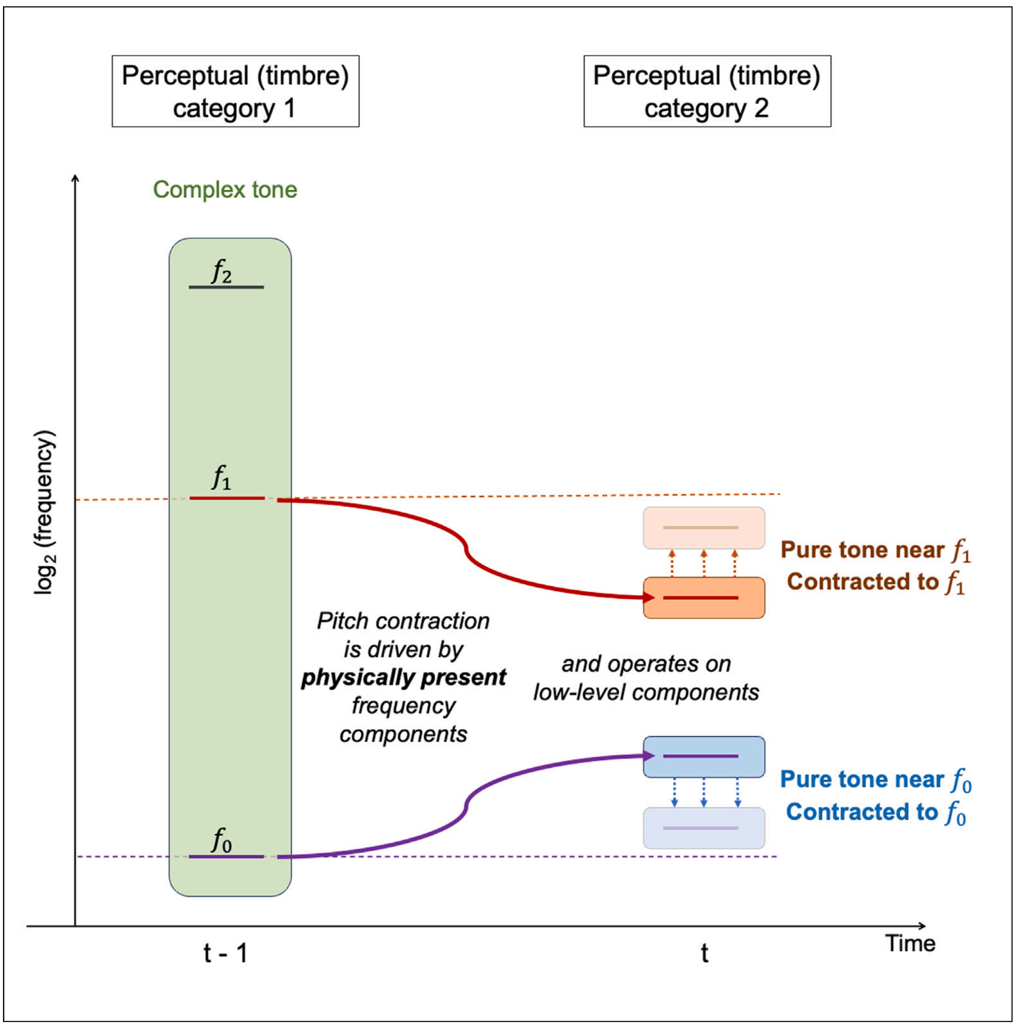
iScience

CellPress
OPEN ACCESS

Article

Frequency-specific contributions to auditory perceptual priors: Testing the predictive-coding hypothesis



Itay Lieder, Aviel Sulem, Merav Ahissar

msmerava@gmail.com

Highlights

Recent history has a large impact on current perception

Low-level representations of recent stimuli form local priors

Local priors operate bottom-up on low-level features of the current stimuli

Contraction driven by local priors contradicts the predictive-coding hypothesis

Lieder et al., iScience *27*, 108946
February 16, 2024
https://doi.org/10.1016/j.isci.2024.108946

iScience

CellPress
OPEN ACCESS

Article

Frequency-specific contributions to auditory perceptual priors: Testing the predictive-coding hypothesis

Itay Lieder,[1,3] Aviel Sulem,[1,3] and Merav Ahissar[1,2,4,*]

SUMMARY

Perceptual priors formed by recent stimuli bias our immediate percept. These priors, expressing our implicit expectations, affect both high- and low-level processing stages. Yet, the nature of the inter-level interaction is unknown. Do priors operate top-down and bias low-level features toward recently experienced objects (predictive-coding hypothesis), or are low-level biases bottom-up driven and formed by local memory circuits? To decipher between these options in auditory perception, we used the "missing fundamental illusion", enabling the dissociation of low-level components from the high-level pitch. Surprisingly, in contrast to predictive coding, when the fundamental frequency was missing, pitch contraction across timbre categories was not found to the previously perceived high-level pitch, but to the physically present frequency. This bottom-up contribution of low-level memory components to perceptual priors, operating independently of recent high-level percepts, may stabilize the perceptual organization and underlie continuity between similar low-level features belonging to different object categories in the auditory modality.

INTRODUCTION

Both perception and action are systematically biased toward previous experiences, recent (termed *serial dependence*, see e.g., Manassi et al.[1]), and older ones. In this study, we focus on perceptual bias driven by recent stimuli (*contraction bias*). Contraction bias is the tendency to perceive the current stimuli as more similar to the recent ones. In their seminal paper on perceptual bias driven by recent stimulus history, Fischer & Whitney[2] suggested that this phenomenon reflects a mechanism for creating a *continuity field*, where the brain biases information to tie together similar things that occur within a short time period, in order to smooth our experience and enhance the stability of our internal representations of the environment.[1,3–6]

During the last decade, a growing body of studies suggests that perceptual priors occur in continuity fields at many levels along the hierarchical brain processing, from the earliest[7] to the highest levels.[8,9] Specifically, both low and high sensory processing levels contribute to perceptual priors.[10] For instance, John-Saaltink et al.[11] demonstrated low-level contributions by showing that BOLD activity in the primary visual cortex (V1) is affected by the orientation of previous stimuli in the same position, even when no response was requested. On the other hand, several studies found contributions of high, object-level perception. For example, the perceived facial identity is strongly biased toward the identity of recently viewed faces.[12] Even in the context of basic features, like orientation, there is evidence for bias at the abstract level. The orientation of Gabor stimuli biases (contracts) the orientation reproduction of dot patterns, although the low-level features are not shared.[13] Furthermore, higher-level, post-perceptual contributions, involving decision-making[14] and working memory[15] have also been shown.

Although the behavioral signature of perceptual priors has been widely investigated, the neural mechanisms underlying continuity fields are still elusive.[16] Specifically, the nature of the interaction between the contributions of low and high levels to perceptual priors remains unclear. A dominant computational approach considers perception as a Bayesian inference that results from integrating responses to current sensory stimulation with prior expectations based on past experiences.[17,18] According to Bayesian theory, the brain constantly uses information from previous experiences to overcome ambiguous or noisy incoming sensory inputs, thus using the past to improve the reliability of predicting the future. Indeed, as the world tends to remain stable, at least over short time windows, a good prediction for the near future is that it will be similar to the present.

A leading theory, which proposes a neural implementation to the Bayesian approach is *predictive coding*. It suggests that predictive models are created at high perceptual levels and are funneled to lower-level sensory areas through feedback connections. Feedforward connections carry only the prediction error which is the difference between the top-down prediction and the actual lower-level activities.[19–21] It follows that the contraction bias toward recently presented low-level features is top-down determined.[22]

[1]The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel
[2]Department of Psychology, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel
[3]These authors contributed equally
[4]Lead contact
*Correspondence: msmerava@gmail.com
https://doi.org/10.1016/j.isci.2024.108946

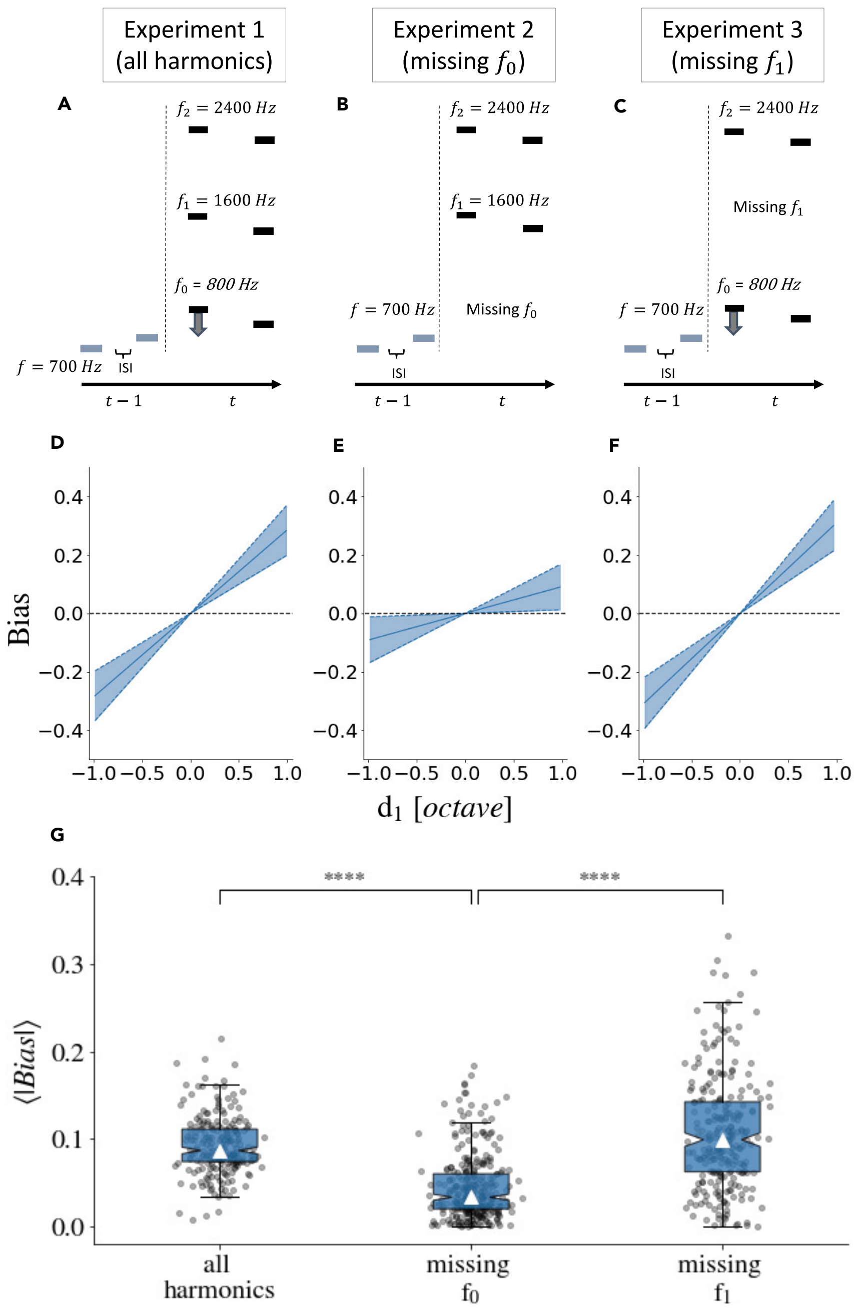


Figure 1. Contraction of the perceived pitch of complex tones by preceding simple-tone trials (simple tone frequency within 500-1000Hz) when the fundamental frequency (f_0) is physically present (A, D, C, F) versus when it is missing (B, E) – There is no contraction of complex tones' pitch when the fundamental is missing

(A–C) Schematic illustrations of simple-tone trials at time $t-1$, followed by a complex-tone trial at time t. The arrows show the direction of the bias, i.e., contraction of the fundamental frequency of the first complex tone in trial t to the first simple tone in trial $t-1$. (A) Complex tones are composed of the three first harmonics, f_0, f_1 and f_2 (Experiment 1). (B) Complex tones include only f_1 and f_2, and miss f_0 (Experiment 2). (C) Complex tones include only f_0 and f_2, and miss f_1 (Experiment 3).

(D–F) Contraction bias magnitudes measured as a function of the frequency distance $d_1 = f_0^t - f^{t-1}$. Using a GAM model (STAR Methods), the magnitude of the bias was assessed on the aggregated data across participants. Error bars indicate standard error. Black horizontal dashed line denotes zero bias.

(G) By adding random effects to the GAM model (STAR Methods), bias magnitudes were assessed at the level of individual participants. The boxplot shows the comparison of the bias magnitudes within individual participants when: left – all frequencies are physically present; middle - f_0 is missing; and right - f_1 is missing. The bias magnitude was calculated for each participant using as a metric the mean of the absolute value of the bias's magnitude. White triangles indicate the median. Error bars show lower to upper quartile values of the data. Most participants showed contraction (the slopes of the bias function were significantly

Figure 1. *Continued*

positive): 227/230 participants in Experiment 1 (W = 8; p <0 .001, Wilcoxon signed-rank), 232/253 in Experiment 3 (W = 942; p <0 .001, Wilcoxon signed-rank), and 239/316 participants when f_0 was missing (Experiment 2), (W = 9288, p <0 .001, Wilcoxon signed-rank). Importantly, the magnitude of the bias is very close to zero (median = 0.033) when the fundamental is missing, and much larger when the fundamental is physically present (median = 0.087 and 0.099, Experiments 1 and 3, respectively). The analysis of the aggregate (D, E, F) and individual data (G) showed that the pitch of complex tones is not contracted toward the preceding simple tones when the fundamental is missing. Furthermore, when the fundamental is physically present, the contraction of the pitch of complex tones is substantially and significantly stronger. **** $p < .001$.

An alternative account that may explain the neuronal mechanism underlying the contraction to low-level features of recent stimuli would be that the effect is local and perhaps driven at each level, including low levels, based on its memory circuits.[23] According to this account, the contraction bias retains recently encountered low-level sensory inputs, even when they are (atypically) inconsistent with high-level perception.

In this study, we tested these two hypotheses asking whether the neural mechanism underlying low-level contributions reflects high-level priors or local memory circuitry. We used the hierarchical relationship between frequency bands (low-level representations) and the perceived pitch (high-level representation) in the auditory modality. Namely, along the fast-ascending pathway from the inner ear to the primary auditory cortex, the processing is largely frequency-specific, whereas later object-level stages are spectrally broad.[24,25] Pitch perception is a high-level feature determined by the fundamental frequency of the stimuli, even when it is physically missing.[26] Consequently, when the fundamental is missing, there is an inconsistency between low-level frequency-specific representations which have no fundamental, and high-level pitch perception which remains invariant. Thus, if contraction is driven by high-level pitch-based predictions, as follows from predictive coding, contraction would be to the perceived pitch even when the fundamental is missing. By contrast, if contraction is based on local memory networks and is not top-down driven, we would expect contraction to occur only toward the specific frequencies that are physically present, even when they differ from the frequency that determines the pitch.

We administered a serial 2-tone pitch discrimination task using trials with either simple or complex tones (see STAR Methods). Simple and complex tones belong to different categories of timbre, yet they have a shared feature – pitch. Each trial was composed of one timbre category (simple or complex), but the two types of trials were presented in random order. In each trial, two tones (220ms) were presented serially, with a silent interstimulus interval (800ms). Participants were asked to determine which of the two tones had a higher pitch. Importantly, we conducted the same task in three independent experiments, using complex tones with different components: In experiment 1, complex tones were composed of the first three harmonics; in experiment 2, they had a missing fundamental; in experiment 3, they had a missing second harmonic. The complex tones in the three experiments were expected to differ only at the level of their spectra and not at the level of their abstract pitch representations.

Mixing simple- and complex-tone trials yields four types of consecutive trials: simple → simple, complex → complex, simple → complex, and complex → simple. The first two are composed of the same timbre category, while the last two are composed of different timbre categories. According to the continuity-field account[4] and to the feature tuning property of serial dependence,[1] the first two types of trials should manifest larger contraction than the third and fourth types since they are composed of similar tones and share the same timbre category. This was indeed the case (Supplemental information Figure S1). However, only the pairs of trials composed of simple and complex tones (across categories) allow us to dissociate between top-down and bottom-up contributions when testing pitch contraction to low-level priors. Simple trials preceding complex trials allow us to dissociate whether the prior formed by simple tones operates on the low-level frequency channels of the complex tones or on the integrated, abstract, perceived pitch (Figures 1A–1C). Complex trials preceding simple trials allow us to dissociate whether simple tones are contracted toward the fundamental frequency or toward the pitch of the complex tones (Figures 2A–2C).

Calculating contraction between simple-tone and complex-tone trials in the three experiments described above, we found that contraction operates on early sensory representations: complex tones with a missing fundamental were not contracted toward simple tones whose frequency was near the fundamental. Crucially, we found no contraction driven by complex tones with a missing fundamental to the fundamental frequency. Rather, we measured contraction to the physically present frequency bands, even though they clearly differed from the perceived pitch. In particular, we showed that simple tones near the complex tones' second harmonic, were contracted toward the second harmonic when the latter was physically present (Experiments 1 and 2). Namely, the contraction is not driven by the abstract pitch, but rather by present neighboring frequency components. These results are not in line with the predictive coding hypothesis which postulates that predictions are generated at high perceptual levels and operate top-down on low-level representations of current stimuli. Based on these results, we propose that low-level contributions to perceptual priors operate through bottom-up pathways, independently of recent high-level percepts.

RESULTS

Perceptual history operates on low-level components

Predictive coding assumes that predictions formed from statistical regularities of recent stimuli are tested against incoming sensory information. We thus first tested whether perceptual history propagates down to low-level representations of pitch, namely whether contraction operates directly on frequency-specific representations. For this aim, we considered simple → complex trial pairs. We analyzed contraction by the simple tone under three experiments: complex tones with three harmonics (Experiment 1, Figure 1A), with a missing fundamental (Experiment 2, Figure 1B), and with a missing second harmonic (Experiment 3, Figure 1C). In each experiment, we measured the contraction driven by the first simple tone of the previous trial (f^{t-1}) on the fundamental frequency of the first complex tone of the current trial (f_0^t). We calculated the magnitude of the biases as a function of the log frequency distance (d_1) between the preceding simple tone (f^{t-1}) and the current fundamental frequency (f_0^t): $d_1 = f_0^t - f^{t-1}$ (Figures 1D–1F, corresponding to Experiment 1, 2 and 3, respectively). Based on previous studies[27] and

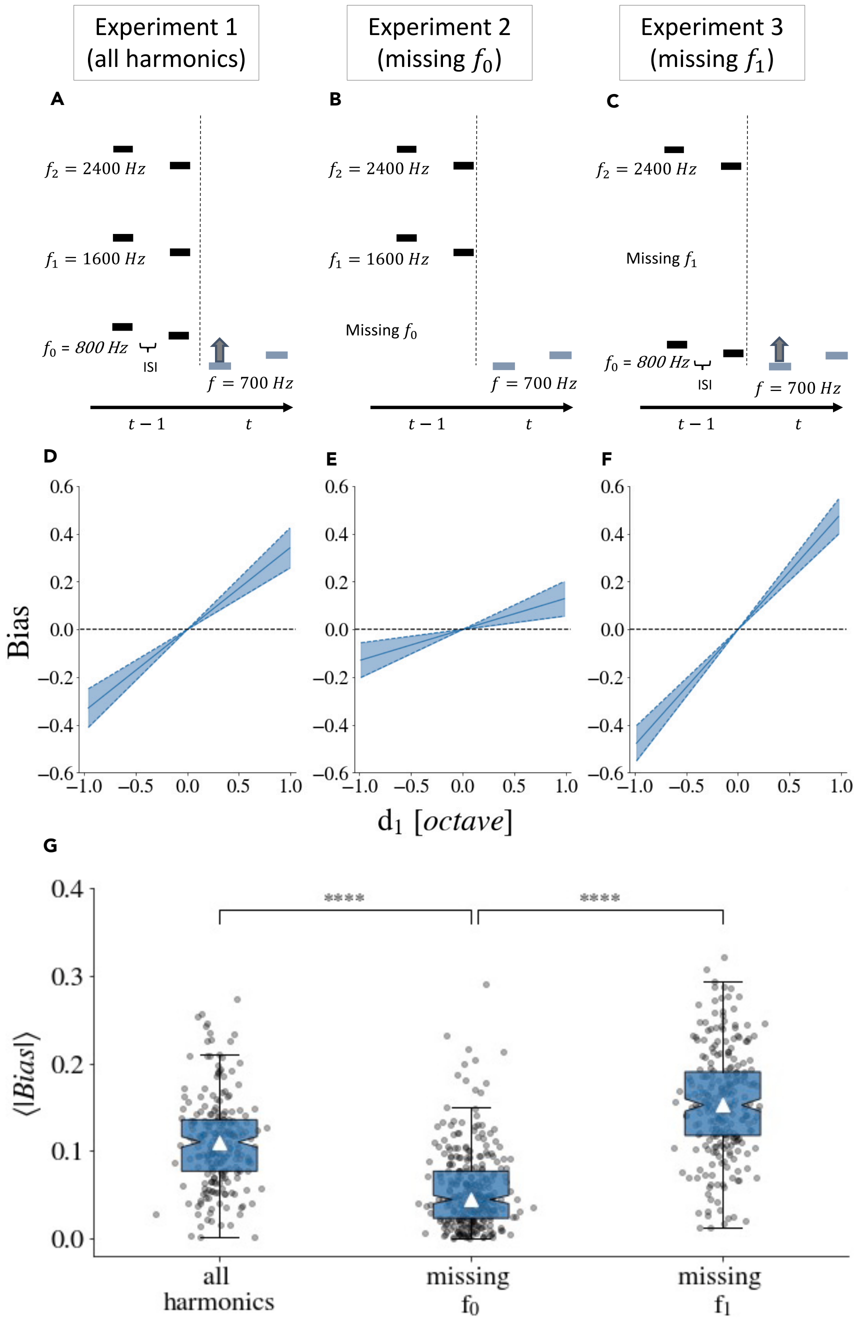


Figure 2. Pitch contraction of simple tones by preceding complex-tone trials (simple tone frequency within 500-1000Hz) when the fundamental frequency (f_0) is physically present (A, C, D, F) versus when it is missing (B , E) – there is no contraction by the pitch of the complex tone when f_0 is missing

(A–C) Schematic illustrations of complex-tone trials at time $t-1$, followed by simple-tone trials at time t, under the three experimental conditions. The arrows show the direction of the bias, i.e., contraction of the first simple tone in trial t to f_0 of the first complex tone in trial $t-1$. (A) Complex tones are composed of the first three harmonics, f_0, f_1 and f_2 (Experiment 1). (B) Complex tones miss the fundamental and include only f_1 and f_2 (Experiment 2). (C) Complex tones are composed of f_0 and f_2 and miss f_1 (Experiment 3).

(D–F) Contraction bias magnitudes measured as a function of the frequency distance $d_1 = f^t - f_0^{t-1}$. Using a GAM model (STAR Methods), the magnitude of the bias was assessed on the aggregated data across participants. Error bars indicate standard error. Black horizontal dashed line denotes zero bias.

(G) By adding random effects to the GAM model (STAR Methods), bias magnitudes were assessed at the level of individual participants. The boxplot shows the comparison of the bias magnitudes within individual participants when: left – all frequencies are physically present; middle – f_0 is missing; and right – f_1 is missing. The bias magnitude was calculated for each participant using as a metric the mean of the absolute value of the bias magnitude. White triangles indicate the

Figure 2. ***Continued***

median. Error bars show lower to upper quartile values of the data. Most participants showed contraction (the slopes of the bias function were significantly positive): 227/230 participants in Experiment 1 ($W = 11$, $p < .001$, Wilcoxon signed-rank), 250/253 in Experiment 3 ($W = 46$, $p < .001$, Wilcoxon signed-rank), and 254/316 when f_0 was missing (Experiment 2), ($W = 6583$, $p < .001$, Wilcoxon signed-rank). Importantly, the magnitude of the bias is very close to zero (median = 0.044) when the fundamental is missing, and much larger when the fundamental is physically present (median = 0.11 and 0.15, Experiments 1 and 3, respectively). The analysis of the aggregate (D, E, F) and individual data (G) showed that perceptual priors were substantially and significantly reduced when simple-tone trials followed complex-tone trials with a missing fundamental, suggesting that contraction is driven by f_0 and not by high-level pitch, which is not in agreement with the predictive coding hypothesis. **** $p < .001$.

our current analysis (Supplemental information, Figure S1) we included only simple tones that were within the same octave as the fundamentals of the complex tones.

If contraction acts directly on low-level representations, as suggested by predictive coding, we expect that complex tones with a missing fundamental will not be contracted. However, if contraction operates on high-level pitch representation, we expect to see no difference in the magnitude of contraction when the fundamental is physically present (Experiments 1 and 3) or missing (Experiment 2), since the high-level representation of pitch is invariant to the presence or absence of the fundamental.

We thus compared the contractions when the fundamental frequency was physically present to its magnitude when it was missing. In the missing fundamental condition, the slope of the bias was flat (Figure 1E), $slope = 0.07 \pm 0.07$ (SE) and not significantly different from zero ($p = .32$). Namely, there was no contraction of the pitch when the fundamental was missing. By contrast, when the fundamental was physically present, the slopes of the bias were substantially steeper and highly significant: $slope = 0.26 \pm 0.08$ (SE), $p < .001$ (Experiment 1, Figure 1D); $slope = 0.33 \pm 0.08$ (SE), $p < .001$ (Experiment 3, Figure 1F), and the predictor d_1 had significant contribution to the model ($\chi^2 = 58.5$, $edf = 41$, $p = .04$; $\chi^2 = 569.0$, $edf = 215$, $p < .001$, Wald test; Experiment 1 and Experiment 3, respectively), but not when f_0 was missing ($\chi^2 = 53.2$, $edf = 76$, $p = .98$, Wald test; Experiment 2).

Analyses at the level of individual participants yielded the same results (Figure 1G shows the mean of the absolute value of the participants' bias magnitude, STAR Methods). Though most participants showed contraction (the slopes were significantly positive) in all three experiments, the magnitude of the bias substantially differed between these conditions ($H = 246.8, p < .001$, Kruskal-Wallis). Post-Hoc tests (Bonferroni correction, Dunn's test) showed that bias was larger when f_0, f_1 and f_2 were physically present (Experiment 1) compared to when f_0 was missing (Experiment 2): ($Z = 12.9, p < .001$). Similarly, the bias was significantly larger when f_0 and f_2 were physically present (Experiment 3) compared to when f_0 was missing (Experiment 2): ($Z = 13.7, p < .001$).

These results confirm that perceptual priors operate on low-level representations, in line with predictive coding.

No contraction to the perceived pitch when the fundamental is missing

A crucial aspect of predictive coding is that perceptual expectations are generated at high perceptual levels and interact with sensory inputs through top-down pathways. We thus tested whether contraction is driven by high-level representation (abstract pitch percept – invariant to missing harmonics), or in contrast, by low-level representation (fundamental frequency f_0 – sensitive to the spectral composition of the stimuli). For this purpose, we examined complex-tone trials preceding simple-tone trials under the three conditions: complex tones with 3 harmonics (Experiment 1, Figure 2A), with a missing fundamental (Experiment 2, Figure 2B), and with a missing second harmonic (Experiment 3, Figure 2C). For each condition, we concentrated on the contraction driven by the fundamental frequency (f_0) and operating on the frequency (f) of simple tones. We calculated the magnitude of the biases as a function of the frequency distance (d_1) between f_0 of the first complex tone in trial $t - 1$, and the f of the first simple tone in trial t ($d_1 = f^t - f_0^{t-1}$) (Figures 2D–2F corresponding to Experiment 1, 2 and 3, respectively). As in the previous analysis, we included only simple tones whose frequencies were in the range of an octave from the fundamental of the complex tones.

Following the same rationale as above, we expect to see no difference in bias magnitude if contraction is driven by high-level representation (pitch), as predicted by predictive coding. However, if contraction is driven by low-level representation (f_0), we expect to measure contraction only when f_0 is physically present.

We thus compared the contractions when f_0 was physically present (Experiments 1 and 3) to its magnitude when it was missing (Experiment 2). When the fundamental was missing, the slope of the bias was almost flat (Figure 2E), $slope = 0.13 \pm 0.08$ (SE), and non-significantly different from zero ($p = 0.095$). Namely, there was no contraction by the high-level pitch when the fundamental was missing. By contrast, when f_0 was physically present, the slopes were substantially steeper and highly significant: $slope = 0.37 \pm 0.09$ (SE), $p < .001$ (Experiment 1, Figure 2D); $slope = 0.53 \pm 0.08$ (SE), $p < .001$ (Experiment 3, Figure 2F). The predictor d_1 had a highly significant contribution to the model ($\chi^2 = 125$, $edf = 78$, $p < .001$; $\chi^2 = 189$, $edf = 103$, $p < .001$, Wald test; Experiment 1 and Experiment 3, respectively), though it was also significant when f_0 was missing ($\chi^2 = 159$, $edf = 113$, $p = .003$, Wald test; Experiment 2), but contraction was very small and close to zero.

These results were also evident at the level of individual participants (Figure 2G shows the mean of the absolute value of the magnitude of biases, STAR Methods). Though most participants showed contraction (the slopes of the bias were significantly positive) in all three experiments, the magnitude of the effect differed between these conditions ($H = 337.5, p < .001$, Kruskal-Wallis). Post-Hoc tests (Bonferroni correction, Dunn's test) showed that the bias was significantly larger when f_0, f_1 and f_2 were physically present (Experiment 1) compared to when f_0 was missing (Experiment 2): ($Z = 10.8, p < .001$). Similarly, the bias was significantly larger when f_0 and f_2 were physically present (Experiment 3) compared to when f_0 was missing (Experiment 2): ($Z = 18.1, p < .001$).

To summarize, we found that cross-category contraction is not driven by the high-level pitch representation, but rather by the fundamental frequency itself when physically present, which is inconsistent with the predictive coding hypothesis.

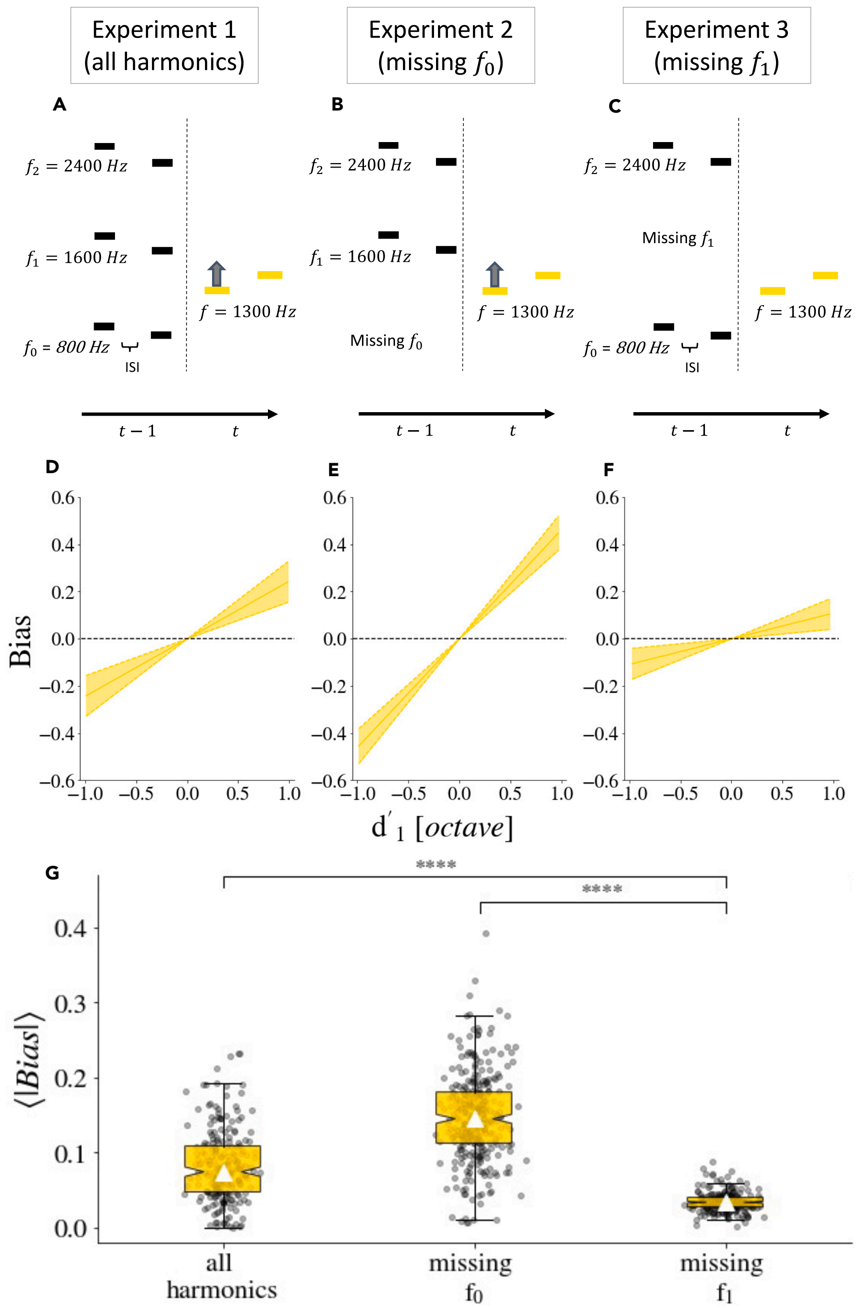


Figure 3. Contraction of the perceived pitch of simple tones (within 1000-2000Hz) driven by the second harmonic (f_1 - in the same frequency range) of preceding complex-tone trials when f_1 is present (A, B, D, E) versus when it is missing (C, F) – simple tones are contracted by f_1 when present, even though it differs from the perceived pitch (determined by f_0)

(A–C) Schematic illustrations of complex tone trials at time $t-1$, followed by a simple tone trial at time t in the same octave as f_1. The arrows show the direction of the bias, i.e., contraction of the first simple tone in trial t to the second harmonic of the complex tone in trial $t-1$. (A) Complex tones are composed of the first three harmonics, f_0, f_1 and f_2 (Experiment 1). (B) Complex tones include only f_1 and f_2, and miss f_0 (Experiment 2). (C) Complex tones include only f_0 and f_2, and miss f_1 (Experiment 3).

(D–F) Contraction bias magnitude measured as a function of the frequency distance $d'_1 = f^t - f_1^{t-1}$. Using a GAM model (STAR Methods), the magnitude of the bias was assessed on the aggregated data across participants. Error bars indicate standard error. Black horizontal dashed line denotes zero bias.

(G) By adding random effects to the GAM model (STAR Methods), bias magnitudes were assessed at the level of individual participants. The boxplot shows the comparison of the bias magnitudes within individual participants when: left – all 3 harmonics are physically present; middle – f_1 and f_2 are present, f_0 is missing; and right - f_0 and f_2 are present, f_1 is missing. The bias magnitude was calculated for each participant using as a metric the mean of the absolute value of the bias's magnitude. White triangles indicate the median. Error bars show lower to upper quartile values of the data. Most participants showed contraction (the slopes of

Figure 3. *Continued*

the bias function were significantly positive): 209/230 participants in Experiment 1 ($W = 853$, $p < .001$, Wilcoxon signed-rank), 313/316 in Experiment 2 ($W = 14$, $p < .001$, Wilcoxon signed-rank), and 252/253 in Experiment 3 ($W = 3.0$, $p < .001$, Wilcoxon signed-rank). Importantly, the magnitude of the bias is very close to zero (median = 0.034) when the second harmonic is missing, and much larger when the second harmonic is physically present (median = 0.074 and 0.14, Experiments 1 and 2, respectively). **** $p < .001$.

Contraction to physically present harmonics even when they differ from the perceived pitch

In the above section, we demonstrated that the fundamental frequency must be present for contraction to the pitch of that fundamental, which is inconsistent with predictive coding. We further ask whether the physical presence of energy in a specific frequency band is not only necessary but also sufficient, even if it differs from the percept. We reasoned that if contraction operates early, through bottom-up pathways, the physical presence of a frequency will induce contraction even when it is inconsistent with the globally perceived pitch. In particular, the analysis of contraction by the second harmonic (f_1), a harmonic that differs from the pitch, validates the interpretation of a bottom-up driven contraction by the harmonics that physically compose complex tones. This verification is important since the pitch of the missing fundamental is perceived by some listeners as higher than the fundamental,[28] and hence the lack of contraction to the fundamental frequency might be attributed to perceiving a higher pitch.

We focused on complex-tone trials preceding simple-tone trials. Importantly, the frequency (f) of the simple tones was in the same octave as the second harmonic (f_1) of the complex tones and thus, within their contraction range.[27] We measured the contraction of f of the first simple tone in the current trial, driven by f_1 of the first complex tone in the preceding trial. We calculated the magnitude of the bias as a function of the frequency distance d_1' to the second harmonic ($d_1' = f^t - f_1^{t-1}$) (Figures 3D–3F corresponding to Experiment 1, 2 and 3, respectively). We compared the bias magnitude when f_1 was physically present (Experiments 1 and 2, Figures 3A and 3B) to when f_1 was missing (Experiment 3, Figure 3C). Crucially, unlike f_0, f_1 does differ from the perceived pitch. Thus, measuring contraction of simple tone driven by f_1 when the latter is physically present would provide strong evidence that contraction is driven by low-level representations via feedforward pathways. If contraction was driven by percept, as predicted by predictive coding, we would expect that no contraction will be observed. Alternatively, if contraction is local, driven by low-level representations, we expect to see contraction when f_1 is present.

When f_1 was missing, the slope of the bias was very shallow (Figure 3F) and not significant: $slope = 0.08 \pm 0.07$ (SE), $p = .25$; namely, there was no contraction driven by f_1 when it was missing, as expected. However, when f_1 was present, the contraction was much stronger and significant: $slope = 0.29 \pm 0.09$ (SE), $p = .0017$ (Experiment 1, Figure 3D); $slope = 0.48 \pm 0.08$ (SE), $p < .001$ (Experiment 2, Figure 3E). Furthermore, the predictor d_1' had a significant contribution to the model when f_1 was present ($\chi^2 = 127$, $edf = 84$, $p = .0016$; $\chi^2 = 234$, $edf = 132$, $p < .001$, Wald test; Experiments 1 and 2, respectively), but not when f_1 was missing ($\chi^2 = 35$, $edf = 31$, $p = .32$, Wald test; Experiment 3).

Statistical analyses at the level of individual participants show the same results (Figure 3G displays the mean of absolute value of the biases, STAR Methods). As in the previous analyses, most participants showed contraction (the slopes were significantly positive) in all three experiments but the magnitude of the biases differed between these conditions ($H = 463.5, p < .001$, Kruskal-Wallis). Post-Hoc tests (Bonferroni correction, Dunn's test) showed that the bias was significantly larger when f_0, f_1 and f_2 were present (Experiment 1) compared to when f_1 was missing (Experiment 3): ($Z = 10.4, p < .001$). Similarly, the bias was significantly larger when f_1 and f_2 were physically present (Experiment 2) compared to when f_1 was missing (Experiment 3) missing: ($Z = 21.5, p < .001$).

To summarize, we found significant contraction of simple tones within the one-octave contraction range, driven by a physically present harmonic which is not associated with the perceived pitch. Thus, the existence of a harmonic is sufficient for contraction. This finding indicates that contraction operates locally at the level of specific frequency bands with energy, in a bottom-up manner, refuting the predictive coding top-down prediction.

DISCUSSION

The present study contributes to deciphering the mechanisms underlying low-level contributions to perceptual priors. Testing the predictive coding hypothesis, we found that simple tones form priors that operate on the low-level, frequency-specific representation of complex tones. This was evidenced by the observation that the pitch of complex tones was not contracted toward simple tones' pitch when the fundamental frequency was missing. Contraction occurred only when the fundamental was physically present. The observation that contraction operates directly on early levels of sensory analysis, before its integration into an abstract percept, is consistent with the predictive coding hypothesis.

However, in contrast to the predictive coding hypothesis, we found that contraction is not driven by high-level percept through top-down pathways, but rather is bottom-up, determined by the physically present frequency-specific bands with energy. This was evidenced by the observation that simple tones were not contracted toward complex tones when the fundamental was missing. Rather, simple tones were contracted toward the second harmonic, which differs from the pitch percept. Importantly, although the pitch illusion of the missing-fundamental is a well-established phenomenon,[26,29,30] some listeners preferentially hear the pitch of missing fundamental stimuli on the basis of the harmonics that are actually present rather than on the basis of the one missing.[28] Hence, additional independent support for the low-level source of contraction is important. We found that simple tones were contracted toward the second harmonic also in the case of complex tones with no missing fundamental. In this case, there is no ambiguity about the pitch of the complex tone, which is determined by the fundamental frequency and not by the second harmonic.

Our results are in line with the core idea of an inferential model developed by Chambers et al.,[31] based on continuity at the level of specific frequency bands and reflecting a temporal binding of successive frequency components. This model explains behavioral results of pitch judgments when the current stimulus is ambiguous. As an experimental paradigm that illustrates such a scenario, the authors used Shepard tones which are complex tones whose components are octave related. When the interval between two serially presented Shepard tones was half an octave (tritone), the participants responded with equal probability to the question of whether the second tone had a higher or lower pitch. Since the frequency distance between the two successive components was equal (half an octave), the pitch comparison was impossible. However, adding another Shepard tone, before the ambiguous trial, biased pitch perception, which they attributed to the contraction of the components of the first tone (in the ambiguous trial) toward the closest components of the preceding tone.

Two previous studies tried to systematically evaluate whether low-level contributions to perceptual priors are driven by top-down mechanisms or by local low-level memory components. The first, Cicchini et al.,[22] used a target grating patch surrounded by four flanker gratings, all tilted 15° away from the orientation of the target. The flankers negatively biased the perceived orientation of the target. They first asked whether the contraction driven by the previous trial operates on the physical or on the perceived orientation in the current trial and found that the contraction operates on the physical orientation, namely at early processing stages, in line with our results. Then they asked whether the contraction of the physical orientation in the current trial is driven by the physical or by the perceived orientation of the previous trial. Here, they found that contraction is driven by the perceived orientation and concluded that this result is consistent with predictive coding. This discrepancy with our results can be overcome by two different accounts: (1) As shown by John-Saaltink et al.,[11] the orientation of previous stimuli affects BOLD activity in the primary visual cortex (V1). The abstract representation of the orientation of the previous trial can be funneled through feedback pathways to V1. The contraction toward the perceived orientation observed in Cicchini et al.[22] does not exclude the possibility that lower-level representations, being affected by the perceived orientation, underly in fact this contraction. In the case of the missing fundamental, there is no energy in the frequency band of the fundamental and therefore no contraction. Note however a recent fMRI study, conducted by Sheehan and Serences,[32] that used a delayed orientation discrimination task, and found that early representations in primary visual cortex are repelled from the previous stimulus, while perceptual judgments are contracted to it. They suggested that when the task includes a working memory delay, serial dependence is driven by post-perceptual or mnemonic circuits and operates on the adapted representation of the current stimulus. Yet, this interpretation does not address the case where there are no extended working memory delay periods. (2) Another possible account is that there is an intrinsic difference between the visual and the auditory modalities. Unlike in the visual modality, stimuli, and deviations from regularities, are tracked automatically in the auditory modality. This modality difference is manifested in the automatic event-related potential (ERP) response to auditory deviants – mismatch negativity,[33] while the visual mismatch negativity is more sensitive to participants' attention.[34] The automaticity in the auditory modality can explain the contribution of local low-level memory components to perceptual priors which apparently may be not the case in vision.

Interestingly, a recent study which investigated perceptual priors in the oculomotor system, found that oculomotor behavior is biased by retinal error signals at early visual processing stages.[7] Specifically, using the Ponzo illusion, which corresponds to the fact that the same object, with a constant retinal size, is perceived as having different sizes depending on its depth level, they showed that the oculomotor system follows the retinal and not the perceived motion of the target. Namely, contraction is driven by low-level representations, in line with our results in the auditory modality. Note that both the oculomotor response to priors and the tracking of regularities and deviants in the auditory modality are highly automatic.

Goettker and Stewart[7] suggested that the mechanisms underlying perceptual priors are different in the motor and the visual systems. Priors operate top-down in the perceptual system,[22] but bottom-up in the oculomotor system.[7] Interestingly, as illustrated in Figure 4, we now show the existence of bottom-up contributions to perceptual priors in the auditory system. This suggests that, across perceptual categories (clearly distinct timbres), contraction between similar local features is driven by local memory components. Moreover, we observed that within perceptual categories (tones that share the same timbre), contraction is stronger, indicating top-down contributions. While top-down contributions to perceptual priors are in line with a predictive coding mechanism, the latter does not consider the bottom-up, low-level contributions to perceptual priors.

While previous stimuli have a substantial effect on the magnitude of the contraction bias, previous decisions are probably not crucial, ruling out a pure decision account.[14] Indeed, in line with Manassi et al.[1] and Goettker and Stewart,[7] perceptual priors occur independently of previous decisions. Had contraction been driven by decisions made in preceding trials, we should have observed contraction in all trial conditions, which was not the case.

In summary, perceptual priors involve both low- and high-level contributions, and probably more than one neural mechanism underlies them.[16] The present study reveals the important contribution of low-level memory components to perceptual priors, which is not captured by the predictive coding model. Neural retention processes operating through local memory networks such as neural adaptation[35] may underly the bias toward recently experienced low-level features.

Limitations of the study

The present study was conducted in controlled experimental conditions, using synthesized tones. This made it possible to decipher the mechanism underlying low-level contributions to perceptual priors. To further generalize the results, it will be useful to address this issue in more ecological environments, including natural stimuli such as musical instruments or human voices.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability

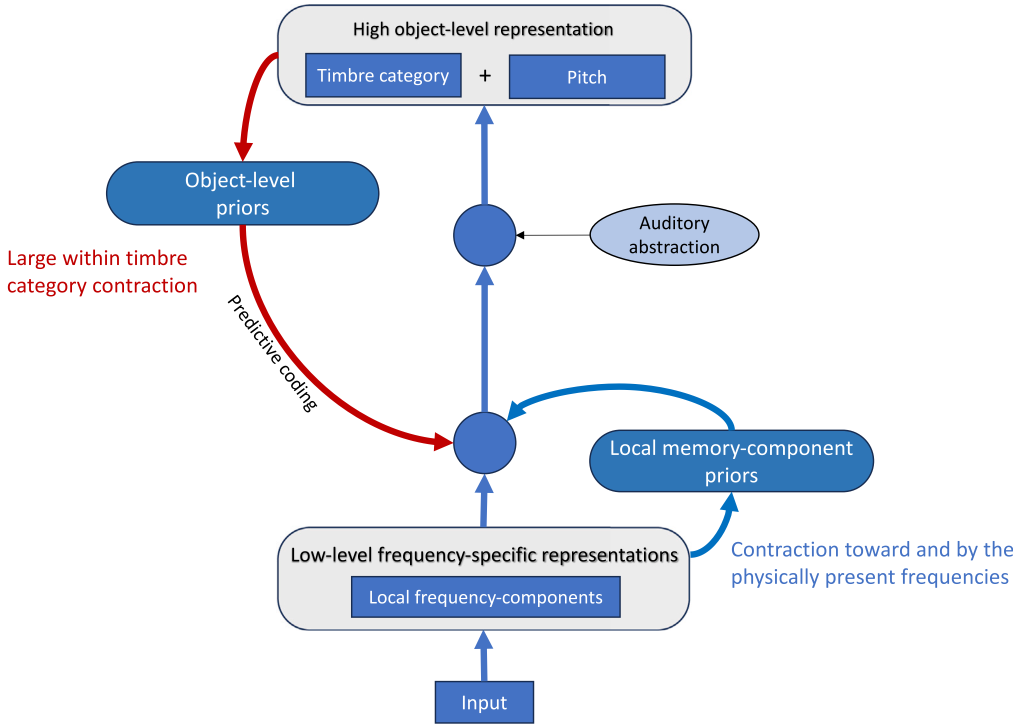


Figure 4. Top-down and bottom-up contributions to auditory perceptual priors

Low-level frequency-specific representations of recent stimuli form priors that operate bottom-up (blue) on low-level features of the current stimuli, independent of high-level predictions. We suggest that this mechanism underlies contraction between similar local features that belong to objects from different perceptual categories. We also found (Figure S1) that contraction is larger within timbre categories, indicating a top-down contribution (red), driven by object-level priors, which is in line with predictive coding. Our analyses cannot determine whether object-level priors operate on low-level representations (illustrated by the red arrow), or on higher object-level representations. The drawing of the general scheme was inspired by Cicchini et al.[22] and Goettker and Stewart.[7]

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.108946.

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 833694) and the Israel Science Foundation (Grant No. 1650/17), both awarded to M.A.

AUTHOR CONTRIBUTIONS

Initialization of the project: I.L.; Design of the experiments: I.L. and M.A.; Conceptualization of the study: I.L., A.S., and M.A. Collection of data: I.L.; Analysis of the data: I.L. and A.S.; Funding acquisition and supervision: M.A. All authors contributed to the interpretation of data and writing of the manuscript. These authors contributed equally: I.L. and A.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 18, 2023
Revised: December 2, 2023
Accepted: January 15, 2024
Published: January 18, 2024

REFERENCES

1. Manassi, M., Murai, Y., and Whitney, D. (2023). Serial dependence in visual perception: A meta-analysis and review. J. Vis. *23*, 18.
2. Fischer, J., and Whitney, D. (2014). Serial dependence in visual perception. Nat. Neurosci. *17*, 738–743.
3. Cicchini, G.M., Mikellidou, K., and Burr, D.C. (2018). The functional role of serial dependence. Proc. Biol. Sci. *285*, 20181722.
4. Liberman, A., Manassi, M., and Whitney, D. (2018). Serial dependence promotes the stability of perceived emotional expression depending on face similarity. Atten. Percept. Psychophys. *80*, 1461–1473.
5. Manassi, M., Liberman, A., Chaney, W., and Whitney, D. (2017). The perceived stability of scenes: Serial dependence in ensemble representations. Sci. Rep. *7*, 1971.
6. Manassi, M., and Whitney, D. (2022). Illusion of visual stability through active perceptual serial dependence. Sci. Adv. *8*, eabk2480.
7. Goettker, A., and Stewart, E.E.M. (2022). Serial dependence for oculomotor control depends on early sensory signals. Curr. Biol. *32*, 2956–2961.e3.
8. Kiyonaga, A., and Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. Psychon. Bull. Rev. *20*, 228–242.
9. Kiyonaga, A., Scimeca, J.M., Bliss, D.P., and Whitney, D. (2017). Serial Dependence across Perception, Attention, and Memory. Trends Cogn. Sci. *21*, 493–497.
10. Manassi, M., Liberman, A., Kosovicheva, A., Zhang, K., and Whitney, D. (2018). Serial dependence in position occurs at the time of perception. Psychon. Bull. Rev. *25*, 2245–2253.
11. St. John-Saaltink, E., Kok, P., Lau, H.C., and De Lange, F.P. (2016). Serial dependence in perceptual decisions is reflected in activity patterns in primary visual cortex. J. Neurosci. *36*, 6186–6192.
12. Liberman, A., Fischer, J., and Whitney, D. (2014). Serial dependence in the perception of faces. Curr. Biol. *24*, 2569–2574.
13. Ceylan, G., Herzog, M.H., and Pascucci, D. (2021). Serial dependence does not originate from low-level visual processing. Cognition *212*, 104709.
14. Pascucci, D., Mancuso, G., Santandrea, E., Della Libera, C., Plomp, G., and Chelazzi, L. (2019). Laws of concatenated perception: Vision goes for novelty, decisions for perseverance. PLoS Biol. *17*, e3000144.
15. Bliss, D.P., Sun, J.J., and D'Esposito, M. (2017). Serial dependence is absent at the time of perception but increases in visual working memory. Sci. Rep. *7*, 14739.
16. Whitney, D., Manassi, M., and Murai, Y. (2022). Searching for serial dependencies in the brain. PLoS Biol. *20*, e3001788.
17. De Lange, F.P., Heilbron, M., and Kok, P. (2018). How Do Expectations Shape Perception? Trends Cogn. Sci. *22*, 764–779.
18. Seriès, P., and Seitz, A.R. (2013). Learning what to expect (in visual perception). Front. Hum. Neurosci. *7*, 668.
19. Friston, K., and Kiebel, S. (2009). Predictive coding under the free-energy principle. Philos. Trans. R. Soc. Lond. B Biol. Sci. *364*, 1211–1221.
20. Rao, R.P., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. *2*, 79–87.
21. Trapp, S., Pascucci, D., and Chelazzi, L. (2021). Predictive brain: Addressing the level of representation by reviewing perceptual hysteresis. Cortex *141*, 535–540.
22. Cicchini, G.M., Benedetto, A., and Burr, D.C. (2021). Perceptual history propagates down to early levels of sensory analysis. Curr. Biol. *31*, 1245–1250.e2.
23. Pearson, J., and Brascamp, J. (2008). Sensory memory for ambiguous vision. Trends Cogn. Sci. *12*, 334–341.
24. Nelken, I., and Ahissar, M. (2006). High-level and low-level processing in the auditory system: the role of primary auditory cortex. In Dynamics of speech production and perception, G. Meyer, S. Greenberg, and P.L. Divenyi, eds., pp. 343–353.
25. Ahissar, M., Nahum, M., Nelken, I., and Hochstein, S. (2009). Reverse hierarchies and sensory learning. Philos. Trans. R. Soc. Lond. B Biol. Sci. *364*, 285–299.
26. Cheveigné, A. de (2010). Pitch perception. In The Oxford Handbook of Auditory Science: Hearing, C.J. Plack, ed. (Oxford University Press), pp. 71–104.
27. Lieder, I., Adam, V., Frenkel, O., Jaffe-Dax, S., Sahani, M., and Ahissar, M. (2019). Perceptual bias reveals slow-updating in autism and fast-forgetting in dyslexia. Nat. Neurosci. *22*, 256–264.
28. Ladd, D.R., Turnbull, R., Browne, C., Caldwell-Harris, C., Ganushchak, L., Swoboda, K., Woodfield, V., and Dediu, D. (2013). Patterns of individual differences in the perception of missing-fundamental tones. J. Exp. Psychol. Hum. Percept. Perform. *39*, 1386–1397.
29. Chialvo, D.R. (2003). How we hear what is not there: A neural mechanism for the missing fundamental illusion. Chaos *13*, 1226–1230.
30. Oxenham, A.J. (2012). Pitch Perception. J. Neurosci. *32*, 13335–13338.
31. Chambers, C., Akram, S., Adam, V., Pelofi, C., Sahani, M., Shamma, S., and Pressnitzer, D. (2017). Prior context in audition informs binding and shapes simple features. Nat. Commun. *8*, 15027.
32. Sheehan, T.C., and Serences, J.T. (2022). Attractive serial dependence overcomes repulsive neuronal adaptation. PLoS Biol. *20*, e3001711.
33. Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. Clin. Neurophysiol. *118*, 2544–2590.
34. Chen, Y., Huang, X., Luo, Y., Peng, C., and Liu, C. (2010). Differences in the neural basis of automatic auditory and visual time perception: ERP evidence from an across-modal delayed response oddball task. Brain Res. *1325*, 100–111.
35. Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. Nat. Neurosci. *6*, 391–398.
36. Raviv, O., Lieder, I., Loewenstein, Y., and Ahissar, M. (2014). Contradictory Behavioral Biases Result from the Influence of Past Stimuli on Perception. PLoS Comput. Biol. *10*, e1003948.
37. Knoblauch, K. (2022). CRAN - Package Psyphy: Functions for Analyzing Psychophysical Data in R (CRAN).
38. Wood, S. (2022). Package 'mgcv': Mixed GAM Computation Vehicle with Automatic Smoothness Estimation (CRAN).
39. Raviv, O., Ahissar, M., and Loewenstein, Y. (2012). How Recent History Affects Perception: The Normative Approach and Its Heuristic Approximation. PLoS Comput. Biol. *8*, e1002731.
40. Ashourian, P., and Loewenstein, Y. (2011). Bayesian inference underlies the contraction bias in delayed comparison tasks. PLoS One *6*, e19551.
41. Hollingworth, H.L. (1910). The Central Tendency of Judgment. J. Philos. Psychol. Sci. Methods *7*, 461.
42. Jaffe-Dax, S., Lieder, I., Biron, T., and Ahissar, M. (2016). Dyslexics' usage of visual priors is impaired. J. Vis. *16*, 10–19.
43. Karim, M., Harris, J.A., Langdon, A., and Breakspear, M. (2013). The influence of prior experience and expected timing on vibrotactile discrimination. Front. Neurosci. *7*, 255.
44. Olkkonen, M., McCarthy, P.F., and Allred, S.R. (2014). The central tendency bias in color perception: Effects of internal and external noise. J. Vis. *14*, 5.
45. Hastie, T., and Tibshirani, R. (1993). Varying-Coefficient Models. J. Roy. Stat. Soc. B *55*, 757–779.
46. Dai, H., and Micheyl, C. (2011). Psychometric functions for pure-tone frequency discrimination. J. Acoust. Soc. Am. *130*, 263–272.

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw and analyzed data	This paper	https://data.mendeley.com/datasets/kpppc634cd/2
Software and algorithms		
Python version 3.7.11	Python Software Foundation	https://www.python.org/
mixed GAM computation vehicle with automated smoothness estimation	mgcv package	https://cran.rproject.org/web/packages/mgcv/index.html.
Psychotoolbox-3	MATLAB toolbox	http://psychtoolbox.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to the lead contact, Merav Ahissar (msmerava@gmail.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The two-tone discrimination experiments were conducted online via the Amazon Mechanical Turk crowd-sourcing platform. The experiments were conducted in JavaScript and administered using web browsers with tones that were preloaded and played using HTML5 Audio. We used Psychotoolbox-3 MATLAB toolbox (*http://psychtoolbox.org/*) for creating the auditory stimuli. Analysis was conducted using the 'mixed GAM computation vehicle with automated smoothness estimation' (mgcv) free package *https://cran.rproject.org/web/packages/mgcv/index.html*.
- All the original code has been deposited at Mendeley Data and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

310 (Experiment 1), 400 (Experiment 2), and 338 participants (Experiment 3) were recruited. All participants (age range 20–50 years old) reported good hearing.

(1) In Experiment 1, their average age was 34.7 years (*SD* = 7.9), 46.1% were females, and they had in average 2.3 years of musical experience (*SD* = 4.0).
(2) In Experiment 2, their average age was 34.0 years (*SD* = 7.6), 44.8% were females, and they had in average 2.4 years of musical experience (*SD* = 4.6).
(3) In Experiment 3, their average age was 34.8 years (*SD* = 8.2), 45.0% were females, and they had in average 2.6 years of musical experience (*SD* = 4.5).

The experiments were conducted via the Amazon Mechanical Turk (M-Turk) platform. All participants (all from the United States) were recruited with previous M-Turk approval >95% and a total number of human intelligence tasks (HITs) > 1,000.

The study was approved by the Ethics Committee of the Psychology Department of the Hebrew University. All participants provided informed consent before their participation.

METHOD DETAILS

Experimental conditions

Our written instructions emphasized that the experiment must be performed: (1) using headphones in a quiet environment; (2) with either a laptop or a desktop computer; and (3) only by people with good hearing who are between the ages of 20 and 50. Each individual could only participate once. All the experiments consisted in a two-tone pitch discrimination task. Participants performed 300 trials (three blocks of 100 trials each) that lasted between 10 and 15 min and were paid US$2.

Exclusion criterion

We excluded participants whose accuracy $\leq$ 55% correct, since this low accuracy (around chance level) suggests that they may not have been performing the task. This yielded 22.9%, 17%, and 19.5% exclusion in Experiments 1, 2, 3, respectively. The mildly high rate of exclusion is largely an outcome of our choice of a non-adaptive protocol. We chose a non-adaptive protocol to avoid introducing correlations between the parameters of consecutive trials.[36] In non-adaptive protocols, the difference between the tones of a given trial is always sampled from the same distribution, regardless of the participant's performance, ensuring that similar stimuli statistics are presented across participants. Since the difficulty is not adapted to the level of each participant, poor listeners obtain lower scores.

In addition, to ensure the consistency of performance throughout the task, we measured the variance of accuracy in 10 windows of 60 trials (overlap of 30 trials) and we excluded participants whose variance exceeded the threshold of 2 (following Lieder et al., 2019). This criterion led to 2.9%, 4%, and 5.6% exclusions in Experiments 1, 2, 3, respectively.

The data of 230 (74.2%) (Experiment 1), 316 (79.0%) (Experiment 2) and 253 (74.9%) (Experiment 3) participants are included.

Two-tone pitch discrimination task

In each trial, two tones (220ms) were presented serially, with an inter-stimulus interval (800ms). Participants were asked which of the two tones had a higher pitch. Feedback was provided to the participants after each trial and also a summary of the accuracy of their responses after each block. Mean accuracy was 80% $\pm$ 1 (SE) in all three experiments.

Importantly, we randomly mixed (with equal probability) two types of trials: simple and harmonic complex tones, as detailed in the section below.

Stimuli

The distribution of the stimuli in each of the three experiments was log-uniform. Two types of trials, randomly selected with equal probability, were intermixed: trials in which both stimuli were simple tones and trials in which both stimuli were harmonic complex tones, composed of three harmonics (fundamental frequency f_0 and the two first overtones: $f_1 = 2\times f_0$ and $f_2 = 3\times f_0$) having equal intensities. Thus, approximately half of the trials were simple tones and half harmonic complex tones. The two tones S_1 and S_2 in each trial were always of the same timbre category (simple or complex). Simple tones were sampled log-uniformly from 500Hz to 2000Hz (2 octaves). The fundamental frequency f_0 of complex tones was sampled log-uniformly from 500Hz to 1000Hz (1 octave). Thus, the first and second overtones f_1 and f_2 ranged from 1000Hz to 2000Hz and from 1500Hz to 3000Hz, respectively. The frequency interval between S_1 and S_2 was sampled log-uniformly from the range 0.4–10.1% (uniformly from 6 to 167 cents). In both cases (simple- and complex-tone trials), the interval between the frequencies of the two tones S1 and S2 of each trial is much smaller than the frequency distance between the first tones in consecutive trials. Importantly, in Experiment 1, the complex tones were composed of all three harmonics (f_0, f_1, f_2). In Experiment 2, the complex tones did not include the fundamental f_0. In Experiment 3, the complex tones did not include the first overtone f_1.

QUANTIFICATION AND STATISTICAL ANALYSIS

Model fitting was performed using the *mgcv* R package.[37,38] Hypothesis tests were performed using the *scikit-learn* and *SciPy* packages in Python. All frequencies are log-transformed.

Calculating the contraction bias function by the most recent trial

Following Lieder et al.,[27] we used a Generalized Additive Model (GAM) for estimating the probability that a participant will respond that the second tone S_2^t on trial t, has a higher pitch.

$$P\left(''S_2^t > S_1^t{}''\right) = \Phi\left(\alpha_p\delta^t + b_0 + b_1\left(d_1^t\right) + b_\infty\left(d_\infty^t\right)\right)$$

$P(''S_2^t > S_1^t{}'')$ is determined by the standard normal cumulative distribution function Φ (the inverse of the probit link function) of the following sum: frequency difference δ^t between the (fundamental, in case of complex tones) frequencies of S_2^t and S_1^t, multiplied by the (pre-fitted) participant's frequency sensitivity α_p (higher α_p means better performance for participant p), a constant offset b_0 (representing a constant response preference across all trials and all participants) and the effect of previous trials (b_1 and b_∞). The constant offset b_0 was never found to be significantly different from zero and so is omitted when the model is discussed elsewhere in the text. The effect of previous trials is composed of two non-parametric (spline-estimated) functions b_1 and b_∞, which were shown to be additive[27]: recent and all other trials, respectively.

(1) $b_1(d_1^t)$ corresponds to the "recent bias",[27] i.e., the contraction to the previous trial, whose magnitude depends on the frequency distance, d_1^t, between the frequency of the first tone in the current trial, S_1^t, and the frequency of the first tone in the previous trial, S_1^{t-1}. We consider the contraction to the first tone of the previous trial rather than to the second or to the mean of the first and the second tones because the frequency interval between the two tones in each trial is much smaller compared to the frequency interval between the trials ($\delta^t << d_1^t$). Namely, the two tones in each trial are approximatively the same compared to the distance between the tones in the previous trial. In addition, we consider the contraction of the first tone in the current trial because its perceptual and memory noise is larger than that of the second stimulus. The first tone is indeed further away than the second from the moment of the participant's

response.[39] All the predictions of $b_1(d_1^t)$, aggregated over all participants, are accompanied by the corresponding standard errors, based on the posterior distribution of the model coefficient $b_1(d_1)$.

(2) $b_\infty(d_\infty^t)$ corresponds to the "longer-term effect"[27] – the contraction to the global mean frequency across the experiment. This term captures the bias toward probable stimuli.[40–44] Its magnitude varies as a function of the frequency distance, d_∞^t, between the frequency of the first tone in the current trial and the mean of the frequencies of the first tones across all trials. Intuitively, d_∞^t reflects the relative position of the current stimuli within the distribution.

In this paper, we are interested in the "recent bias" function, and will not address the longer-term effect. However, we introduced it to our model since it was shown that both biases significantly contribute to the model fit.[27] Importantly, this model has the advantage of separating the effect of recent history from the central tendency induced by the overall average of all previous trials.

For simplicity, we replaced $b_1(d_1^t)$ by linear regressions $w_1 \cdot d_1^t$, as they performed as well as the spline functions in these cases.

Thus, for simple (both octaves) after complex trials, the model takes the following form:

$$P\left(''S_2^t > S_1^t''\right) = \Phi\left(\alpha_p \delta^t + w_1 d_1^{t,(0.5k-1k)} + w_1' d_1'^{t,(1k-2k)} + b_\infty(d_\infty)\right)$$

where $d_1^{t,(0.5k-1k)}$ is the frequency distance between the first simple tone in the lower octave and the fundamental frequency of the first complex tone,

$d_1'^{t,(1k-2k)}$ is the frequency distance between the simple tone in the upper octave and the second harmonic of the first complex tone,

w_1 and w_1' are the corresponding slopes of the recent bias.

There are two components for the recent bias, describing the contraction by the fundamental frequency f_0 of the simple tone is in the lower octave (500-1000Hz) and the contraction by the second harmonic f_1 of the simple tone is in the upper octave (1000-2000Hz).

Similarly, for complex after simple (lower octave) trials, the model takes the following form, using the same notation:

$$P\left(''S_2^t > S_1^t''\right) = \Phi\left(\alpha_p \delta^t + w_1 d_1^{t,(0.5k-1k)} + b_\infty(d_\infty)\right)$$

where there is one component for the recent bias, describing the contraction by the simple tone (f) of the fundamental frequency (f_0) when the latter is in the lower octave.

Fitting single participants (random effects)

In the regression analysis described above, we assumed shared bias functions across participants. This assumption is a statistical necessity since we do not have sufficient individual data to estimate the model individually. To fit single participants, we assumed a shared parameter across participants but allowed small individual deviations (termed *random effects*). Specifically, we added two random effect terms, factorized by individual participants:

(1) A random effect on the bias intercept, namely a possible individual response preference bias b_0 that reflects a systematic tendency of a given participant to prefer "$S_2^t > S_1^t$ or vice versa regardless of the stimuli.

(2) A random effect on the bias slope, that permits comparing the bias magnitudes between individuals, by tuning the bias function. Specifically, we used a metric quantifying the influence of each bias fit – the mean of absolute value of the participant's bias magnitude: $\langle |b_1^t| \rangle$, where $b_1^t = b_1(d_1^t)$ is the bias magnitude in trial t.

These random effect terms allowed us to compare serial dependence magnitudes quantitatively, and account for inter-individual variability. Because assumed deviations are small, inference in these models is tractable.

Varying-coefficient models and lapse rate

A single model was fit for all participants, with common bias functions, but with participant-specific sensitivities α_p. We combined the pre-fitted α-participant pairing in the model using the varying-coefficients models method.[45] This method assumes linearity in the regressors, but their coefficients are allowed to change smoothly with the value of other variable (alphas in our case): $g(\alpha_p)\delta_p^t$. Here α_p is the pre-fitted sensitivity parameter, δ_p^t is the difference between the target tones for each participant p, and g is some smooth fitted function. We added a lapse parameter λ to account for occasional inattentiveness[46]: $P(''S_2^t > S_1^t'') = \frac{1}{2}\lambda + (1-\lambda)\Phi$. We set λ to a fixed 0.05.

Statistical tests

In each of the three subsections of the Results, we conducted statistical tests at the group level and at the level of individual participants:

At the group level, using a GAM model, we calculated the slope of the bias, based on the aggregated data over all participants. We performed Wald tests to assess whether the slope is statistically different from zero. We found that the slope is negligible (about zero) and not significant when specific harmonics were missing.

Adding random effects to the GAM model, enabled us to calculate a slope estimate for each individual participant. This procedure makes it possible to obtain approximative estimates at the level of individual participants, when there is no sufficient data associated with each participant for estimating the model individually. This was particularly insightful for comparing the distribution of slopes between the three experiments (conditions), and testing whether they differ significantly. Indeed, by conducting a Kruskal-Wallis test we showed that the distributions corresponding to the three experiments differed significantly. Then, the Dunn-Bonferroni test validated that contraction was much larger when the specific harmonic was present compared to when it was missing.

We also performed ANOVA for nested GAM models to validate whether the predictors d_1 and d_1' contribute significantly to the model.

In addition, the statistical test that we reported in the figure captions corresponds to the Wilcoxon signed-rank test on the slope distributions that were calculated at the individual participant level with random effects. We found that the slope distribution of each of the three experiments significantly differed from zero. However, most importantly, the magnitude of the bias was always much smaller in the case of contraction to a missing harmonic, compared to a physically present one, which is in line with the results obtained at the group level. Indeed, panel g in Figures 1, 2, and 3, shows that in the case of a missing harmonic, a large majority of the dots are very close to zero. This is not the case for the other two experiments, where the dots are symmetrically distributed around the median which is clearly above zero. More precisely, we report here the values of the medians of the individual slopes (the means have very similar values): (a) Contraction of f_0 to pure tones (first subsection of the Results), median equals 0.087 and 0.099 in Experiments 1 and 3, respectively, while median equals 0.033 in Experiment 2. (b) Contraction of pure tones to f_0 (second subsection of the Results), median equals 0.11 and 0.15 in Experiments 1 and 3, respectively, while median equals 0.044 in Experiment 2. (c) Contraction of pure tones to f_1 (third subsection of the Results), median equals 0.074 and 0.14 in Experiments 1 and 2, respectively, while median equals 0.034 in Experiment 3.