Neural responses of Interval Judgment in the Tritone Paradox

Subeen Kim[1] (maisie.been@snu.ac.kr), Jusung Ham[2], Inyong Choi[2], Kyogu Lee[1,3,4] (kglee@snu.ac.kr)
[1]Graduate School of Convergence Science and Technology, Seoul National University, Republic of Korea
[2]Department of Communication Sciences and Disorders, University of Iowa, United States
[3]Interdisciplinary Program in Artificial Intelligence, Seoul National University, Republic of Korea
[4]Artificial Intelligence Institute, Seoul National University, Republic of Korea

Abstract

The Tritone Paradox is an auditory illusion in which a sequence of two complex tones is perceived as either ascending or descending, depending on the individual. It presents an interesting phenomenon for investigating pitch perception in contexts. however, no neurophysiological study has been conducted. This study identified event-related potential (ERP) correlates of pitch judgments under different pitch contexts. Twenty-seven participants judged whether the tritone pair was perceived as ascending or descending after listening to a sequence of ascending or descending tone pairs. Cortical auditory evoked responses to the second tone of the tritone pair were compared across contexts. In the Rise context, standard stimuli evoked larger responses at Fp1; in the Fall context, deviant stimuli elicited stronger responses across all sites. These results suggest that frontal and central brain regions are involved in processing ambiguous pitch stimuli, and that ERP responses reflect the interaction between stimulus context and perceptual.

Keywords: Tritone paradox; Shepard tones; Auditory illusion; Pitch perception; Event-related potentials (ERP); N1-P2; Auditory Decision-making; Perceptual ambiguity; context effects

Introduction

Auditory illusions, unlike the widely recognized optical illusions, remain relatively less known. These phenomena arise when the brain misinterprets sound, causing a gap between actual and perceived input. Notable examples include mondegreens—misheard lyrics caused by phonetic ambiguity (Beck, Kardatzki, & Ethofer, 2014), and the Shepard tone illusion, where octave-spaced tones create an illusion of endlessly ascending or descending scale due to overlapping spectral components (R. N. Shepard, 1964). Such illusions arise from auditory processing of amplitude and frequency information, paralleling the principles underlying optical illusions in visual perception (Davidson, 2012).

Pitch perception involves both pitch height and pitch class, yet neither dimension alone provides sufficient information to determine whether a sound is perceived as higher or lower. In music perception, pitch height follows a logarithmic relationship with frequency and is perceived as a linear dimension. In contrast, the pitch class is a cyclic dimension, divided into 12 semitones per octave in western music (R. Shepard, 1982). To explore how pitch height and class interact, Shepard synthesized tones containing only pitch-class information, hypothesizing that listeners would perceive them as invariant in pitch height. However, empirical findings indicate that listeners differentiate pitch based on the relative proximity between tones, suggesting that pitch perception operates through relative shifts from a reference.

The tritone paradox arises when two Shepard tones, spaced six semitones apart (a tritone), are played in succession, resulting in varying pitch perceptions among listeners. Due to their ambiguity, listeners perceive either direction based on internal pitch class templates (R. Shepard, 1982). The experimental results of previous studies indicate that these judgments differ between individuals, showing specific tendencies or significant patterns in their responses (Deutsch, 1987, 1991).

Several theories have been proposed to explain individual differences in pitch perception in the tritone paradox. A widely accepted explanation is the pitch class comparison mechanism, shaped by individually developed pitch class templates (Deutsch, 1991, 2007; Deutsch, Henthorn, & Dolson, 2004). Cross-linguistic studies support this hypothesis by showing that early linguistic exposure significantly influences pitch perception. For example, comparisons between English speakers from California and southern England (Deutsch, 1991, 1994, 2007), Mandarin-speaking adolescents (Deutsch, Jiang, Henthorn, & Zhou, 2014), and Vietnamese-English bilinguals (Deutsch et al., 2004) provide evidence that early language experience affects pitch perception. These studies revealed group-level differences in the direction of perceived pitch change: for example, Californian and Southern British English speakers showed systematically reversed patterns of 'Rise' and 'Fall' judgments in response to identical tone pairs. In addition, speakers of tone languages such as Mandarin and Vietnamese exhibited greater consistency in their responses. Such findings suggest that early exposure to the prosodic and tonal characteristics of one's native language can shape the internal formation of pitch-class templates and influence how ambiguous pitch stimuli are interpreted.

Psychoacoustic theories offer another view (Malek & Sperschneider, 2018). Virtual pitch theory (Terhardt, 1979) suggests that complex tones with multiple frequencies are perceived as a single auditory event, which can be mathematically modeled. This concept has been extended to Shepard tone perception, supporting a probabilistic processing model of the tritone paradox.

In a study by Malek and Sperschneider (2018), systematic adjustments to the spectral envelope of Shepard tones showed

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

that participants' reference pitch shifted by about five semitones, aligning with (1997), who reported a six-semitone shift. challenging the assumption of fixed absolute pitch class templates.

However, Malek and Sperschneider (2018) argue that their probabilistic model does not fully contradict the pitch class comparison mechanism. Their model includes a threshold function that highlights specific frequency ranges in auditory perception. Since this function varies among individuals, predictions depend on whether a person's preferred frequency range is well-defined. When distinct frequency preferences exist, their model aligns with Deutsch (1991, 1994), further supporting the link between pitch perception and linguistic exposure. This suggests that early auditory environments, particularly phonetic features of language, shape preferred frequency ranges.

Although prior theories have emphasized the role of pitch class templates and frequency-based filtering in explaining individual differences in pitch perception, few studies have examined how these internal representations manifest at the neural level during perceptual decision-making. This study addresses this gap by investigating how the brain encodes directional pitch judgments in the tritone paradox using event-related potentials (ERPs). Importantly, we manipulate the preceding pitch context—presenting each tone pair within either a globally rising or falling sequence—to examine how listeners integrate sequential auditory information to resolve ambiguity. This design enables us to test whether neural responses reflect probabilistic auditory processing, where context dynamically guides perceptual resolution based on pitch-class templates or frequency preferences shaped by early experience. We hypothesized that deviant stimuli—tones that contradicted the expected pitch direction implied by the preceding global context (i.e., rising or falling sequences)—would elicit stronger ERP responses, reflecting prediction error signals associated with context violation.

Methods

Participants

Thirty adults were initially recruited for the study. However, three participants were excluded due to EEG equipment malfunction, resulting in a final sample of 27 participants (11 males, 16 females; *M* age = 26.78 years, *SD* = 6.38). All participants included in the analysis reported normal or corrected-to-normal vision and no hearing impairments.

Among them, 19 participants (70.4%) had prior musical training, including three music majors (11.1%) specializing in piano or composition. The music majors had an average training duration of 25.33 years (*SD* = 7.32), while non-majors reported an average of 8.25 years (*SD* = 5.17). Reported instruments included piano, flute, violin, trumpet, trombone, and traditional Korean instruments such as the gayageum, danso, and haegeum.

All procedures were approved by the Institutional Review Board of Seoul National University.

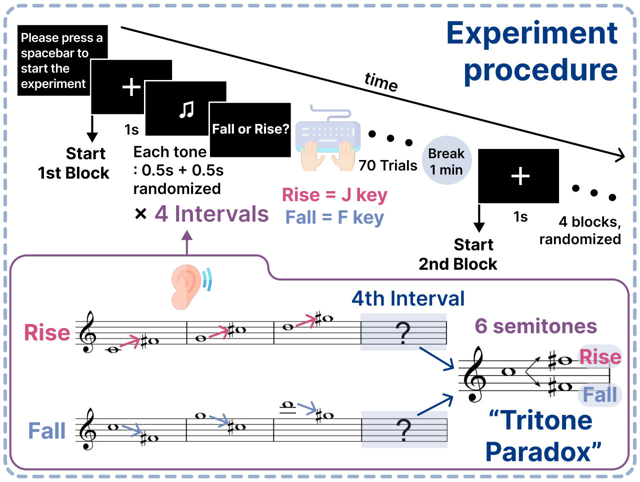

Figure 1: Procedure of the Tritone Paradox experiment. Participants judged pitch direction (rise/fall) in the Tritone Paradox (4th interval tone pairs)

Auditory Stimuli

To avoid uncomfortable auditory responses and reduce pitch ambiguity, we restricted the stimulus frequency range to the mid-frequency band of C3–C6 (130.81–1046.50 Hz). This range is musically and linguistically natural and avoids the sharpness and perceptual unreliability associated with higher frequencies (Oxenham, 2012; Moore, 2012; Bendor & Wang, 2005).To account for differences in auditory sensitivity across frequencies, we adjusted the sound pressure levels of the stimuli based on the equal-loudness contour (Fletcher & Munson, 1933), thus reducing excessive emphasis on high frequencies and maintaining perceptual balance.

The stimuli were Shepard tones, created by superimposing multiple sine waves across octaves that share the same pitch class. Each component was modulated by amplitude using a fixed Gaussian spectral envelope, with the central octave presented at the highest intensity and the adjacent octaves attenuated, forming a perceptually unified tone. Each stimulus consisted of two 500 ms tones separated by a fixed interval of a tritone (6 semitones). This interval, known as a half-octave in Western music theory, represents a perceptually ambiguous region due to the lack of shared harmonics, and this ambiguity is further amplified when combined with Shepard tones (R. N. Shepard, 1964; Deutsch, 1987).

The stimuli were categorized into three types: Tritone Rise, Tritone Fall, and Paradox. Rise and Fall each included 31 stimuli, and the Paradox set contained 24 stimuli (with C3-based items excluded to equalize the number of octave layers between pitch classes). All stimuli were presented in a randomized order.

Procedure

Before the experiment, participants completed an online demographic questionnaire assessing age, gender, hearing and vision status, and musical training. A fixed sound level was used, and participants confirmed it was comfortable. A black screen with a white fixation cross was displayed throughout the session to help maintain visual focus and alertness.

The auditory task was implemented using Psychtoolbox and consisted of 280 trials, divided into four blocks of 70 trials each. The entire session lasted approximately 50 minutes. Before starting the task, participants read the instructions on the screen and pressed the space bar to start the experiment.

To ensure comprehension and verify participants' ability to perceive pitch direction, four practice trials were presented before the main experiment. Each practice trial contained four tone pairs moving in the same pitch direction (e.g., Rise-Rise-Rise-Rise or Fall-Fall-Fall-Fall). After hearing the final pair, participants pressed the 'J' key if they perceived the pitch as ascending and the 'F' key if descending. Immediate feedback ("Correct" or "Incorrect") was given after each response.

During the main experiment, each trial consisted of four tone pairs, following the same format as the practice session. However, unlike in the practice trials, only the final (fourth) pair was a Tritone Paradox stimulus, while the first three pairs served as contextual cues in either the Rise or Fall direction. At the end of each block, participants were given a one-minute break and shown their progress on the screen.

After completing all four blocks, EEG recording was terminated and the experimental session concluded.

EEG Recording and Data Processing

The experiment was carried out in a soundproof EEG booth designed to minimize noise and visual distractions. The participants viewed stimuli on a 24-inch Samsung monitor and responded using a standard keyboard. Memory foam in-ear earphones ensured a secure fit and reduced external noise interference.

EEG data were recorded using the BIOSEMI ActiveTwo system with 32 scalp electrodes following the 10–20 system. Additional electrodes included CMS and DRL for signal stabilization, EXG4 and EXG5 on bilateral mastoids for reference, and EXG6 below the left eye to monitor eye movements and blinks.

EEG pre-processing involved the following steps to ensure signal quality and artifact reduction. First, the signals were re-referenced to the bilateral mastoid electrodes (EXG4, EXG5) to reduce reference noise. A band-pass filter of 1–20 Hz was applied to remove low-frequency drifts and high-frequency noise while preserving neural signals of interest. The continuous data were then segmented into epochs ranging from -300 to 1300 ms relative to the onset of the event corresponding to the presentation of the paradox stimulus ("condition 3"). The dataset was subsequently down-sampled from 2048 Hz to 256 Hz to optimize computational efficiency without compromising temporal resolution.

Bad channels were identified by visual inspection and persistently noisy channels were removed and interpolated. To eliminate excessive artifacts, a preliminary manual trial rejection was performed by scrolling through raw EEG data and removing trials that contained significant distortions. Subsequently, an independent component analysis (ICA) was performed to remove artifacts, including those related to eye movement, using the electro-oculogram (EOG) channel (EXG 6) placed below the left eye as a reference for component identification. The EOG channel was excluded from further analysis after correction of artifacts.

After ICA correction, baseline correction was applied using the -200 to 0 ms pre-stimulus window. To verify that the baseline correction was performed correctly, epochs with an absolute baseline deviation exceeding 0.5 µV were rejected. Additional artifact rejection was based on a voltage threshold of 50-70 µV, a peak-to-peak amplitude criterion of 100 µV, and a signal-to-noise ratio (SNR) threshold of 5.0, with trials falling below this threshold manually removed.

ERP Data Analysis

The N1–P2 analysis window was defined as 620–830 ms, time-locked to the onset of the second tone in each trial. This window accounted for a 50 ms auditory device delay and a ± 30 ms buffer to accommodate individual differences in neural response latency. We focused on the second tone because perceptual ambiguity in the Tritone Paradox arises from interpreting the interval between tones rather than the tones themselves. The window was defined based on the grand-averaged ERP waveform to ensure consistency across participants.

Paired t-tests were also computed across all 32 scalp electrodes for exploratory visualization. Since not all electrodes satisfied the normality assumption, no inferential claims were made outside Fp1, Fz, and Cz. The resulting t-values were plotted as scalp topographies to illustrate the spatial distribution of condition differences. For completeness, an FDR correction was applied across the 32 channels at $q = 0.05$.

ERP data analysis focused on frontal electrodes associated with auditory processing. To determine the suitability of parametric testing, we assessed the normality of N1–P2 peak-to-peak amplitude differences between standard and deviant conditions at Fp1, Fz, and Cz for each participant using the Lilliefors test ($\alpha = 0.05$) These three electrodes satisfied the assumption of normality across all participants and were analyzed using paired t-tests.

Behavioral Response Analysis

To analyze perceptual judgments, we calculated the proportion of 'standard' and 'deviant' responses for each individual in the Rise and Fall conditions. A 'standard' response matched the direction of the preceding context (e.g., judging the Paradox pair as rising after a Rise context), whereas a "Deviant" response did not. These response proportions, bounded between 0 and 1, were compared within each condition using paired t-tests. Before performing the t tests, we verified the normality of the difference scores (e.g., Rise-standard minus Rise-deviant) using the Shapiro-Wilk test. The test results indicated that the differences did not deviate significantly from normality, justifying the use of parametric tests. All statistical analyzes were performed with Python.

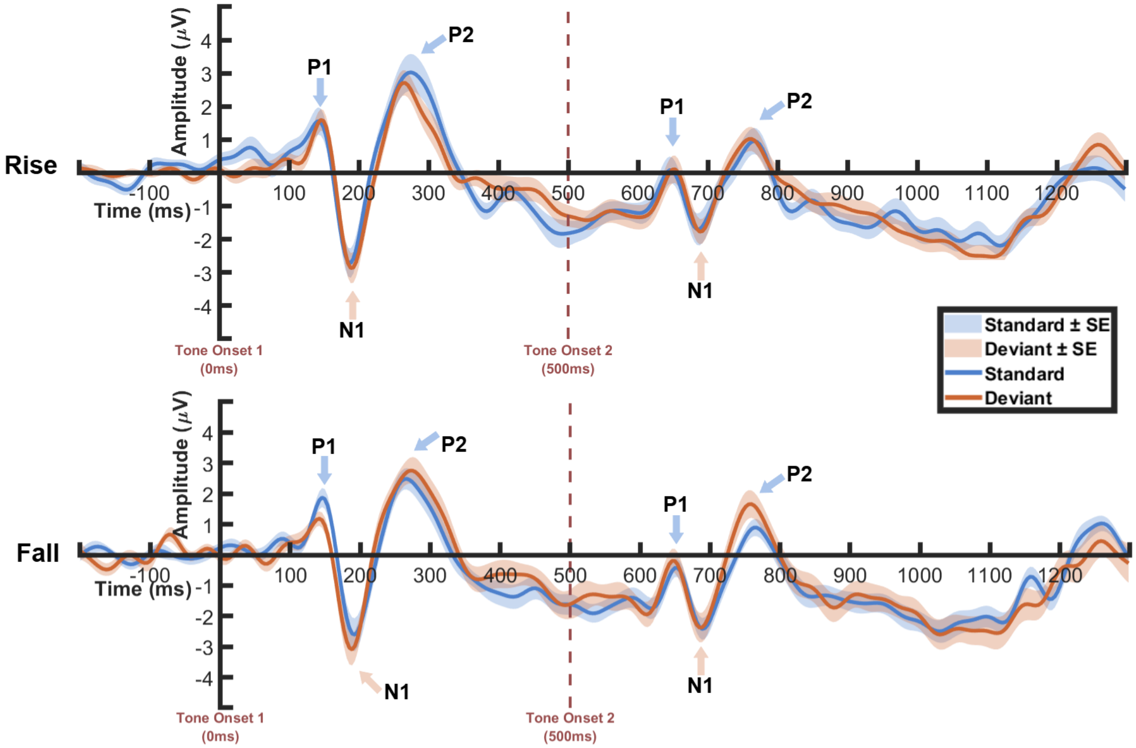

Figure 2: Fp1, grand average ERP waveforms for standard and deviant stimuli (±SE) in the Rise (top) and Fall (bottom) conditions. Shaded areas represent standard error (SE) bands.

Result

Participant Response Trends in Contextual Judgments

The mean proportion of "Deviant" responses in the Rise context was 0.29 ($SD = 0.08$), while the mean proportion of "Standard" responses in the Fall context was 0.31 ($SD = 0.09$). In contrast, "Standard" responses in the Rise context and "Deviant" responses in the Fall context were lower, with means of 0.19 ($SD = 0.07$) and 0.18 ($SD = 0.09$), respectively (see Table 1).

Table 1: Mean and standard deviation of response proportions for "Standard" and "Deviant" judgments in Rise and Fall contexts.

Condition	Standard (M ± *SD*)	Deviant (M ± *SD*)	*p*
Rise	0.19 ± 0.07	0.29 ± 0.08	$< .001$
Fall	0.31 ± 0.09	0.18 ± 0.09	$< .001$

Topographical Analysis of T-Scores

Figure 3 shows the scalp topography of t-scores computed via paired t-tests across all 32 EEG electrodes. These t-tests were conducted for exploratory visualization, with statistical inference limited to electrodes that met the normality assumption. After FDR correction ($q < 0.05$), electrode Fp1 exhibited significant differences in both rise ($t = 2.17$, $p = .039$) and fall conditions ($t = -3.90$, $p < .001$).

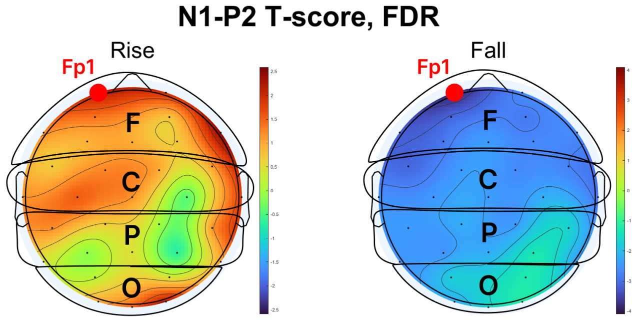

Figure 3: Scalp topographies of t-scores for N1–P2 amplitude differences (standard vs. deviant; FDR-corrected). F, C, P, O = frontal, central, parietal, occipital. Red dots indicate electrodes with significant, normally distributed effects.

N1-P2 Peak-to-Peak Amplitude Differences

N1–P2 peak-to-peak amplitudes were analyzed within a time window of 620-830 ms, time-locked to the second tone onset.

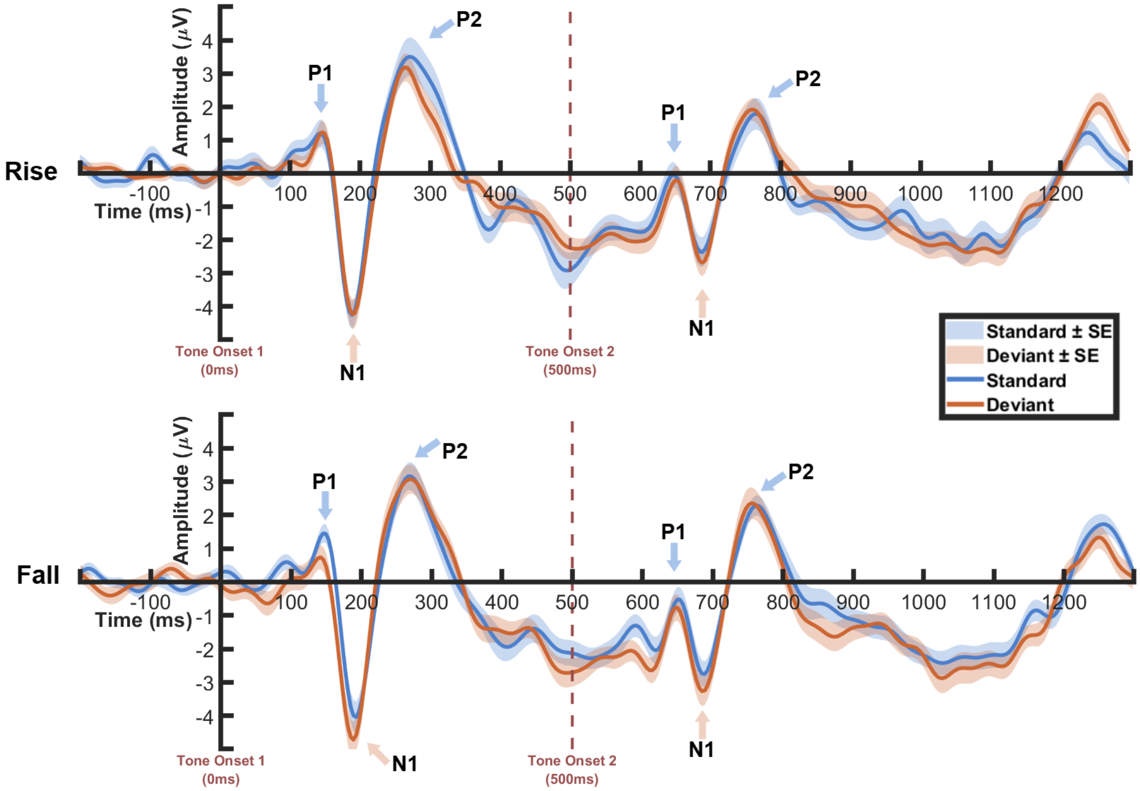

Figure 4: Fz and Cz, Grand average ERP waveforms for standard and deviant stimuli in the Rise (top) and Fall (bottom) conditions. Shaded areas represent standard error (SE) bands.

Significant differences were observed in both the Rise and Fall conditions, In the *Rise* condition, a significant difference was observed only at the Fp1 electrode, where standard stimuli elicited stronger N1–P2 amplitudes than deviant stimuli (M = 5.10 ± 2.06 vs. 4.51 ± 1.76; $t = 2.17$, $p = .039$, Cohen's $d = 0.43$). This effect is visually evident in Figure 2, which shows a more prominent N1–P2 peak following the second tone onset (approximately 500 ms) in the standard condition at Fp1. No significant differences were found at Fz or Cz (Fz: $p = .339$, $d = 0.42$; Cz: $p = .259$, $d = 0.09$).

Table 2: Paired t-test results for N1–P2 peak-to-peak amplitudes at Fp1, Fz and Cz electrodes (*Mean* ± standard deviation(*SD*) of individual amplitudes).

Condition (Channel)	**Standard** (*M* ± *SD*)	**Deviant** (*M* ± *SD*)	***p*-value**	**Cohen's *d***
Rise (Fp1)	5.10 ± 2.06	4.51 ± 1.76	.039*	0.43
Rise (Fz)	6.63 ± 2.91	6.27 ± 2.40	.339	0.42
Rise (Cz)	6.15 ± 2.43	5.84 ± 2.16	.259	0.09
Fall (Fp1)	4.57 ± 2.17	5.89 ± 2.11	.001***	0.76
Fall (Fz)	6.67 ± 2.70	7.82 ± 3.05	.023*	0.24
Fall (Cz)	6.19 ± 2.37	7.27 ± 2.94	.024*	0.25

* $p < .05$, *** $p < .001$. Units: μV.

In contrast, the Fall condition showed significant differences at all three electrodes—Fp1, Fz, and Cz—with deviant stimuli eliciting stronger responses than standard stimuli. The corresponding statistics are reported in table 2. These differences are visually observable in Figure 2 (Fp1) and Figure 4 (Fz and Cz), where the N1-P2 amplitudes for deviant stimuli appear consistently larger across all sites following the second tone onset.

Discussion

We hypothesized that deviant stimuli would elicit stronger neural responses when they violated the preceding context (i.e., rising or falling pitch sequence), reflecting prediction error signals. Contrary to this hypothesis, participants consistently showed a behavioral bias toward perceiving falling pitch directions regardless of context. In contrast, ERP responses—particularly N1–P2 amplitudes—were significantly larger when participants perceived rising pitch sequences. This neural effect was most pronounced at the frontal site (Fp1) in the Rise condition, and across all three sites (Fp1, Fz, Cz) in the Fall condition. These findings suggest a dissociation between perceptual judgment and neural sensitivity to pitch direction.

Frontal Involvement in Pitch Perception

As a result of the analysis, the ERP effect at the frontal site (Fp1) was particularly pronounced in the Fall condition ($d = 0.76$). This effect was substantially larger than those observed at central electrodes (Fz, Cz; $d < 0.3$), suggesting enhanced prediction error processing in the frontal cortex during perceptual uncertainty. This is consistent with research showing the prefrontal cortex supports auditory decision-making under ambiguity. (Alain, Snyder, He, & Reinke, 2006; Wild et al., 2012). The deviance-related ERP response observed in the frontal region may also reflect early sensory processing, as it aligns with phenomena such as frontal negativity and auditory N1 (Näätänen, Paavilainen, Rinne, & Alho, 2007).

Interestingly, the fact that this effect was more prominent in the Fall condition may indicate that descending auditory patterns provide greater perceptual salience within Shepard tone-based Tritone Paradox stimuli, or alternatively, that predictive mechanisms in the frontal cortex respond more sensitively to specific auditory contexts.

Future research should employ high-density EEG systems (e.g. 64-128 channels) to examine functional differentiation within frontal subregions. Additionally, source localization techniques (e.g., sLORETA, eLORETA) are recommended to more precisely identify the origins of ERP signals, thereby clarifying the specific functional roles of frontal regions in pitch perception.

Auditory Decisions under Perceptual Ambiguity

Analysis of N1–P2 peak-to-peak amplitudes revealed condition-specific (Rise vs. Fall) and electrode-dependent ERP response patterns. Particularly in the *Fall* condition, deviant stimuli elicited significantly greater amplitudes than standard stimuli at frontal (Fp1) and central electrodes (Fz, Cz). Conversely, the *Rise* condition showed a limited effect restricted to the frontal site (Fp1), where standard stimuli evoked stronger responses.

This alignment between the rarity of behavioral responses and ERP amplitudes suggests that the brain is sensitive not only to physical stimulus properties but also to subjective probability violations. That is, auditory decision-making under perceptual ambiguity appears to be influenced by probabilistic processing based on expectations formed through preceding context.

This interpretation aligns with probabilistic models explaining the tritone paradox, such as the *virtual pitch theory* and frequency-sensitivity-based predictive models. Specifically, Malek and Sperschneider (2018) proposed that the perception of Shepard tones is dynamically determined by individual frequency preferences and contextual cues, rather than fixed pitch-class templates.

Furthermore, these patterns may reflect *prediction error signals* arising from discrepancies between expected and actual auditory inputs, highlighting the neural basis of decision-making under contextual influences. This perspective is consistent with findings that rare or unexpected stimuli elicit larger ERP components, such as the P3b, which are associated with the processing of infrequent and task-relevant events (Polich, 2007; Donchin & Coles, 1988).

Inconsistency in Contextual Pitch Judgment

Subject responses exhibited inconsistency between *Rise* and *Fall* contexts. In the *Rise* context, standard responses were less frequent and elicited stronger ERP responses, whereas in the *Fall* context, deviant responses were less frequent and associated with larger ERP amplitudes.

This response pattern allows for several interpretations. First, participants may have perceived the judgment tone as deviating from the preceding context in the *Rise* condition, leading to a higher frequency of deviant responses. Conversely, in the *Fall* condition, participants tended to perceive the judgment tone as consistent with the preceding context, resulting in more standard responses. This suggests that perception is sensitive to the structural flow of preceding stimuli.

Second, this asymmetry may relate to inherent characteristics of the tritone paradox. Notably, both the *Rise-Deviant* and *Fall-Standard* conditions, which were more frequently endorsed, may reflect a perceptual bias or enhanced salience for descending tones.

However, as the present study employed equal-loudness contours to control for intensity differences, it is unlikely that these response patterns are solely attributable to physical loudness variations. Consequently, the observed asymmetry may stem from either (1) context-based directional contrast effects or (2) an intrinsic perceptual salience for descending tones. The current study design does not allow for a definitive dissociation between these possibilities.

Future research should consider employing a simplified experimental session that presents judgment tones without preceding context. This approach would help isolate the effects of directional salience from contextual influences, thereby clarifying the underlying mechanisms of pitch perception under ambiguity.

Conclusion

This study investigated how auditory perception and brain responses interact under ambiguous pitch conditions using Shepard tone-based Tritone Paradox stimuli. Participants tended to perceive falling pitch sequences across contexts. However, their neural responses—especially N1–P2 amplitudes—were stronger when they perceived rising pitch sequences, most notably at the frontal site (Fp1).

These results suggest that, although participants judged based on familiarity, the brain responded more strongly to less expected pitch directions, revealing how it processes uncertainty. This indicates that perceptual judgments are shaped by both external context and internal prediction mechanisms.

Our findings highlight the importance of combining neural and behavioral approaches when studying how people perceive ambiguous sounds. They also offer a deeper understanding of how the brain deals with uncertainty and unexpected events in hearing.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00461617)

References

Alain, C., Snyder, J. S., He, Y., & Reinke, K. S. (2006, June). Changes in auditory cortex parallel rapid perceptual learning. *Cerebral Cortex*, *17*(5), 1074–1084. Retrieved from `https://doi.org/10.1093/cercor/bhl018`

Beck, C., Kardatzki, B., & Ethofer, T. (2014). Mondegreens and soramimi as a method to induce misperceptions of speech content – influence of familiarity, wittiness, and language competence. *PLOS ONE*, *9*, 1–5.

Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, *436*, 1161-1165.

Davidson, G. (2012). *Apparent motion and the tritone paradox: An eeg investigation of novel bistable stimuli*. Doctoral dissertation, Reed College.

Deutsch, D. (1982). The processing of pitch combinations. *Psychological Review*.

Deutsch, D. (1987). The tritone paradox: Effects of spectral variables. *Perception Psychophysics*, *41*, 563–575.

Deutsch, D. (1991). The tritone paradox: An influence of language on music perception. *Music Perception*, *8*, 335–347.

Deutsch, D. (1994). The tritone paradox and the pitch range of the speaking voice: Reply to repp. *Music Perception: An Interdisciplinary Journal*, *12*, 257–263.

Deutsch, D. (2007). Mothers and their offspring perceive the tritone paradox in closely similar ways. *Archives of Acoustics*, *32*, 3–14.

Deutsch, D., Henthorn, T., & Dolson, M. (2004). Speech patterns heard early in life influence later perception of the tritone paradox. *Music Perception*, *21*, 357–372.

Deutsch, D., Jiang, C., Henthorn, T., & Zhou, S.-Y. (2014). The tritone paradox among chinese children aged 12 and 13. In *Proceedings of a conference*.

Donchin, E., & Coles, M. G. H. (1988). Is the p300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*(3), 357–374. doi: 10.1017/S0140525X00058027

Fletcher, H., & Munson, W. A. (1933). Loudness, its definition, measurement and calculation. *The Journal of the Acoustical Society of America*, *5*(2), 82–108. doi: 10.1121/1.1915637

Malek, S., & Sperschneider, K. (2018). Aftereffects of spectrally similar and dissimilar spectral motion adaptors in the tritone paradox. *Frontiers in Psychology*, *9*, 677. doi: 10.3389/fpsyg.2018.00677

Moore, B. C. (2012). *An introduction to the psychology of hearing*. Brill.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (mmn) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544-2590. doi: https://doi.org/10.1016/j.clinph.2007.04.026

Oxenham, A. (2012). Pitch perception. *The Journal of Neuroscience*, *32*, 13335-13338.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. Retrieved from `https://doi.org/10.1016/j.clinph.2007.04.019` doi: 10.1016/j.clinph.2007.04.019

Repp, B. H. (1997). Spectral envelope and context effects in the tritone paradox. *Perception*, *26*, 645-665.

Shepard, R. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review*, *89*, 305-333.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *The Journal of the Acoustical Society of America*, *36*, 2346-2353.

Terhardt, E. (1979). Calculating virtual pitch. *Hearing Research*, *1*, 155-182.

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: The processing of degraded speech depends critically on attention. *Journal of Neuroscience*, *32*(40), 14010–14021. doi: 10.1523/JNEUROSCI.1528-12.2012