Title
The Effects of Musical Factors on the Perception of Auditory Illusions

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Choi, Ahyeon
Bang, Younyoung
Park, Jeong Mi
et al.

Publication Date
2024

Peer reviewed

The Effects of Musical Factors on the Perception of Auditory Illusions

Ahyeon Choi[1,*] **Younyoung Bang**[2,*] **Jeong Mi Park**[2] **Kyogu Lee**[1,3,4]

[1] Department of Intelligence and Information, Seoul National University
[2] Department of Musicology, Seoul National University
[3] Interdisciplinary Program in Artificial Intelligence, Seoul National University
[4] Artificial Intelligence Institute, Seoul National University
chah0623, bang1381, mensa04, kglee@snu.ac.kr

Abstract

This study delves into how various musical factors influence the experience of auditory illusions, building on Diana Deutsch's scale illusion experiments and subsequent studies. Exploring the interaction between scale mode and timbre, this study assesses their influence on auditory misperceptions, while also considering the impact of an individual's musical training and ability to discern absolute pitch. Participants were divided into non-musicians, musicians with absolute pitch, and musicians with relative pitch, and were exposed to stimuli modified across three scale modes (tonal, dissonant, atonal) and two timbres (same, different). The findings suggest that scale illusions occur less frequently with different timbres and vary with scale mode. Crucially, the absolute pitch ability appears to have a more significant impact on the perception of illusions than the duration of musical training. This research contributes to understanding the complex interplay between various factors in auditory perception and the mechanisms behind the experience of auditory illusions.

Keywords: Auditory illusion; Scale illusion; Timbre; Mode; Musical training; Absolute pitch

Introduction

Auditory illusions offer a fascinating glimpse into the complexities of how we perceive sound, showcasing moments when our auditory system interprets sounds differently from their actual sources. In the face of perceptual complexity, we are capable of distinguishing and analyzing sound attributes such as pitch(Stainsby & Cross, 2009a), timbre(McAdams, 2013), and tempo(Quinn & Watt, 2006), demonstrating our advanced engagement with and interpretation of auditory information. This highlights the critical importance of exploring the mechanisms behind auditory perception, providing insight into the complex relationship between sound attributes and how we perceive them.

Building on this foundation, the study of auditory illusions is pivotal for unraveling the underlying principles of the auditory system. Characterized by auditory misperception, these phenomena have been the focus of extensive scientific scrutiny. Groundbreaking research into Shepard Tones(Shepard, 1964), the Octave Illusion(Deutsch, 1974), the Scale Illusion(Deutsch, 1975a), and the Speech-to-Song Illusion(Deutsch, Henthorn, & Lapidis, 2011) has propelled our understanding forward. These studies shed light on the intricate rules that govern auditory perception, significantly enhancing our ability to interpret complex sounds.

Among various auditory illusions, Diana Deutsch's scale illusion(Deutsch, 1975a) stands out as a particularly compelling example of auditory misperception. In her experiment, sounds from a continuous scale were presented alternately to each ear, arranging discontinuous and distant sounds to be heard in one ear. This innovative approach led to participants perceiving ascending notes in one ear and descending notes in the other simultaneously(Deutsch, 1975a). Contrary to the expected perception of interleaved zigzag patterns, which would be predicted by traditional auditory processing theories, most listeners reported experiencing cohesive, sequential scale patterns across both ears. This unexpected outcome challenges conventional understanding and highlights the complexity of auditory perception.

This unexpected outcome has been linked to the brain's grouping mechanisms(Deutsch, 1999), illustrating how our auditory system organizes sounds into coherent streams. These grouping principles, which rely on various sound attributes like frequency, amplitude, temporal and spatial positioning, and even timbre(Deutsch, 1999), enable listeners to construct meaningful auditory experiences from complex acoustic inputs. This insight reveals the sophisticated nature of auditory perception, highlighting our innate ability to integrate and interpret intricate sound patterns through perceptual grouping.

The scale illusion is interpreted as the auditory system processing scales by grouping closer notes, rather than associating more distant notes(Deutsch, 1999). However, subsequent research indicates that changes in the characteristics of scale stimuli or differences in participant group traits can lead to variations in the degree or tendency of experiencing the illusion(Radvansky, Hartmann, & Rakerd, 1992; Smith, Hausfeld, Power, & Gorta, 1982). This suggests that specific musical factors and individual characteristics play a significant role in these perceptions. However, the precise impact of these factors remains unclear, necessitating further research to systematically vary them. This approach can identify the specific elements contributing to auditory illusions, enhancing our understanding of the intricate interplay between musical stimuli and individual characteristics in perception.

Reflecting on the foundational works of Diana Deutsch(Deutsch, 1975a, 1975b), which focused on stimuli primarily in C major, this approach highlighted the importance of proximity in the perception of scale illusions but left open the question of the influence of the scale mode's familiarity. As people are primarily exposed to and

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

educated in 'tonal music'—music that is organized around a central key and involves structured hierarchical relationships between pitches, providing resolution and closure(Grove, Sadie, Tyrrell, & Levy, 1980)—it's not surprising that 'atonal music', which avoids these tonal centers, can be more challenging to identify. This is evidenced by research showing that individuals generally find it more challenging to identify the musical structure of atonal music compared to tonal music, with this effect being most pronounced among non-music major students(Daynes, 2011). This gap suggests a need for exploration across a wider range of scales, from well-known to those less familiar, including atonal sequences, to assess the impact of musical memory on auditory perception.

The study by Radvansky et al. (1992) expanded upon Deutsch's findings by altering the note patterns within scales, revealing that the scale illusion depends not only on the pitch proximity but also on the continuity of note patterns. This investigation, which experimented with varying the number of notes, pointed out that changes in note patterns could alter perceptual grouping strategies. However, the introduction of dissonant sounds in traditionally consonant scales underscores the complexity of musical elements contributing to auditory illusions, suggesting the need for further scrutiny into how alterations in scale composition affect perception.

Additionally, Smith et al. (1982) delved into the effects of altering various elements such as pitch and timbre, and categorized participants into musicians and non-musicians. This study found that auditory experiences varied significantly with changes in timbre and melody, particularly between individuals with different levels of musical training. The choice of instruments with markedly different timbres for the experiments introduces another layer of complexity, indicating the potential influence of timbre familiarity on the perception of auditory illusions and pointing towards the value of including more commonly associated timbres, like that of the violin, in future studies. To be more specific, when a different timbre is introduced, the auditory system of the brain processes it as distinct from the original sound, thereby diminishing the inherent coherence of the auditory perception(McAdams, 2013). This implies that cognitive changes can occur depending on the complexity of timbre.

Based on these findings, our study aims to delve deeper into the impact of specific musical factors—scale mode and timbre—and individual characteristics, such as the presence of musical training and absolute pitch ability. We have designed our experiments to assess the influence of three types of scale modes—tonal-consonant, tonal-dissonant, and atonal—alongside variations in timbre(piano + piano vs. piano + violin), exploring how these factors interact to shape the perception of auditory illusions. Participants are divided based on their musical background, and a pitch discrimination test is administered to further categorize them into groups with either absolute or relative pitch.

We hypothesize that the familiarity of scale mode will influence how listeners group sounds and perceive the strength of auditory illusions, with less familiar modes weakening the illusion effect. Additionally, variations in timbre are expected to diminish the illusion's intensity by altering the cohesive auditory perception. Moreover, we posit that absolute pitch ability, beyond musical training, will significantly affect the experience of auditory illusions, with those possessing this skill demonstrating unique responses due to their refined pitch identification capabilities. This comprehensive approach enables us to examine the intricate effects of musical stimuli's scale mode and timbre on auditory illusions, alongside investigating how an individual's musical training and pitch perception skills play a role in their auditory experiences.

Methods

Participants

Forty-four healthy normal hearing individuals (aged 20-30) participated in our study. They were categorized into three groups based on musical experience and an absolute pitch ability[1]. The non-musician group, consisting of 15 members (average age 23.6, 8 males, 7 females), had no formal education at professional music institutions and an average musical training duration of 1.4 years. The second group included 16 musicians with absolute pitch (average age 25.56, 2 males, 14 females), with an average musical training duration of 11.31 years. The third group comprised 13 musicians with relative pitch (average age 26.15, 1 male, 12 females), sharing a similar training duration (average 11.08 years) as the AP group.

Classification into AP was determined by scoring 90% or higher on the absolute pitch test, while those not meeting this criterion were identified as having RP. All musicians, both AP and RP, had diverse musical training backgrounds, primarily in piano, composition, violin, flute, clarinet, and vocals. Apart from their pitch perception abilities, there were no significant differences in age or overall musical training between the AP and RP groups. All participants gave informed written consent before their participation and received financial compensation. Ethics approval was obtained from the Institutional Review Board of Seoul National University.

Stimuli

As mentioned earlier, musical stimuli were composed of three scale modes according to tonal familiarity: Tonal, Dissonant, and Atonal(see Figure 1). Tonal music closely aligns with cognitive schemas and syntactic expectations, providing the most familiar and structured auditory patterns. Dissonant music, while still adhering to some conventional patterns, introduces elements that partially defy these expectations, creating moderate tension. Atonal music, on the other hand, completely breaks away from established schemas, offering a listening experience that challenges and diverges entirely from conventional musical rules.

[1]We used the absolute pitch test developed by Diana Deutsch. https://deutsch.ucsd.edu/psychology/pages.php?i=6215

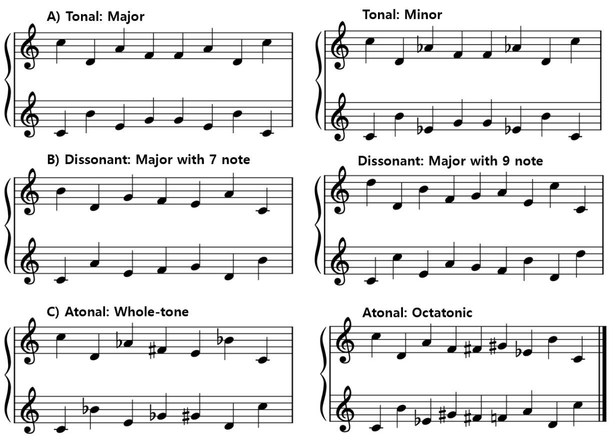

Figure 1: Scale modes; A)Tonal, B)Dissonant, C)Atonal

These include two types each of tonal consonant scales, tonal dissonant scales, and atonal scales. The tonal consonant scales are basic major and minor scales with eight notes. The tonal dissonant scales modify the arrangement and number of notes in the basic major scale, resulting in seven-note and nine-note scales. The atonal scales consist of a whole-tone scale with seven notes and an octatonic scale with nine notes. To present the scale modes across various pitch ranges, we transposed them starting from C up to B, totaling 12 keys$(C, C\sharp, D, E\flat, E, F, F\sharp, G, A\flat, A, B\flat, B)$.

These 72 stimuli were further altered into two versions to compare the effects of the timbre. One version replicated the stimuli in piano timbre for both ears, as in the existing literature. The other version played the stimuli with piano timbre in one ear and violin timbre in the other, resulting in a total of 144 excerpts. All stimuli were initially notated and transposed in Musescore (version 4.0) in MIDI format, and then the timbres were modified using acoustic piano and acoustic violin virtual instruments in Logic Pro X.

Procedure

Before the experiment, participants filled out a questionnaire to collect demographic data, including age, gender, and hearing status, along with information about their musical training. Afterward, they completed the absolute pitch test. Participants were then briefed on the experimental procedure and experimented in a soundproof room. The study used a 13-inch MacBook Pro and AirPods Pro for presenting stimuli and recording responses. The overall experimental procedure is depicted in Figure 2.

Participants first listen to a musical stimulus lasting 5 seconds (Figure 2A), and then click on the number that matches the melodic pattern heard in either the left or right ear (Figure 2B). The contour patterns for the scales were modeled after Radvansky et al. (1992), offering a clear visualization of potential melodic contours. Unlike previous studies that used alphabetic labels (A to K), our setup utilized numeric values from 0 to 9 to facilitate more intuitive matching.

'1' represented a 'Monotone' pattern with no pitch change. The 'Linear' patterns, represented by '2' and '3', were chosen to indicate a sequential pitch ascent or descent, which participants were likely to select in the event of experiencing a scale illusion, reflecting their perception of continuous pitch movement in one direction. Similarly, '4' and '5', denoting 'Arc' patterns, involved pitches rising and then falling around a central point, and were also indicative of scale illusion experiences, where participants perceived pitches organizing around a peak or trough.

'6' to '9' were 'Jagged' patterns of varying pitches, with '6-7' and '8-9' specifically matching the scale's actual contours, differentiated by the number of notes and intersection points. These options suggested a more complex perception of the melody, potentially indicating a lack of scale illusion. Participants were advised to select '0' if none of the options seemed fitting or if they were unsure.

The musical stimuli were presented in a random order, and the instructions for choosing the melody heard in either the right or left ear were counterbalanced and randomized. Each stimulus lasted about 5 seconds, and the total duration for responding to all stimuli was approximately 40 minutes.

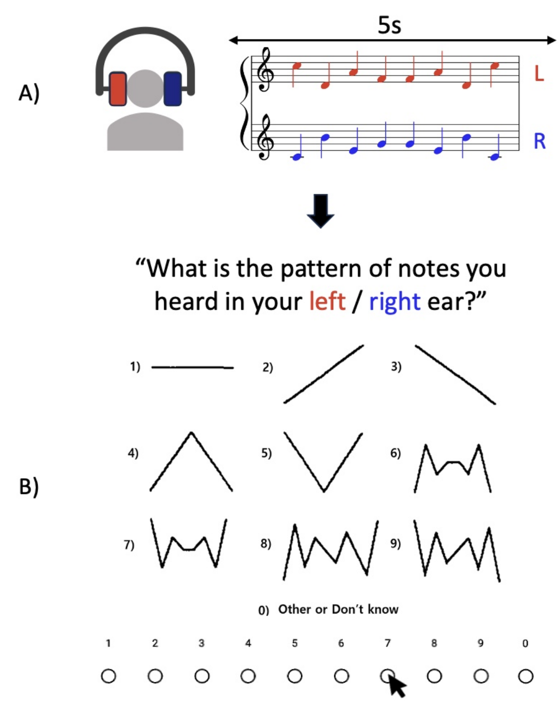

Figure 2: Experiment design. **A)** Listen to a musical stimulus. **B)** Select the scale contour patterns; 1) monotone, 2) ascending linear pattern, 3) descending linear pattern, 4) ascending then descending arc pattern, 5) descending then ascending arc pattern, 6) the jagged pattern, 7) the complimentary jagged pattern, 8) another jagged pattern, 9) another complimentary jagged pattern, 0) other & don't know.

Data Analysis

In our study, we employed a mixed-design Analysis of Variance (ANOVA) to investigate the effects of multiple factors on the scale illusion. The between-subject factors included three distinct groups: non-musicians, musicians with absolute pitch, and musicians with relative pitch. The within-subject factors were Pattern (categorized as 'jagged', 'arc', and 'linear'), Timbre (divided into 'same' and 'different'), and Mode (categorized as 'tonal', 'dissonant', and 'atonal'). The dependent variables were the frequencies of 'arc', 'linear', or 'jagged' responses. These were analyzed to assess how group distinctions and musical variables influence the perception of the scale illusion.

Results

The sounds presented to each ear were jagged patterns, but perceptions varied based on the pattern type, timbre, mode factor, and the presence of absolute pitch(Table 1).

Table 1: Descriptive statistics for responses, *M* (*SD*).

$N = 44$

	Timbre	Mode	Non-Mus. (n = 15)	RP-Mus. (n = 13)	AP-Mus. (n = 16)	Total
Jagged	Same	Tonal	9.1(8)	9.9(7.7)	13.1(9.5)	
		Diss	9.1(7.4)	8.7(7.2)	13.1(9)	
		Atonal	9.7(7.9)	8.7(6.7)	12.9(8.2)	86.7(35.9)
	Diff	Tonal	16.2(5.9)	18.6(5)	20.8(4.9)	
		Diss	16.5(6)	16(6.5)	20.6(4.2)	
		Atonal	17.3(6)	17.5(7.1)	20.4(4.1)	
Arc	Same	Tonal	10.3(6.2)	12.1(6.7)	10.7(9.4)	
		Diss	8.9(5.7)	9(6.4)	8.1(8.1)	
		Atonal	8.6(6.5)	9.6(6.2)	9.3(7.9)	39.6(29.4)
	Diff	Tonal	4.7(4.2)	4.4(4.7)	3.1(4.9)	
		Diss	4.7(3.9)	3.6(4.4)	2.2(2.9)	
		Atonal	4.4(4.8)	3.4(4.7)	2.3(3.5)	
Linear	Same	Tonal	3.4(3.3)	1.4(2.2)	0.1(0.3)	
		Diss	5.3(3.3)	4.7(4.9)	2(2.4)	
		Atonal	4.6(4)	3.9(3.8)	1.2(1.8)	13.3(14.0)
	Diff	Tonal	3.1(3.2)	0.8(1.5)	0.1(0.3)	
		Diss	2.4(2.7)	3.5(4.3)	0.4(1)	
		Atonal	2(2.7)	1.7(2.2)	0.5(0.9)	

To analyze how different groups perceive the same sound as various patterns, a mixed ANOVA was conducted with pattern type (jagged, arc, linear) as a within-subject factor and three groups (Non-Mus., RP-Mus., and AP-Mus.) as a between-subject factor. The results indicated a significant main effect for the perceived pattern type ($F(2,82) = 52.58$, $p < .001$, $\eta_p^2 = .562$), with the most responses for the jagged pattern, followed by the arc, and then the linear ($ps < .001$). However no significant main effect for group ($F(2,41) = 0.39$, $p = .680$, $\eta_p^2 = .019$) or interaction effect was found ($p = .179$). Interestingly, among musicians with absolute pitch, cases of responding with "linear" were exceedingly rare(see Figure 3). To provide a more detailed description of the effects of each factor, we will explain them according to each pattern of response.

Jagged pattern responses

In all participant groups, responses for the jagged pattern were prevalent, but statistically significant differences in responses were observed depending on whether the sounds presented to both ears had the same or different timbres ($F(1,41) = 83.58$, $p < .001$, $\eta_p^2 = .671$). Specifically, when the same timbre was presented to both ears, participants were more likely to report perceiving the Arc or Linear patterns, suggesting a strong tendency for timbre-based grouping that overrides pitch-based grouping in these scenarios. Conversely, the introduction of different timbres led to a marked increase in responses grouping according to timbre rather than pitch, underscoring the dominant influence of timbral differences in auditory perception.

Arc pattern responses

The Arc pattern was predominantly responded to as an illusion phenomenon, yet this varied depending on the Mode and Timbre factors. Firstly, a main effect due to the Mode factor was observed ($F(2,82) = 8.52$, $p < .001$, $\eta_p^2 = .172$). The Bonferroni post-hoc analysis revealed that the frequency of responses identifying as the Arc pattern was statistically significantly higher in the Tonal condition compared to the Dissonant ($p = .003$) or Atonal conditions ($p = .013$).

Secondly, a main effect related to Timbre was also observed ($F(1,41) = 60.61$, $p < .001$, $\eta_p^2 = .596$). As indicated in the results for the jagged pattern, there was a significantly higher occurrence of responses identifying as the Arc pattern when the timbre was the same.

Lastly, an interaction effect between Mode and Timbre factors was also observed ($F(2,82) = 5.97$, $p = .004$, $\eta_p^2 = .127$). As can be seen in Figure 3B. In the case of dissonance, when the timbre is the same, there is less experience of auditory illusion compared to when it is tonal. However, when the timbre is different, the effect of mode is less pronounced.

Linear pattern responses

Although responses to the linear pattern were fewer compared to the jagged or arc patterns, the effects of various factors appeared interestingly. Firstly, a main effect for mode was observed ($F(2,82) = 11.50$, $p < .001$, $\eta_p^2 = .219$). According to the Bonferroni-adjusted post-hoc analysis, responses for the linear pattern were significantly higher in the dissonant ($p = .001$) and atonal cases ($p = .044$) compared to the tonal case, with the dissonant case showing more linear responses than the atonal case ($p = .013$).

Secondly, a main effect for timbre was also observed ($F(1,41) = 16.85$, $p < .001$, $\eta_p^2 = .291$). Consistent with the previous two results, there were significantly more linear responses when the timbre was the same.

Thirdly, a significant main effect of the group was also observed in the linear responses ($F(2,41) = 7.38$, $p = .002$, $\eta_p^2 = .265$). According to the Bonferroni post-hoc correction, musicians with absolute pitch had significantly fewer responses compared to non-musicians ($p = .002$) and musicians with relative pitch ($p = .046$). This indicates that an individual's ability to perceive pitch, rather than their musical training, may play a significant role in influencing their auditory illusion in this context.

Lastly, a significant interaction between mode and timbre was observed ($F(2,82) = 9.59$, $p < .001$, $\eta_p^2 = .190$). This appears to suggest that the segments with fewer responses to dissonance in the Arc pattern are compensated by the linear pattern, with a similar observation that the difference becomes more pronounced when the timbre is consistent.

Table 2: F,p, and eta squared values from statistical analysis.

	Factors	F	p	η_p^2
Jagged	Mode	0.59	.559	.014
	Timbre	83.58***	.000	.671
	Group	2.10	.135	.093
	Mode x Group	0.78	.543	.037
	Timbre x Group	0.09	.910	.005
	Mode x Timbre	0.57	.568	.014
	Mode x Timbre x Group	0.60	.664	.028
Arc	Mode	8.52***	.000	.172
	Timbre	60.60***	.000	.596
	Group	0.22	.807	.010
	Mode x Group	0.58	.676	.028
	Timbre x Group	0.74	.482	.035
	Mode x Timbre	5.97**	.004	.127
	Mode x Timbre x Group	0.29	.882	.014
Linear	Mode	11.50***	.000	.219
	Timbre	16.85***	.000	.291
	Group	7.38**	.002	.265
	Mode x Group	2.45	.053	.107
	Timbre x Group	1.11	.338	.052
	Mode x Timbre	9.59***	.000	.190
	Mode x Timbre x Group	1.61	.180	.073

***$p < .001$, **$p < .01$

Discussion

There are various factors that can induce illusions, which can be explained by several principles of Gestalt psychology. Gestalt principles primarily deal with visual illusions, which are also applied to music. Notable principles include the law of similarity and the law of proximity. These principles were similarly observed in the results, but interestingly, a special principle of grouping unique to music was also discovered.

General findings

According to the law of similarity in Gestalt principles, visually, there is a tendency to perceive items of the same color as grouped together(Wertheimer, 1938). In music, timbre plays a similar role(Deutsch, 1999), and our results showed a significant tendency to group sounds of the same timbre, even when different timbres were presented to each ear. This grouping by timbre is an important aspect of musical perception, crucial when following the vocals in an ensemble or selectively listening to individual instruments.

Gestalt principle of proximity suggests that objects located close to each other are typically perceived as a collective group(Wertheimer, 1938). In music, proximity can be substituted by the pitch distance between sounds(Heise & Miller, 1951). That is, even if the original sounds are jagged, the higher cognitive functions of the brain synthesize these to perceive them as grouped by their closeness. The results of this study showed that even when the stimuli to each ear were jagged, they were perceived as smoothly connected arcs or linearly, with this tendency being stronger when the timbres were the same. These results are consistent with previous studies. However, a difference between vision and music was discovered; visually, linear connections are more naturally formed, and thus illusions typically appear in this form, whereas in this study, arcs were more prevalent than linear forms.

Specific characteristics in music

Why are there more illusions perceived as arcs rather than linearly in music or auditory perception? In vision, linear continuity might seem more natural, but in auditory perception, higher pitches are amplified in the auditory pathway, which is crucial for accurate language perception(Horwitz, Ahlstrom, & Dubno, 2008). In music, since the highest part often forms the main melody, when sounds occur simultaneously, attention tends to be more focused on the highest part, a phenomenon also observed at the subcortical level(Lee, Skoe, Kraus, & Ashley, 2009). This can lead to grouping the highest notes together, aiding in the separate perception of melodies that typically appear in the higher registers. The finding that arcs are more prevalent in music, especially in tonal modes compared to dissonant or atonal contexts, supports the interpretation that familiar systems or hierarchy of sound continuity are cognitively grouped into arcs of upper parts.

Furthermore, the statistical results of arc and linear perceptions showed significant findings in the interaction between mode and timbre, suggesting that different timbres lead to grouping by timbre rather than schema, and similar timbres promote schema-based grouping. In music, separating each instrument for listening is a critical aspect of music appreciation, thus prominently displaying this effect, whereas in the case of the same timbre, tonality becomes crucial, and the phenomenon of arc perception by grouping the highest parts of tonal stimuli is more pronounced. These findings highlight the complexity of how the auditory system integrates both the timbral quality and the tonal context to form a coherent musical experience. The interaction effects between timbre and mode demonstrate that our perception of music is not a mere direct reflection of the auditory stimuli but an active, reconstructive process that intricately combines sensory cues with cognitive expectations and experiences.

Interestingly, arc perception was distinctly more evident among musicians with absolute pitch. In this study, musicians with absolute pitch rarely experienced linear perception, likely due to their ability to translate the melody of the highest parts into absolute pitch names, a result first revealed in this research. Their perception appears to rely on the absolute positioning of high and low notes, indicating they navigate away from auditory illusions by identifying the actual note positions. This corresponds with previous studies indicating that in pitch memory tasks, individuals with Absolute Pitch perceive and categorize pitches based on their absolute frequencies, whereas those with Relative Pitch utilize a method of relative spatial mapping of pitch information(Schulze, Gaab, & Schlaug, 2009). This suggests that further research is needed on how individuals with abso-

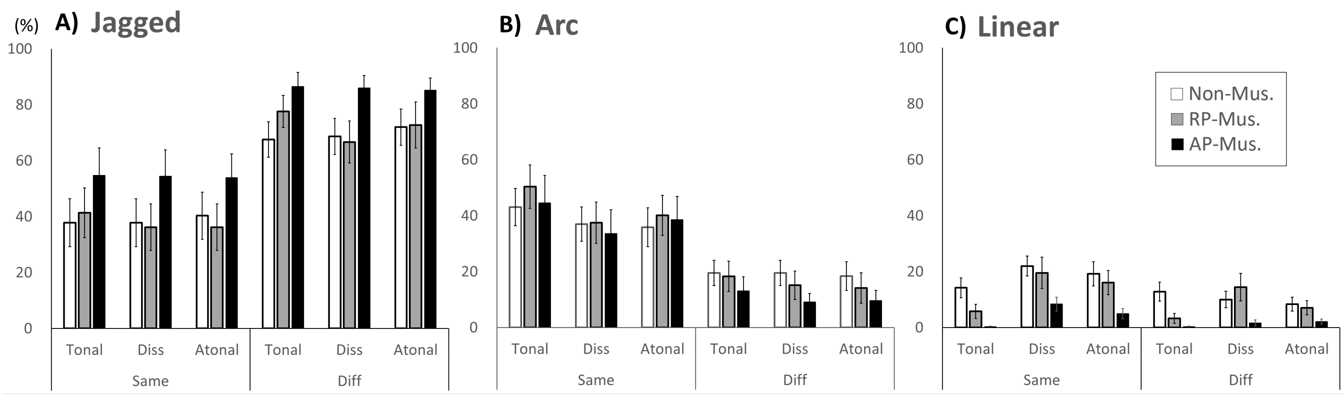

Figure 3: Frequency of responses and standard error for jagged (A), arc (B), and linear (C) patterns by group according to timbre (same, different), and mode (tonal, dissonant, atonal) conditions

lute pitch may experience different illusions.

This finding is particularly insightful when considering that the duration of musical training was similar for both absolute and relative pitch groups in our study. Prior research has indicated a difference in the perception of auditory illusions between musicians and non-musicians, attributed to the extent of musical training(Smith et al., 1982). However, our results suggest that the presence or absence of absolute pitch, rather than the duration of musical training, may have a more profound impact on susceptibility to such illusions. This underscores the importance of pitch perception capabilities in influencing how musical sequences are cognitively organized and perceived, highlighting a distinct cognitive processing pathway in individuals with absolute pitch who label tones more accurately.

This comparison becomes even more intriguing when set against visual illusions, in which visual cognition naturally groups elements that are spatially proximate, a phenomenon that most people experience(Smith et al., 1982). However, in the case of auditory illusions, there is a notable aspect that individuals who perceive pitch absolutely tend to experience auditory illusions linearly far less often. Moreover, visually, people tend to perceive intersecting lines more as linear rather than arc shapes due to good continuity, following the strategy of grouping most simply and predictably. In contrast, auditorily, there is a tendency to perceive in an arc rather than a linear form. This could be influenced by the tendency in everyday environments to hear higher pitches better than lower ones, potentially affecting how higher pitches are grouped(Stainsby & Cross, 2009b).

However, comparisons between these groups are limited by the small number of participants in our study, making it difficult to generalize the findings. Therefore, to explore the phenomenon of auditory illusions related to absolute pitch capabilities more clearly, experiments need to be conducted with larger groups and should also verify whether the phenomenon occurs consistently with different types of stimuli. Moreover, to ascertain the effects of pitch perception more definitively, it is crucial to distinguish between musical training and absolute pitch capabilities. For instance, comparing the outcomes in non-musicians with absolute pitch to those who have undergone musical training could provide valuable insights. This approach would illuminate whether the patterns observed are inherent to absolute pitch perception itself, or if they emerge as a consequence of musical training.

Conclusion

This study meticulously reviewed previous experiments on scale illusions and developed a new experimental design to explore how various musical factors and individual characteristics influence auditory illusions. By introducing variations in the traditional scale illusion stimuli, we observed predominantly jagged patterns, followed by arc, and the least being linear patterns. The significant effect of timbre was evident, as changes in timbre notably reduced the occurrence of illusions, particularly in responses to arc and linear patterns, suggesting a complex interaction with scale mode.

Our most intriguing finding was that musicians with absolute pitch were significantly less likely to respond with linear patterns, primarily perceiving the actual jagged pattern rather than experiencing auditory illusions. This suggests that absolute pitch may play a crucial role in perceiving auditory illusions, marking a departure from traditional differences observed between musicians and non-musicians. However, the limited number of participants restricts the generalizability of our findings, indicating the need for further research with larger and more diverse groups to distinguish clearly between musical training and absolute pitch abilities.

Despite these limitations, our study significantly contributes to the understanding of music cognition and auditory illusions. By investigating how various musical elements and personal traits impact auditory perceptions, our research provides valuable insights for fields such as music education, auditory science, and neuroscience. Furthermore, by setting the stage for advanced studies on auditory illusions, our research encourages further academic exploration in this intriguing field and opens avenues for new cognitive science research, particularly in comparing auditory with visual illusions.

Acknowledgments

This work was partly supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. 2021-0-0268, Artificial Intelligence Innovation Hub (Artificial Intelligence Institute, Seoul National University, 1/2) and National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2023-00219429, 1/2).

References

Daynes, H. (2011). Listeners' perceptual and emotional responses to tonal and atonal music. *Psychology of Music*, *39*(4), 468–502.

Deutsch, D. (1974). An auditory illusion. *Nature*, *251*(5473), 307–309.

Deutsch, D. (1975a). Musical illusions. *Scientific American*, *233*(4), 92–105.

Deutsch, D. (1975b). Two-channel listening to musical scales. *The Journal of the Acoustical Society of America*, *57*(5), 1156–1160.

Deutsch, D. (1999). Grouping mechanisms in music. In *The psychology of music* (pp. 299–348). Elsevier.

Deutsch, D., Henthorn, T., & Lapidis, R. (2011). Illusory transformation from speech to song. *The Journal of the Acoustical Society of America*, *129*(4), 2245–2252.

Grove, G., Sadie, S., Tyrrell, J., & Levy, M. (1980). The new grove dictionary of music and musicians. *(No Title)*.

Heise, G. A., & Miller, G. A. (1951). An experimental study of auditory patterns. *The American Journal of Psychology*, *64*(1), 68–77.

Horwitz, A. R., Ahlstrom, J. B., & Dubno, J. R. (2008). Factors affecting the benefits of high-frequency amplification.

Lee, K. M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *Journal of Neuroscience*, *29*(18), 5832–5840.

McAdams, S. (2013). Musical timbre perception. *The psychology of music*, 35–67.

Quinn, S., & Watt, R. (2006). The perception of tempo in music. *Perception*, *35*(2), 267–280.

Radvansky, G. A., Hartmann, W. M., & Rakerd, B. (1992). Structural alterations of an ambiguous musical figure: The scale illusion revisited. *Perception & psychophysics*, *52*, 256–262.

Schulze, K., Gaab, N., & Schlaug, G. (2009). Perceiving pitch absolutely: comparing absolute and relative pitch possessors in a pitch memory task. *BMC neuroscience*, *10*(1), 1–13.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *The journal of the acoustical society of America*, *36*(12), 2346–2353.

Smith, J., Hausfeld, S., Power, R. P., & Gorta, A. (1982). Ambiguous musical figures and auditory streaming. *Perception & Psychophysics*, *32*, 454–464.

Stainsby, T., & Cross, I. (2009a). The perception of pitch. *The Oxford handbook of music psychology*, 47–58.

Stainsby, T., & Cross, I. (2009b). The perception of pitch. *The Oxford handbook of music psychology*, 47–58.

Wertheimer, M. (1938). Gestalt theory.