequency pure tones (>4000 Hz) increases with increasing level. This finding was obtained by a relatively indirect method, mainly with one subject, in which two tones of different frequencies were presented and the subject was instructed to change the second tone's level so that the two pitches matched. A similar finding for low frequencies was mentioned by Fletcher (1934) and shown with another method by Snow (1936), who asked subjects to rank two tones of different levels and frequencies as higher or lower in pitch; a lack of effect on complex tones was also mentioned (but not shown). These results were later confirmed with more subjects (Morgan et al., 1951; Terhardt, 1974b; Terhardt, 1975), although with substantial inter-individual variability, using a pitch matching method – the subject adjusted the frequency of the second tone to match the pitch of the first tone. At 200 Hz, when the tone level was increased from 40 dB to 80 dB SPL, the pitch shifted down by an amount equivalent to about half a semitone (Terhardt, 1974b), well above the just noticeable difference (for tones of identical level). This was confirmed more exhaustively with two subjects by Verschuure and van Meeteren (1975). Terhardt (1975) reported small pitch shifts with complex tones, but the shifts varied markedly across participants, and the statistical significance of the shifts was not assessed.

The goals of this psychophysical study were (1) to show that the reported changes with level truly reflect the level dependence of melodic pitch (as opposed to procedural biases or confusion of the perceptual dimensions of pitch and loudness); (2) to determine the level dependence of the pitches of pure tones and complex tones with identical fundamental frequency; (3) to determine whether the phenomenon is influenced by musical experience. Using pitch- and interval-matching experiments, we found that the pitch of low frequency pure tones, but not complex tones, depends on level regardless of musical experience, and that the measured phenomenon reflects small but actual changes in melodic pitch that partially transfer to the perception of melodic intervals.

2. Materials and Methods

2.1 Subjects and equipment

Ethical approval was granted by the local ethics committee (Comité pour la Protection des Personnes Ile de France). All subjects were fully informed about the goal of the study and provided written consent before their participation. All subjects had normal hearing (<20 dB hearing loss (HL) between 100 and 8000 Hz), and were 18-35 years old. Subjects in the non-musician group had never played an instrument or only briefly played one (<2 years) during childhood. Subjects in the musician group had 7-22 years of musical training, and played at least 1 hour per day at the time of the experiments (Table 1). Experiments 1 and 3 included 4 non-musicians and 4 musicians; experiment 2 included 2 musicians and 2 non-musicians from the same pool; experiment 4 included 6 other musicians.

Stimuli were generated digitally at a sampling rate of 44.1 kHz. Stimuli were presented diotically via a RME Fireface UC soundcard and a Sennheiser HD580 headphone. Sound levels were calibrated for each tone frequency with a sound pressure level meter, giving estimated sound levels at the eardrum. Subjects were seated individually in a double-walled, sound-insulated booth.

Table 1. *List of subjects and their musical experience.*

Subject	Age	Sex	Musical experience	Experiments
SAG	22	M	Drums, 12 years	1, 3
PES	22	F	None	1, 2, 3
PSS	25	M	None	1, 2, 3
PAM	22	F	Piano, 14 years	1, 2, 3
PJC	33	M	Guitar, 15 years	1, 2, 3
PTC	24	F	Trumpet, 15 years	1, 3
PAL	20	M	None	1, 3
SYZ	33	F	None	1, 3
LHK	27	F	Piano and violin, 10 years	4
PPG	27	M	Violon, 7 years	4
SLC	21	F	Violin, 15 years	4
SLJ	20	M	Piano, guitar and saxophone, 10 years	4
SEC	28	F	Violin, 22 years	4
SOP	23	M	Trumpet, 10 years	4

2.2 Experiments 1 and 2: Pitch matching of pure tones

Each trial began with a 500-ms reference pure tone followed by a 300-ms gap and a 500-ms comparison tone. Listeners were asked to adjust the frequency of the comparison tone until its pitch matched that of the reference tone. The starting frequency of the comparison tone was randomly chosen from a uniform distribution on a discrete semitone scale with a range of 4 semitones around the frequency of the reference tone. After each trial, listeners could adjust the frequency of the comparison tone up or down by 2, 0.5 or 0.125 semitones, without exceeding ±4 semitones around the reference frequency, could elect to hear the same tone pair again, or could indicate that they were satisfied with the pitch match. Listeners were encouraged to start with a big step size and then change to a smaller step size, and to adjust the comparison tone below and above the chosen frequency before making a final decision. No feedback was provided. The reference pure tone had a frequency of 200, 1000 or 4000 Hz. The level of reference tone varied from 20 to 70 dB SPL in steps of 10 dB. In experiment 1, the level of comparison pure tone was set to 40 dB SPL. In experiment 2, it also varied from 20 to 70 dB SPL. Each combination of reference and comparison levels was presented 10 times.

For experiment 1, the comparison level was fixed, while reference levels and frequency conditions were randomized between trials. For experiment 2, comparison levels were randomized across listening sessions. In each listening session, the comparison level was fixed but reference levels and frequency conditions were randomized. Each session contained 18 different conditions (6 reference levels times 3 frequency conditions), and lasted 5-20 minutes depending on the subject. Experiment 1 consisted of 10 sessions per subject (i.e., 10 trials for each condition), while experiment 2 consisted of 60 sessions per subject (6 comparison levels times 10 trials) – on average about 10 hours in total per subject.

2.3 Experiment 3: Pitch matching of harmonic complex tones

The pitch-matching procedure was the same as for Experiment 1, except that both the reference and comparison tones were harmonic complex tones, composed of 6 or 12 consecutive harmonics (order 1~6 or 1~12) with equal amplitude and random phase. The level of the comparison tones was fixed at 30 dB SPL per component (overall level was thus 37.8 or 40.8 dB), and the level of reference tones was varied from 10 to 60 dB SPL per component. The initial fundamental frequency of the comparison tone was randomly chosen from a uniform distribution on a discrete semitone scale with a range of 4 semitones around the fundamental frequency of the reference tone (200, 1000 or 4000 Hz).

2.4 Experiment 4: Pitch interval matching

An interval consisted of a pair of 500-ms pure tones presented in sequence (no gap). Each trial began with a reference interval, followed by a 300-ms gap and then a comparison interval. The resulting sequence of 4 tones always increased in frequency. Subjects were instructed to adjust the frequency of the last tone of the comparison interval until its pitch interval matched that of the reference interval, with the same procedure as in the pitch-matching experiments. To prevent listeners from memorizing the reference interval, its size was set to 2 or 3 semitones with equal probability. To ensure that listeners were performing the task by comparing intervals rather than the pitches of individual tones, the frequency of the first tone of the reference interval was set randomly at 200 Hz ±3 semitones (uniform distribution on a semitone scale), and the frequency of the first tone of the comparison interval was set randomly between 0 and 3 semitones above that of the second tone of the reference interval. The levels of the two tones in the reference pair were as follows: (1) increasing level, 40 dB SPL then 70 dB SPL; (2) decreasing level, 70 dB SPL then 40 dB SPL; (3) fixed level, 55 dB SPL then 55 dB SPL. The level of the two tones of the comparison interval was fixed at 55 dB SPL.

3. Results

3.1 Level dependence of the pitch of pure tones

Fig. 1 shows the results of experiment 1 for 8 subjects, of whom 4 were musicians (dashed lines). For the low-frequency reference tone (200 Hz), all subjects lowered the matching frequency of the comparison tone when the reference tone's level increased (Fig. 1A; ANOVA, F_5=19.68, p<0.001, effect size η^2=0.71). These results indicate that the reference tone sounded lower in pitch when its level was increased, consistent with Stevens' rule (Stevens, 1935) and previous observations with 300-Hz tones (Verschuure and van Meeteren, 1975). Averaged over subjects, the mean pitch shift exceeded 0.6 semitone (3.5% frequency change) at 70 dB SPL (Fig. 1D, black). At 1000 Hz, there was a smaller but significant downward shift in pitch with increasing level (Fig. 1B; ANOVA, F_5=6.87, p<0.001, effect size η^2=0.45), reaching about 0.2 semitone (1.1%) at 70 dB (Fig. 1D, blue). At 4000 Hz, the mean pitch shift was small and not significant (Fig. 1C and 1D, red; ANOVA, F_5=0.81, p=0.55, effect size η^2=0.09). Thus the pitch of high-frequency pure tones does not depend on level, in contradiction with Stevens rule. However, we cannot exclude the possibility that a significant effect would be seen with more subjects, and a subset of subjects did show the upward shift observed by Stevens for one subject. Stevens also observed upward pitch shifts at high levels for pure tones with frequencies higher than 4000 Hz (up to 12 000 Hz). We chose to restrict our study to tones with repetition rate lower than 4000 Hz because only these elicit a pitch sensation salient enough to carry melodic information (Attneave and Olson, 1971).

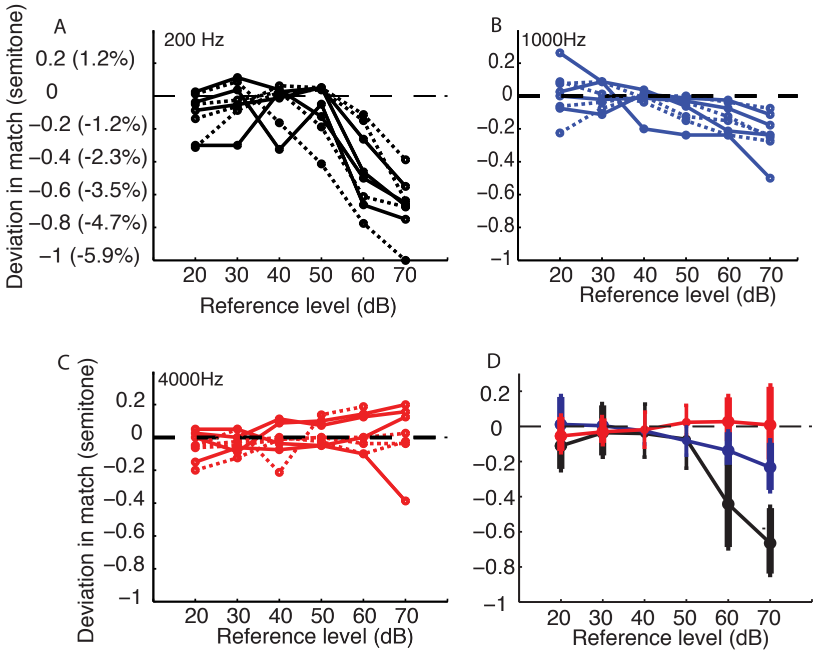

Figure 1. Level-dependence of the pitch of pure tones. A-C. Subjects adjusted the frequency of a 40 dB SPL tone so as to match the pitch of a reference tone. Each curve represents the frequency shift in semitones at the matching point relative to the frequency of the reference tone (A: 200 Hz; B: 1000 Hz; C: 4000 Hz), as a function of reference tone level, for one subject (solid lines: musicians; dashed lines: non musicians). D. Average frequency shift across subjects for reference tones at 200 Hz (red), 1000 Hz (blue) and 4000 Hz (red). Error bars represent ±1 standard deviation (SD) across subjects.

Psychophysical experiments on pitch perception are often done with musically trained subjects, because their results tend to be more consistent and therefore require fewer trials. We tested whether the same effects were seen for non-musicians and musicians. There was no significant difference in pitch shift between the two groups (t-test, $p = 0.9$): for 200 Hz tones; musicians (more than 10 years of musical training, Table 1) and non-musicians showed a negative pitch shift of the same magnitude with increasing level. However, the precision of pitch matching was markedly better for musicians (t-test of standard deviations, SDs, $p<0.001$), consistent with previous studies of pitch discrimination (Kishon-Rabin et al., 2001; McDermott et al., 2010): SDs were 0.11 (200 Hz), 0.075 (1000 Hz) and 0.07 semitones (4000 Hz) for musicians (corresponding to 0.64, 0.43 and 0.41% frequency change), vs. 0.18, 0.084 and 0.13 semitones for non-musicians (1, 0.49 and 0.75%), averaged across all levels.

We selected 2 musicians and 2 non-musicians for an in-depth analysis of the level dependence of pitch (experiment 2). We wanted to make sure that performance was indeed determined by the pitches of the two tones taken individually and was not influenced by some procedural aspect. For example, one might hypothesize that the pitch of the second tone is influenced by properties of the first tone, via some adaptive or context-dependent process (Chambers and Pressnitzer, 2014). To this end, we varied the level of both the comparison and reference tones between 20 and 70 dB SPL in steps of 10 dB, yielding a matrix of pitch shifts (Table 2). Each curve in Fig. 2A-C shows the pitch shift as a function of reference tone level for a particular comparison tone level.

Table 2. *Frequency shifts in semitones for all combinations of reference and comparison levels (experiment 2). Columns: reference level; rows: comparison level.*

200 Hz

	Reference level					
Comparison level	20 dB SPL	30 dB SPL	40 dB SPL	50 dB SPL	60 dB SPL	70 dB SPL
20 dB SPL	0.02±0.13	-0.02±0.11	-0.10±0.11	-0.02±0.13	0.32±0.16	0.39±0.20
30 dB SPL	0.07±0.10	-0.02±0.14	-0.05±0.20	0.04±0.14	0.34±0.13	0.47±0.15
40 dB SPL	0.12±0.12	0.10±0.13	-0.01±0.13	0.16±0.14	0.42±0.17	0.56±0.13
50 dB SPL	0.03±0.14	-0.05±0.12	-0.02±0.15	0.06±0.14	0.32±0.16	0.48±0.09
60 dB SPL	-0.28±0.20	-0.39±0.13	-0.47±0.14	-0.28±0.11	0.00±0.18	0.30±0.11
70 dB SPL	-0.42±0.18	-0.43±0.19	-0.65±0.18	-0.39±0.16	-0.17±0.20	0.02±0.19

1000 Hz

	Reference level					
Comparison level	20 dB SPL	30 dB SPL	40 dB SPL	50 dB SPL	60 dB SPL	70 dB SPL
20 dB SPL	0.00 ±0.06	0.00±0.08	0.05±0.09	0.10 ±0.09	0.15 ±0.16	0.20±0.12
30 dB SPL	0.02 ±0.07	-0.05±0.06	0.01±0.09	0.09 ±0.10	0.17±0.14	0.21 ±0.07
40 dB SPL	-0.04±0.07	-0.04±0.03	-0.04±0.06	0.01 ±0.05	0.15±0.07	0.20±0.08
50 dB SPL	-0.14±0.17	-0.11±0.10	-0.09±0.07	0.00±0.07	0.07 ±0.07	0.12 ±0.06
60 dB SPL	-0.21±0.08	-0.18±0.08	-0.14±0.07	-0.06±0.07	-0.01±0.06	0.06±0.05
70 dB SPL	-0.19±0.07	-0.16±0.06	-0.26±0.04	-0.08±0.04	-0.04±0.06	0.00 ±0.06

4000 Hz

	Reference level					
Comparison level	20 dB SPL	30 dB SPL	40 dB SPL	50 dB SPL	60 dB SPL	70 dB SPL
20 dB SPL	-0.02±0.08	0.01±0.11	-0.02±0.12	0.00±0.13	-0.02±0.11	-0.05±0.07
30 dB SPL	0.03±0.09	0.03±0.10	-0.03±0.10	-0.03±0.12	-0.01±0.11	-0.04±0.06
40 dB SPL	0.00±0.08	-0.01±0.09	0.02±0.09	-0.01±0.09	-0.04±0.10	-0.05±0.08
50 dB SPL	0.01±0.07	-0.02±0.10	0.02±0.09	0.01±0.10	-0.04±0.10	-0.05±0.07
60 dB SPL	-0.07±0.10	-0.03±0.09	0.04±0.06	0.00±0.11	0.00±0.11	-0.02±0.06
70 dB SPL	0.04±0.08	0.03±0.09	0.02±0.15	0.05±0.09	0.03±0.07	0.04±0.10

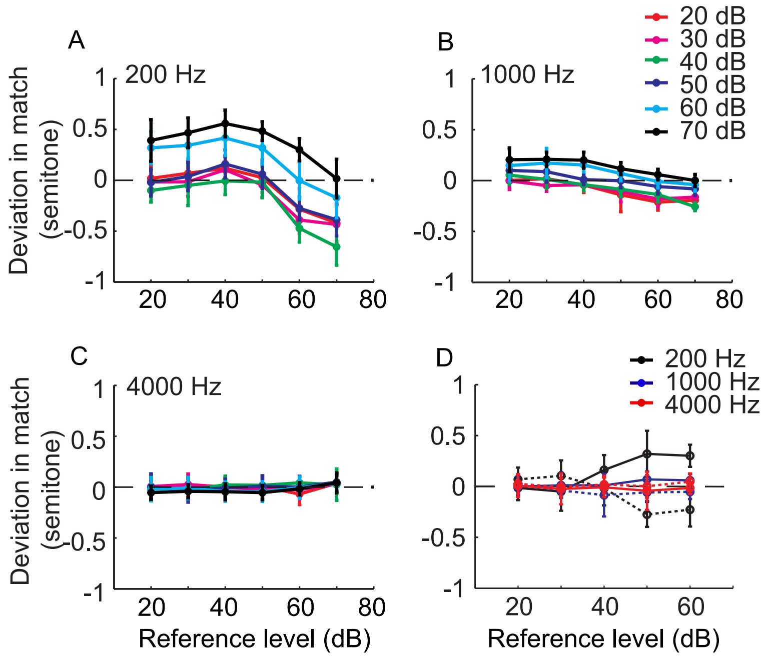

Figure 2. Detailed analysis of the level dependence of pure tone pitch for 4 subjects. A-C. average frequency shift as a function of reference level, for different comparison tone levels (20 to 70 dB SPL; A: 200 Hz, B: 1000 Hz, C: 4000 Hz). Errors bars indicate ±1 sd. D. Average frequency shift as a function of reference level, when the comparison level was 10 dB higher (solid line) or lower (dashed line) than that of the reference, for the 3 frequencies (same color code as in Fig. 1).

If the observed frequency shift required to equate the pitch of two tones with different levels, A and B, were entirely determined by the difference in pitch of the two tones, then it should be possible to express this shift S(A,B) (in semitones) as a difference S(A,B) = Pitch(B) – Pitch(A). This implies in particular anti-symmetry (S(A,B) = -S(B,A)) and transitivity (S(A,C) = S(A,B) + S(B,C)). An alternative hypothesis could be, for example, that the observed shift is a function of the difference in levels between the two tones because of contextual effects: S(A,B) = f(B-A). Our data do not follow this hypothesis, as shown in Fig. 2D, where the observed pitch shift is plotted as a function of reference level when the level of the comparison tone was 10 dB above (solid) or below (dashed) that of the reference tone. The difference in level was the same for all data points in each of the two sets. At 200 Hz, it appears that the pitch shift depends on reference level and not only on level difference (repeated measures ANOVA, $p<0.01$; no significant effect at 1000 Hz and 4000 Hz).

We now examine the hypothesis that the observed shift can indeed be expressed as a difference S(A,B) = Pitch(B) – Pitch(A). As there were no significant shifts for high frequency tones (4000 Hz, Fig. 2C), we only analyzed the results for reference tones of 200 Hz and 1000 Hz. First, a two-way repeated measures ANOVA on the observed shifts vs. the levels of reference and comparison tones showed no significant interaction ($p=0.9$), while each of the two levels had a significant effect on the reported shift ($p<0.001$ for reference and comparison levels, at 200 Hz and 1000 Hz). This means that the data are statistically consistent with the hypothesis that the observed shift is a sum of two quantities determined by each of the two tones: $S(A,B) = f_1(A) + f_2(B)$. Second, Fig. 2A shows that the pitch shift was near 0 when the reference and comparison tones had the same level, i.e., $S(A,A) = 0$ (mean ± s.d.: 0.01 ± 0.05 semitones at 200 Hz; -0.01 ± 0.06 semitones at 1000 Hz; 0.01 ± 0.07 semitones at 4000 Hz). Therefore, for all tones, $f_1(A) + f_2(A) = 0$, i.e. $f_1 = -f_2$. It follows that $S(A,B) = f_2(B) - f_2(A)$, i.e., the observed shift can indeed be expressed as a difference between two identical functions of tone level. This property also appears in Fig. 2A: curves showing the relationship between pitch shift and reference tone level all have the same shape, and differ by a constant shift that only depends on the level of the comparison tone.

Overall, this analysis strongly supports the claim that the pitch of low frequency pure tones decreases when the level increases above 50 dB SPL.

3.2 Level-dependence of the pitch of complex tones

In experiment 3, the reference tone was a harmonic complex tone with fundamental frequency 200 Hz, composed of either the first 6 harmonics (200 – 1200 Hz) or the first 12 harmonics (200 – 2400 Hz). In the first case, all harmonics were resolved, whereas in the second case the higher harmonics were not. All subjects reported that they heard the complex tone as a whole without hearing out individual components.

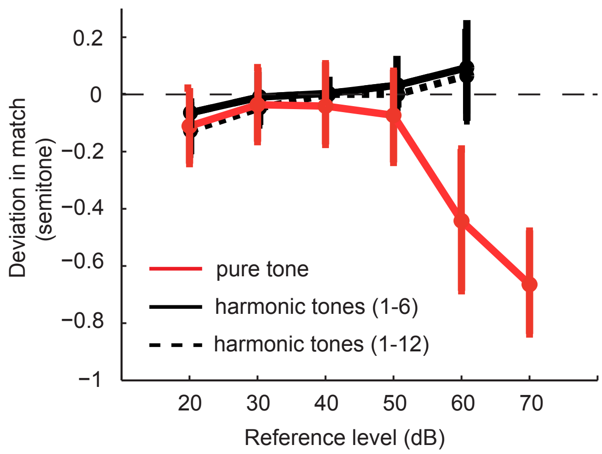

Figure 3. Level dependence of the pitch of complex tones. Black: average frequency shift at matching point as a function of reference level (in dB SPL per component) for complex tones with fundamental frequency = 200 Hz, composed of the first 6 harmonics (solid line) or the first 12 harmonics (dashed line). The comparison tone was a complex tone with a level of 30 dB SPL per component. Red: frequency shift for 200 Hz pure tones (Fig. 1D).

As shown in Fig. 3, there was no significant effect of level on the pitch of the complex tones, whether they contained only resolved harmonics or also unresolved harmonics (ANOVA, p=0.13). This result is important because for the resolved complex tone, the pitch of each individual component was presumably sensitive to level when presented in isolation, as experiments 1 and 2 suggest.

3.3 Level dependence of interval perception

Non-musically trained listeners can discriminate differences in musical intervals of around 1 semitone (Burns and Ward, 1978; Burns and Campbell, 1994; McDermott et al., 2010), but precision is greater for musicians (McDermott et al., 2010). As experiment 1 showed that the level dependence of pure tone pitch was similar for musicians and non-musicians, all 6 subjects of experiment 4 were musicians. In this experiment, listeners had to match a comparison interval with a fixed level to a reference interval in which the level of the two tones could increase or decrease (see Methods).

As the pitch of low-frequency pure tones decreases when level increases, we expected that the reference interval would sound smaller when the level increased than when it was fixed. Fig. 4 shows the results of the interval-matching procedure for the three level conditions (solid line). Subjects were able to accurately match interval size when all four tones had the same level (fixed level, error: -0.08±0.05 semitone). The matched interval size decreased in the increasing level condition by about 0.4 semitone (2.3%) ($p<0.01$), and conversely increased in the decreasing level condition ($p<0.01$). These results conform to our expectations based on the pitches of isolated pure tones, although the effect was smaller than predicted (dashed line).

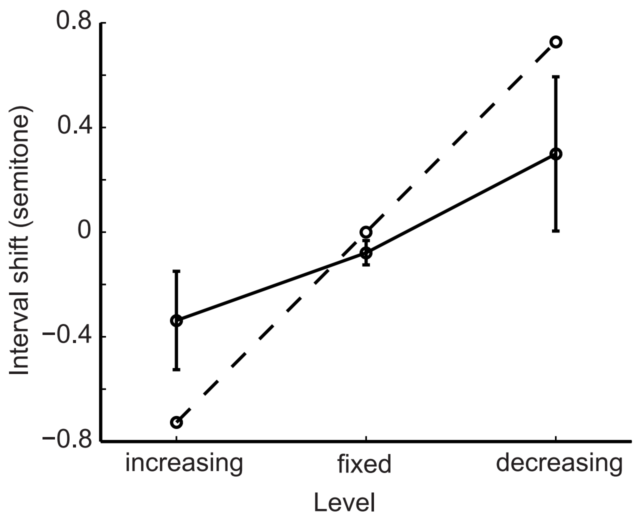

Figure 4. Level dependence of frequency adjustments required for interval matching, expressed as deviation from expected interval. Experimental results (solid line, error bars show ±1 SD across subjects) are compared to predictions based on pure tone pitch (dashed).

4. Discussion

Taken together, our results confirm that the pitch of low frequency pure tones (200 Hz) depends on level. Specifically, the pitch decreased by an amount equivalent to a change in frequency of more than half a semitone (3.5%) when the level was increased from 40 to 70 dB SPL. This is of course small compared to the effect of repetition rate on pitch, especially given the very large change in level, but it is significant and systematic. We found no dependence on level for high frequency tones (4000 Hz) or complex tones. The analysis for many different pairs of comparison and reference levels indicates that the frequency adjustments in the matching task reflect changes in perceived pitch as opposed to contextual or procedural effects. Additionally, similar, although smaller, pitch changes were measured in an interval matching task. Finally, we found that musicians were more precise than non-musicians in pitch matching tasks but showed the same pitch shifts with level.

Overall, our results confirm and complement previous studies (Fletcher, 1934; Stevens, 1935; Snow, 1936; Morgan et al., 1951; Terhardt, 1974b; Verschuure and van Meeteren, 1975) with a more exhaustive and controlled analysis. A number of studies had already shown the effect of level on low frequency pure tones, a few of them with a pitch matching procedure. We have replicated these findings. Our analysis of pitch matching with many different pairs of comparison and reference levels (Experiment 2) confirms that the observed changes in frequency deviation have the expected properties of pitch changes (e.g. antisymmetry, additivity); a similar analysis was performed in one study by Verschuure and van Meeteren (1975) with two subjects. We have complemented this analysis by using our interval matching task, which showed effects of level in the expected direction, but of smaller magnitude; it is not clear why level had less influence on interval perception than expected from the results of the pitch matching experiments (Fig. 4). An absence of an effect of level on the pitch of complex tones was mentioned by Snow (1936) but not shown. Terhardt (1975) reported small pitch shifts with level for complex tones, but the shifts varied markedly across participants and the mean shift was probably not statistically significant. In particular, a small negative pitch shift was reported for complex tones with fundamental frequency 200 Hz and harmonics 1-5, similar to one of the stimuli we used (harmonics 1-6). This shift was interpreted as consistent with pitch shifts of the individual harmonics, although it did not appear to be statistically significant. We applied the same procedure to complex tones with the same fundamental frequency, and we observed no significant pitch shifts, even though pitch shifts were observed with the with pure tones of 200 and 1000 Hz. Finally, it is known that musical experience can influence psychophysical performance in pitch tasks (for example discrimination), but previous studies did not mention the musical experience of listeners. We have analyzed the data for musicians and non-musicians separately, and we found that even though musicians were more precise in pitch matching, the two groups showed the same average dependence of pitch on level. We conclude that the level-dependence of the pitch of low frequency tones is genuine and independent of musical experience.

We found that musical training has an effect on the precision of pitch matching, but not on the average effect of level. This observation reinforces the claim that the pitch of low frequency pure tones truly depends on level by ruling out one alternative interpretation of the experimental observations, namely that the level-dependence of the deviation in match reflects the level-dependence of uncertainty in the task. That is, one could argue that pitch is more uncertain at low levels and therefore listeners are biased towards some prior, which could reflect for example the spectrum of natural sounds. This could predict level-dependent responses, with larger shifts for frequencies far from the prior. If this were true, we would expect smaller level-dependent shifts when responses are more precise, but this was not the case for musicians.

The results confirm that there is no exact one-to-one mapping between repetition rate and pitch, even for pure tones. This finding is related to other surprising phenomena in pitch perception (Hartmann, 1998). Diplacusis, which exists in normal hearing subjects, is the phenomenon that the pitch of a tone can differ slightly but significantly between the two ears (Jeffress, 1944; van den Brink, 1975; Burns, 1982). Another known phenomenon is octave stretch, the fact that when subjects are asked to adjust the frequency of a tone so that it sounds an octave higher than a reference tone of frequency f, they tend to set a frequency greater than 2f. This phenomenon is also observed for listeners with absolute pitch (Ward, 1954). Pitch shifts can also be induced by masking noise (Houtsma, 1981), by a previous adapting tone of similar frequency (Larkin, 1978; Rakowski and Hirsh, 1980), or by changes in the envelope of the tone (Rossing and Houtsma, 1986; Hartmann, 1978).

The level dependence of pitch is rather difficult to explain by standard temporal (Licklider, 1951) or place (Terhardt, 1974a) theories of pitch. The implications of this phenomenon on theories of pitch have been discussed previously (Hartmann, 2004; Moore, 2012), and we therefore only give a general overview. Temporal theories of pitch based on the periodicity of the sound's waveform, predict that pitch is independent of sound level. Significantly, our results show that the pitch of low frequency tones, but not of complex tones with the same periodicity, depends on level; it is not straightforward to see how temporal theories of pitch could predict differences between these two cases. It has been shown that the most frequent all-order interspike interval in auditory nerve recordings corresponds to the pitch of complex tones, and varies little with level (Cariani and Delgutte, 1996), in contrast with estimates from first-order intervals (successive spikes). A possible explanation for the level dependence of the pitch of low frequency pure tones is that the estimate from interspike intervals can deviate from the stimulus period because of the refractory period of auditory nerve fibers (Ohgushi, 1983). At higher levels, this deviation could become more important as the firing rate increases. However, at low frequencies, this effect is only seen in first-order intervals but not in all-order intervals (McKinney and Delgutte, 1999).

Place theories of pitch in which pitch is indicated by the place of maximal activation of the cochlea, suffer from the fact that auditory nerve fibers saturate at high levels and the locus of maximal activation varies dramatically with level (Sachs and Young, 1979; Kim, 1980; Chatterjee and Zwislocki, 1997; Cedolin and Delgutte, 2005; Versteegh et al., 2011), by an amount equivalent to about 0.5 octave for a level change from 40 to 80 dB (McFadden, 1986; Zwislocki and Nguyen, 1999; Moore et al., 2002). It has been proposed that pitch could be indicated by the low frequency edge of the activation pattern (Zwislocki and Nguyen, 2009), but the predicted variation of pitch with level far exceeds that measured psychophysically (Temchin and Ruggero, 2014).

In template theories of pitch (Terhardt, 1974a), it is postulated that the cochlear activation profile associated with each pitch is learned, which could potentially include level-dependent effects. In this context, it could be argued that the pitch of low-frequency tones, but not of complex tones, shows level-dependence because the auditory system is mostly exposed to complex tones (in particular voices) while pure tones are less natural. In this conceptual framework, the observation that musical training enhances precision but has no effect on the average level-dependence of the pitch of low frequency tones is not straightforward to interpret. Indeed, what underlies the enhanced precision of musicians? A natural explanation in the framework of template theory is that increased exposure to templates yields better discrimination via learning. In this case, we would expect that a reduction in bias accompanies an enhancement of precision, but this was not the case. Thus if template theory is correct, then the learning that underlies the level independence of the pitch of complex tones must be of a different type than the learning that underlies the enhancement in precision in musicians.

A few models of pitch (Loeb et al., 1983; Cedolin and Delgutte, 2010; Laudanski et al., 2014) and of tone detection (Carney et al., 2002) are based on timing differences between the spike trains in different fibers. For example, Loeb's model proposes that the frequency of a pure tone can be estimated by comparing signals across the basilar membrane: the distance that separates places that vibrate in phase is related to the tone's frequency. Because the phase of the basilar membrane response to tones depends on level (Robles and Ruggero, 2001), pitch would depend on level according to this model. However, a quantitative study using guinea pig auditory nerve responses (Carlyon et al., 2012) showed that in Loeb's model (Loeb et al., 1983), the predicted variation of pitch with level far exceeds that measured by psychophysical measurements: the change corresponds to more than two octaves for a 40 dB level change.

The structural theory of pitch (Laudanski et al., 2014) is a generalization of the models of Licklider and Loeb (Licklider, 1951; Loeb et al., 1983), in which pitch is postulated to be the perceptual correlate of the regularity structure of the vibration pattern on the basilar membrane, across place and time. Regularity structure at the level of cochlea includes periodicity within a cochlear channel (mathematically, $S(x,t) = S(x,t-\delta)$ for all t, where $S(x,t)$ is the mechanical signal at position x on the basilar membrane), named within-channel structure, and identities across frequency channels of the form $S(x,t) = S(y,t-\delta)$ for all t (x and y are two fixed cochlear places), named cross-channel structure. Within-channel structure is level-independent, but cross-channel structure is level-dependent if the basilar membrane responds nonlinearly. As in Licklider's model, delays δ in the regularity structure are matched by conduction (e.g. axonal) delays. If there is a physiological upper bound on these delays ($\delta<\delta_{max}$), then low frequency pure tones have only cross-channel structure (more precisely, within-channel structure is not usable), while complex tones also have within-channel structure (Laudanski et al., 2014). Thus, the theory predicts level dependence for low frequency pure tones but not complex tones if it is postulated that within-channel structure dominates cross-channel structure. The extent of level dependence depends on the spacing between channels and therefore on the specific instantiation of the theory: if channels are widely spaced, then the model would suffer from the same problem as Loeb's model (Carlyon et al., 2012). The model produces weak level dependence for narrowly spaced channels, for tone frequencies lower than $1/\delta_{max}$.

In conclusion, the pitch of low frequency tones depends weakly but significantly on level, and this finding poses an interesting and unresolved challenge for theories of pitch.

Acknowledgments

This work was supported by the European Research Council (ERC StG 240132). We thank the reviewers and the editor for their help.

References

Attneave F, Olson RK (1971) Pitch as a medium: a new approach to psychophysical scaling. Am J Psychol 84:147–166.

Bernstein JG, Oxenham AJ (2003) Pitch discrimination of diotic and dichotic tone complexes: harmonic resolvability or harmonic number? J Acoust Soc Am 113:3323-3334.

Burns EM (1982) Pure-tone anomalies. I. Pitch-intensity effects and diplacusis in normal ears. J Acoust Soc Am 72:1394-1402.

Burns EM, Ward WD (1978) Categorical perception--phenomenon or epiphenomenon: evidence from experiments in the perception of melodic musical intervals. J Acoust Soc Am 63:456-468.

Burns EM, Campbell SL (1994) Frequency and frequency-ratio resolution by possessors of absolute and relative pitch: examples of categorical perception. J Acoust Soc Am 96:2704-2719.

Cariani PA, Delgutte B (1996) Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. J Neurophysiol 76:1698.

Carlyon RP, Long CJ, Micheyl C (2012) Across-channel timing differences as a potential code for the frequency of pure tones. J Assoc Res Otolaryngol 13:159-171.

Carney L, Heinz MG, Evilsizer ME, Gilkey RH, Colburn HS (2002) Auditory phase opponency: A temporal model for masked detection at low frequencies. Acta Acustica 88:334-347.

Cedolin L, Delgutte B (2005) Pitch of complex tones: rate-place and interspike interval representations in the auditory nerve. J Neurophysiol 94:347-362.

Cedolin L, Delgutte B (2010) Spatiotemporal representation of the pitch of harmonic complex tones in the auditory nerve. J Neurosci 30:12712-12724.

Chambers C, Pressnitzer D (2014) Perceptual hysteresis in the judgment of auditory pitch shift. Atten Percept Psychophys 76:1271-1279.

Chatterjee M, Zwislocki JJ (1997) Cochlear mechanisms of frequency and intensity coding. I. The place code for pitch. Hear Res 111:65-75.

de Cheveigne A (2005) Pitch perception models. In: Springer handbook of auditory research, vol 24 Springer, New York, pp 169–233.
Fletcher H (1934) Loudness, pitch and the timbre of musical tones and their relation to the intensity, the frequency and the overtone structure. J Acoust SocAm 6:59-69.
Hartmann WM (1978) The effect of amplitude envelope on the pitch of sine wave tones. The Journal of the Acoustical Society of America 63:1105–1113.

Hartmann WM, Doty SL (1996) On the pitches of the components of a complex tone. J Acoust Soc Am 99:567-578.
Hartmann WM (2004) Signals, Sound, and Sensation. Springer Science & Business Media, New York.

Houtsma AJM (1981) Noise-induced shifts in the pitch of pure and complex tones. The Journal of the Acoustical Society of America 70:1661 – 1668.

Jeffress LA (1944) Variations in Pitch. Am J Psychol 57:63-76.
Kim DO (1980) Cochlear mechanics: implications of electrophysiological and acoustical observations. Hear Res 2:297-317.
Kishon-Rabin L, Amir O, Vexler Y, Zaltz Y (2001) Pitch discrimination: are professional musicians better than non-musicians? J Basic Clin Physiol Pharmacol 12:125-143.
Larkin WD (1978) Pitch Shifts Following Tone Adaptation. Acta Acustica united with Acustica 41:110–116.

Laudanski J, Zheng Y, Brette R (2014) A structural theory of pitch. eneuro:ENEURO 0033-14.2014.
Licklider JC (1951) A duplex theory of pitch perception. Experientia 7:128-134.
Loeb GE, White MW, Merzenich MM (1983) Spatial cross-correlation. A proposed mechanism for acoustic pitch perception. Biol Cybern 47:149-163.
McDermott JH, Keebler MV, Micheyl C, Oxenham AJ (2010) Musical intervals and relative pitch: frequency resolution, not interval resolution, is special. J Acoust Soc Am 128:1943-1951.
McFadden, D., 1986. The curious half octave shift: Evidence for a basalward migration of the travelling-wave envelope with increasing intensity. In: Salvi, R. J., Henderson, D., Hamernik, R. P., Colletti, V. (Eds.), Basic and Applied Aspects of Noise-Induced Hearing Loss, Plenum, New York, pp. 295-312.
McKinney MF, Delgutte B (1999) A possible neurophysiological basis of the octave enlargement effect. The Journal of the Acoustical Society of America 106:2679–2692.

Moore, B. C. J., Alcántara, J. I., Glasberg, B. R., 2002. Behavioural measurement of level-dependent shifts in the vibration pattern on the basilar membrane. Hear. Res. 163, 101-110.

Moore BCJ (2012) An Introduction to the Psychology of Hearing. BRILL, The Netherlands.

Morgan CT, Garner WR, Galambos R (1951) Pitch and intensity. J Acoust Soc Am 23:658-663.
Ohgushi K (1983) The origin of tonality and a possible explanation of the octave enlargement phenomenon. The Journal of the Acoustical Society of America 73:1694–1700.

Oxenham AJ (2012) Pitch perception. J Neurosci 32:13335-13338.
Oxenham AJ, Micheyl C, Keebler MV, Loper A, Santurette S (2011) Pitch perception beyond the traditional existence region of pitch. Proceedings of the National Academy of Sciences of the United States of America 108:7629-7634.

Rakowski A, Hirsh IJ (1980) Poststimulatory pitch shifts for pure tones. The Journal of the Acoustical Society of America 68:467–474.

Robles L, Ruggero MA (2001) Mechanics of the Mammalian Cochlea. Physiol Rev 81:1305-1352.

Rossing TD, Houtsma AJM (1986) Effects of signal envelope on the pitch of short sinusoidal tones. The Journal of the Acoustical Society of America 79:1926–1933.

Sachs MB, Young ED (1979) Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate. J Acoust Soc Am 66:470-479.

Shamma S, Klein D (2000) The case of the missing pitch templates: how harmonic templates emerge in the early auditory system. J Acoust Soc Am 107:2631-2644.

Snow WB (1936) Change of Pitch with Loudness at Low Frequencies. J Acoust Soc Am 8:14-19.

Stevens SS (1935) The relation of pitch to intensity. J Acoust Soc Am:150-154.

Temchin AN, Ruggero MA (2014) Spatial irregularities of sensitivity along the organ of Corti of the cochlea. J Neurosci 34:11349–11354.

Terhardt E (1974a) Pitch, consonance, and harmony. J Acoust Soc Am 55:1061-1069.

Terhardt E (1974b) Pitch of pure tones: Its relation to intensity. In E Zwicker & E Terhardt (Eds),Facts and models in hearing (pp 353-360) Heidelberg: Springer-Verlag.

Terhardt E (1975) The influence of intensity on the pitch of complex tones. Acustica 33, 344-348.

van den Brink G (1975) The Relation between Binaural Diplacusis for Pure Tones and for Complex Sounds under Normal Conditions and with Induced Monaural Pitch Shift. Acta Acustica united with Acustica 32:159–165.

Verschuure J, van Meeteren AA (1975) The effect of intensity on pitch. Acustica 32:33-44.

Versteegh CP, Meenderink SW, van der Heijden M (2011) Response characteristics in the apex of the gerbil cochlea studied through auditory nerve recordings. J Assoc Res Otolaryngol 12:301-316.

Ward WD (1954) Subjective Musical Pitch. J Acoust Soc Am 26:369-380.

Zwislocki JJ, Nguyen M (1999) Place code for pitch: a necessary revision. Acta Otolaryngol 119:140–145.