forum acusticum 2026
8-12 September · Graz, Austria

Crossmodal correspondences between user interface sounds and material sensations: A cross-cultural investigation

Anjali Bhatara[1,2], Juliette Brun[2], Tevy Sault[2], David Morizet[2]

[1] *Ircam Amplify, Paris, France*
[2] *L'Oréal R&I, Clichy, France*

This paper is a revised version of the authors' submitted manuscript that was peer-reviewed by at least two anonymous reviewers.

Abstract

This study investigated crossmodal correspondences across cultures between acoustic properties of user interface sounds and five sensations: finish, weight, opacity, texture and physical state. Correspondences between these sensations and mean fundamental frequency (f0), melodic and loudness contour, and dissonance in 30 sounds were compared across three countries: France, China, and the USA. Results showed links between mean f0 and two sensations, finish and opacity, in France and the USA but not China. Similarly, many links were shown between sensations and melodic and loudness contour and dissonance for France and the USA, whereas correspondences were rarer for China. Potential mechanisms for this cross-cultural difference include linguistic and cultural associations. These crossmodal associations (where they exist) have important implications for product sound design.

Keywords: *crossmodal correspondences, user interface sounds, cross-cultural comparisons, sound design*

1. Introduction

Crossmodal correspondences are associations commonly reported between two modes or senses (see [1] for an overview). The present study focuses on the auditory domain and crossmodal associations between acoustic characteristics of user interface (UI) sounds and five sensations relevant for physical product development: Visual texture (rough/smooth), finish (shiny/matte), state (solid/liquid), weight (light/heavy), and opacity (transparent/opaque). We test these associations with participants in three countries: France, the USA, and China.

Corresponding author: *david.morizet@loreal.com*

1.1 Texture

The words "rough" and "smooth" can be used for visual and tactile texture as well as sounds in English, French, and Mandarin. "Roughness" in sound is a psychoacoustic measure of perceived abrasiveness [2]. An association between dissonance and visual roughness has been demonstrated in Greek participants [3], but dissonance and auditory roughness were dissociated among the Tsimane', a native Amazonian society [4], so this association may be culturally variable. Associations have also been shown between pitch height and texture. Higher pitch has been linked to the word "smooth" [5] and to tactile softness [6], though tactile smoothness was only linked to higher pitch for blind individuals [6].

1.2 Weight

High-pitched isolated tones have been linked to small size [6], [7], [8] and lightness in weight [6]. Lower pitch is linked to heavier objects [9]. Higher-pitched music was found to be associated with the words *thin* and *light* (as opposed to heavy) [5].

1.3 Opacity, Finish & State

Little evidence exists for crossmodal associations between acoustic aspects and these three sensations. However, higher pitches and spectral centroids have been linked to colors with higher luminance in humans [10] as well as in chimpanzees [11]. For opacity and finish, more transparent and shinier materials transmit or reflect more light. Hence, they could have higher luminance and be linked with higher pitches. For physical state, one study on semantic rating of sound mass demonstrated that sounds featuring short "granular" sound events in the mid-high register or slow, continuous motion were rated high on a "liquid-crystal" dimension and pieces with loud dynamics, spectral saturation and low register were high on the "voluminous-solid" dimension [12]. Hence, we can hypothesize that moving melodic contours or higher-pitched sounds may be more linked to liquid than solid.

1.4 Emotional mediation

The emotional mediation theory posits that associations between complex sounds and colors or odors may be modulated by emotion; if they each evoke a similar emotion, they will be more likely to be associated together [13] [14]. This highlights the potential importance of emotion in crossmodal correspondences. Our previous work [15] showed strong emotional effects of melodic and loudness contour direction. For example, ascending melodic contours were rated as the most positive, and constant the most negative. Much of the crossmodal auditory research to date has focused on single pitches or music with frequent changes in pitch direction. The association between direction of pitch or loudness contours and textures or sensations has not yet been thoroughly explored, and we hypothesize that it will be related to the contour's emotional impact.

1.5 Cross-cultural crossmodal correspondences

Cross-cultural studies examining crossmodal correspondences including the auditory domain are rare. One study found subtle differences between Taiwanese and North American participants in the well-known "Bouba-Kiki effect" [16]. Another study examined crossmodal associations with individual notes from Chinese musical instruments and found that Chinese participants matched higher pitches to silk than to terrycloth – the silk was both smoother and lighter than the terrycloth [17]. However, they did not examine any other culture's perception of these instruments.

1.6 Auditory dimensions of interest

The following auditory dimensions are of interest, based on [15] and an internal pilot study:

Mean fundamental frequency (f0) – a correlate of pitch height and timbral brightness [17]

Median spectral centroid – a link with timbral brightness, which may be related to visual brightness [10]

Melodic contour – this was shown to have strong emotional and functional connotations in the previous work cited above

Loudness contour – see melodic contour

Dissonance category – These short melodies or chords have one to three intervals in them, depending on note count. Using Schellenberg & Trehub's dissonance hierarchy and categorization of interval frequency ratios [18], each sound was categorized by its most dissonant interval. If it fell in the perfectly or imperfectly consonant category, (octaves, P5, P4, M3, m3, M6, m6) it was considered to be consonant, and if it fell in the dissonant category (m2, M2, m7, M7, and tritones) the sound was considered to be dissonant. Sounds with one note repeated were considered consonant.

Though the existing literature is sparse, we can hypothesize a few results: 1) High pitch has been linked with smaller size, so higher average f0 will be linked with lighter weight. 2) Dissonant (vs. consonant) intervals may be linked to roughness. 3) Ascending melodic and loudness contours should be linked to more "positive" sensations — potentially smoothness, shininess, lightness, and transparency.

2. Method

2.1 Stimuli

Novel UI sounds were created by our sound designer, based on existing sounds from household devices, largely using synthesized tones with occasional sampled instruments. They were created to have a variety of melodic and loudness contours and intervals among notes. Thirty of these were selected for the present test. Ten were chosen for each melodic contour (ascending, descending, or constant). The other variables were balanced to the best of our ability. Interactions between loudness contour and melodic contour are shown in Table 1. Dissonance was the least balanced, with the majority of the sounds being consonant (n = 21) (note that constant melodies, i.e., unchanging pitches, were coded as consonant). Mean f0 was calculated for each of the melodies (range: 156 – 1886 Hz, $M = 539$, $SD = 369$). One of the melodies had a noisy timbre with no clearly identifiable pitch, so it was excluded from the regression analyses, leaving a total of 29 stimuli, and only nine with a descending melodic contour. Total duration for these 29 sounds ranged from 431 ms to 3430 ms, $M = 1150$, SD = 503.

Table 1. Melodic and loudness contour interactions

		Melodic Contour		
		Ascending	Constant	Descending
Loudness Contour	Getting Louder	4	2	3
	Static	3	4	4
	Getting Softer	3	4	3

2.2 Participants

Participants were tested in the US (N = 301; 150 F), France (N = 308; 178 F), and China (N = 485; 253 F). They ranged in age from 18-55. Age was recorded in five-year brackets. Participant recruitment was coordinated by the agency Igonogo with international panel providers, who relied on local field agencies in each country to recruit participants meeting the predefined inclusion criteria. Participants were compensated for their time, informed consent was obtained from all participants, and all testing was done in accordance with the Declaration of Helsinki.

2.3 Procedure

Participants responded to the study on a smartphone. After the consent document, they completed a set of screening questions verifying that they were frequent users of smartphones, reported no problem in hearing, vision or cognitive function, and were currently wearing headphones or earbuds. Screening questions (counting the number of sounds at different frequencies and

reporting whether a sound came from the left or right ear) verified that they had headphones/earbuds.
During the test, they heard a UI sound, replayed it as many times as they wished, and responded to a series of questions on emotions, sensations and other associations evoked by this sound. This was repeated twice more, for a total of three sounds per person, resulting in at least 90 ratings per sound. Only the sensation data is reported here. In a series of untimed screens, they saw the question "If this sound were a sensation, what would it be?" for four of the items and "If this sound were a weight, what would it be?" for the weight question. Each was accompanied by a pair of images and a one-word description. They responded by clicking on the corresponding image (Fig. 1).

2.4 Analyses

Data preprocessing and analyses were performed in R [19]. First, the data were examined for relations among the predictors using correlations for continuous predictors, Fisher's Exact Tests for categorical predictors and ANOVAs for categorical/continuous predictors. The only significant result was a high Pearson correlation ($r = .85$) between median spectral centroid and mean f0. Because of the extensive crossmodal literature on pitch height, we chose to keep mean f0 rather than spectral centroid to be able to add to existing research.

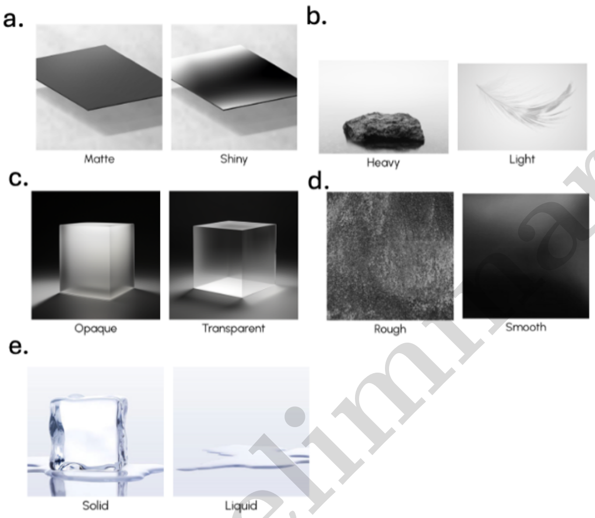


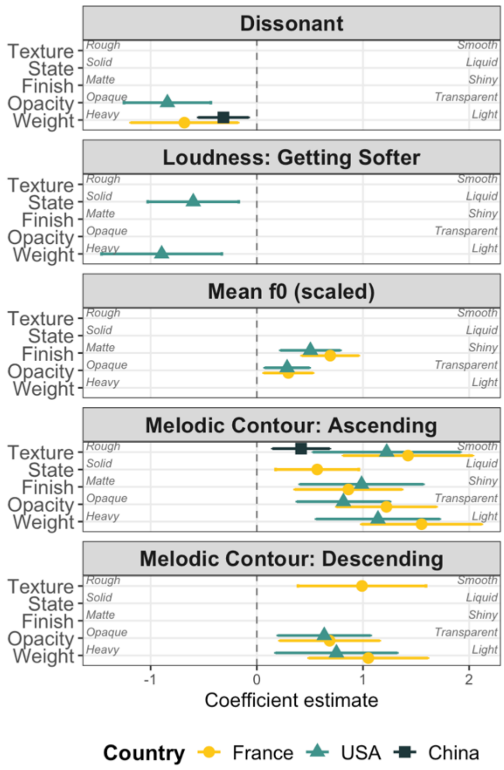
Figure 1. Test images participants saw for each sensation: a. Finish, b. Weight, c. Opacity, d. Texture, and e. State. Words were presented in French and Mandarin for the France and China groups.

Mixed-effect logistic regressions were performed with scaled mean f0 as a continuous fixed factor and three categorical fixed factors: melodic contour (reference level = Constant), loudness contour (reference level = Static), and dissonance (reference level = Consonant). Random factors were stimulus and participant. One regression was performed for each sensation/country combination using the glmer function in lme4 [20]. This model returned a boundary singular warning, indicating that the estimated variance of a random effect was at or near zero. Given that each participant heard only three of the 30 sounds, the model was tested with and without participant using likelihood ratio tests. These were nonsignificant in 14 out of 15 of the models, indicating negligible contribution of participant as a random effect in all but the China Texture model, which was significant at $p = .009$.
To keep consistent models across countries, participant was eliminated from all models. The results of the China Texture model are reported with and without participant. Because multiple models were run in each country, a Benjamini-Hochberg correction was applied to all results.

3. Results

3.1 Regression analysis

The significant results after correction are shown in Table 2 and Figure 2. Adding participant as a random factor to the model for China Texture did not change the results (melodic contour remained the only significant fixed effect), so the results without participant are reported here for all models. For dissonant (vs. consonant) sounds, *opaque* was chosen more often in the USA, and *heavy* was chosen more often in France and China. For loudness contour, the only effects were in the USA: When getting softer (relative to static loudness contour), participants chose solid and heavy more often. For mean f0, in both France and the USA, higher f0 was associated with more frequent selection of *shiny* and *transparent*. For



melodic contour, ascending melodic contours (relative

Figure 2. Significant results from the regression analyses after correction. Bars indicate 95% CI.

to constant) were associated with more frequent selection of *smooth* (in all three countries), *liquid* (in France only), and *shiny*, *transparent*, and *light* weight (in both France and the USA). Descending melodic contours were also associated with selection of *smooth* (in France) and *transparent* and *light* weight (in France and USA).

Table 2. Significant regression results from the 15 models, corrected for multiple comparisons (Asc = ascending, Desc = descending).

Country	Sensation	Acoustic factor	Estimate	Std. Err.	Z value	Corr. p value
China	Texture	Mel: Asc	0.42	0.14	3.01	0.02
	Weight	Dissonant	-0.31	0.12	-2.60	0.04
France	Finish	f0 (scaled)	0.69	0.14	5.03	<.001
		Mel: Asc	0.86	0.26	3.34	0.01
	Opacity	Mel: Asc	1.22	0.24	5.08	<.001
		Mel: Desc	0.69	0.24	2.85	0.02
		f0 (scaled)	0.30	0.12	2.50	0.05
	State	Mel: Asc	0.57	0.20	2.85	0.02
	Texture	Mel: Asc	1.42	0.31	4.56	<.001
		Mel: Desc	0.99	0.31	3.22	0.01
	Weight	Mel: Asc	1.55	0.29	5.36	<.001
		Mel: Desc	1.05	0.29	3.66	0.003
		Dissonant	-0.68	0.26	-2.64	0.04
USA	Finish	f0 (scaled)	0.50	0.14	3.50	0.005
		Mel: Asc	0.99	0.30	3.32	0.01
	Opacity	Dissonant	-0.84	0.21	-4.03	0.001
		Mel: Asc	0.82	0.22	3.64	0.003
		Mel: Desc	0.63	0.22	2.84	0.02
		f0 (scaled)	0.29	0.11	2.64	0.04
	State	Get softer	-0.60	0.22	-2.74	0.03
	Texture	Mel: Asc	1.22	0.36	3.44	0.01
	Weight	Mel: Asc	1.14	0.30	3.84	0.002
		Get softer	-0.90	0.29	-3.10	0.012
		Mel: Desc	0.75	0.29	2.56	0.04

4. Discussion

Results showed strong associations in France and the USA between several acoustic measures of UI sounds and five sensations. Two associations were shown in China as well: sounds with dissonant intervals were rated as heavier weight, and ascending melodic contour was rated as smoother, but the Chinese sample showed few crossmodal associations with the UI sounds. A few effects showed up in only France or only the USA (associations with dissonance and loudness contour, for example) but the two countries mapped together for f0 and much of melodic contour.

Regarding our initial hypotheses, some were confirmed, at least partially, whereas others were not. High pitch was associated with transparency/shininess (both linked to higher luminance) in France and the USA, but not in China. Ascending melodic contours being linked to the intuitively more positive sensations was confirmed in France and the USA for "shiny," "light," and "transparent," in France for liquid, and across all three countries for "smooth." However, high pitch was not associated with lighter weight as predicted, and dissonance had no significant roughness association. These predicted effects may be absent due to variability in sound stimuli, size of pitch range, or factors inherent to the sensation measures such as their visual rather than tactile nature.

4.1 Potential mechanisms underlying associations.

Statistical associations in nature may underlie certain correspondences: Melodic motion vs a repeated note may be suggestive of lightness and liquidity, as a lighter object is more likely to be mobile than a heavy, solid one. As discussed above, the findings connecting high pitch to shininess and transparency could be related to higher luminance in colors. Spence & Deroy speculate that the origin of this link could be an environmental association between small size, high pitch, and being found higher in the sky (birds and insects vs. large land or water animals), where sunlight comes from [21]. However, luminance in colors and light reflection or transparency are not equivalent, so it would be worthwhile to explore this link further.

Associations with dissonance and the associations between melodic motion and smoothness, transparency and shininess fit better with the emotional mediation hypothesis [14]. One of our predictions for the present study was that ascending melodic contours would be associated with the emotionally positive poles of the sensation binaries. The results of the study confirm this, while also showing that it is not only upward motion but motion in general that appears to have more positive associations, though the effects were stronger and more universal for ascending relative to descending contours.

The loudness contour effect (sounds that get softer over time being heavier and more solid for USA participants), has no obvious explanation. Auditory looming, the impression of an object approaching when a sound is getting louder and receding when a sound is getting softer [22] has no direct connotation of weight or solidity. Even so, we would predict that an approaching object would be more associated with heaviness and solidity than a receding object, which is the opposite of what was shown here.

4.2 Potential explanations for the cross-cultural difference

We find no obvious explanation for the loudness contour results being significant only in the USA. It could have to do with semantic or sound symbolism specific to English, or American cultural factors. This should be further explored in future studies.

Regarding the major differences we see between China and France/USA, one explanation could be linguistic factors. The labels under the pictures were necessary to ensure participants understood which contrast was indicated. However, these labels may have contributed to the cross-cultural result pattern. First, there may be semantic differences in the way pitches or sounds are described that could affect crossmodal correspondences. However, Mandarin is one of the languages that use the same word to describe "high" pitch and "high" in space, like English. Though this does not eliminate the possibility that other semantic differences may exist which could affect crossmodal association, the way pitch height is described in the two languages is unlikely to contribute to different perception. Words used to describe the sensations "rough" and "smooth" also have auditory connotations, and this was confirmed across the three languages via conversations among native speakers on the research team. Subtle differences in these descriptions could have influenced responses but are unlikely to have had an impact as large as shown here.

The more likely linguistic explanation is that Mandarin is a tonal language. Two of the factors we examined, mean f0 and melodic contour, depend on pitch height and patterns, and there is evidence that speakers of tonal languages process pitch in a unique way. A meta-analysis of tonal language speakers' musical ability showed improved melodic discrimination [23]. However, tone contours may be processed in a different way because of language experience—Mandarin speakers may show differential sensitivity to the four tone categories present in their language [24]. This automatic encoding by category could interfere with the development of crossmodal associations.

Alternatively, other cultural or musical factors may affect how sounds are processed, both semantically and emotionally. Traditional Chinese music varies in many ways from Western music [25], though participants have likely exposed to Western music.

4.3 Implications for product sound design

UI sounds should be designed with a number of constraints. For example, 1) they must communicate useful information about the current state of the product or device 2) they should be audible but not too obtrusive, evoking neutral to mild positive emotions and avoiding annoyance (alarms are the exception to this rule) and 3) they should be coherent with the product/device and its intended market and brand image. This third constraint is the one for which these crossmodal correspondences are especially important to keep in mind. Multisensory and emotional congruence have been demonstrated in previous studies to facilitate cognitive processing [26], potentially reducing mental load for consumers, and to enhance customer evaluations in retail settings [27]. Hence, UI sounds that are congruent with a device's physical appearance likely have positive impacts on the user experience. This study also highlights the importance of taking cultural differences into account for product design.

4.4 Limitations and future directions

This study was performed on a small set of UI sounds, with mean f0 and spectral centroid too correlated to examine separately. It would be worthwhile to examine crossmodal correspondences in noisier or atonal sounds, which could allow easier separation of these two factors (though c.f. [28] which shows that loudness-weighted spectral centroid affects pitch perception in non-harmonic sounds). Data on individual factors such as participant age and musical experience were not collected for this study (only age brackets were noted) and each could potentially play a role in perception.

Another limitation of the study is that all sensations were presented visually. This is notable especially for texture and weight, which are not normally visually determined, at least not exclusively. Follow-up studies integrating tactile perception would be useful to further explore these perceptions.

4.5 Conclusions

This study has demonstrated that, across three disparate cultures, ascending melodic contour in a UI sound is associated with smoother (visually presented) textures. It demonstrated numerous other associations among acoustic aspects and sensations, but these varied by culture. Many effects of melodic motion were significant in both France and the USA but not China: *shiny* for ascending melodies, and *transparent* and *light* for both types of melodic motion. Similarly, higher mean f0 was linked with *shiny* and *transparent* in both France and the USA but not China. Dissonant chords were associated with *heavy* in France and China, but not the USA, and other effects appeared only in France or the USA. This study's novel contributions are testing pitch height and pitch motion in short UI sounds and adding to the literature on crossmodal associations with consonance and loudness contours across three cultures. The results of this study provide implications for UI sound design (in countries where associations are demonstrated) and congruence of products and experiences across the senses.

References

[1] C. Spence, "Crossmodal correspondences: A tutorial review," Atten. Percept. Psychophys., vol. 73, no. 4, pp. 971–995, 2011, doi: 10.3758/s13414-010-0073-7.

[2] E. Zwicker and H. Fastl, Psychoacoustics–Facts and models, 2nd Edition. Heidelberg: Springer-Verlag, 1999.

[3] K. Giannos, G. Athanasopoulos, and E. Cambouropoulos, "Cross-Modal Associations Between Harmonic Dissonance and Visual Roughness," *Music Sci.*, vol. 4, p. 20592043211055484, 2021, doi: 10.1177/20592043211055484.

[4] J. H. McDermott, A. F. Schultz, E. A. Undurraga, and R. A. Godoy, "Indifference to dissonance in native Amazonians reveals cultural variation in music perception," Nature, vol. 535, no. 7613, pp. 547–550, Jul. 2016, doi: 10.1038/nature18635.

[5] Z. Eitan and R. Timmers, "Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context," Cognition, vol. 114, no. 3, pp. 405–422, Mar. 2010, doi: 10.1016/j.cognition.2009.10.013.

[6] G. Hamilton-Fletcher, K. Pisanski, D. Reby, M.

Stefańczyk, J. Ward, and A. Sorokowska, "The role of visual experience in the emergence of cross-modal correspondences," *Cognition*, vol. 175, pp. 114–121, 2018, doi: 10.1016/j.cognition.2018.02.023.

[7] E. Ben-Artzi and L. E. Marks, "Visual-auditory interaction in speeded classification: Role of stimulus difference," Percept. Psychophys., vol. 57, no. 8, pp. 1151–1162, Nov. 1995, doi: 10.3758/BF03208371.

[8] K. K. Evans and A. Treisman, "Natural cross-modal mappings between visual and auditory features," J. Vis., vol. 10, no. 1, pp. 1–12, 2010, doi: 10.1167/10.1.6.

[9] P. Walker, G. Scallon, and B. Francis, "Cross-Sensory Correspondences: Heaviness is Dark and Low-Pitched," Perception, vol. 46, no. 7, pp. 772–792, Jul. 2017, doi: 10.1177/0301006616684369.

[10] A. Anikin and N. Johansson, "Implicit associations between individual properties of color and sound," Atten. Percept. Psychophys., vol. 81, no. 3, pp. 764–777, Apr. 2019, doi: 10.3758/s13414-018-01639-7.

[11] V. U. Ludwig, I. Adachi, and T. Matsuzawa, "Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (Pan troglodytes) and humans," Proc. Natl. Acad. Sci., vol. 108, no. 51, pp. 20661–20665, Dec. 2011, doi: 10.1073/pnas.1112605108.

[12] J. Noble, E. Thoret, M. Henry, and S. McAdams, "Semantic Dimensions of Sound Mass Music," Music Percept., vol. 38, no. 2, pp. 214–242, Nov. 2020, doi: 10.1525/mp.2020.38.2.214.

[13] C. A. Levitan, S. A. Charney, K. B. Schloss, and S. E. Palmer, "The Smell of Jazz: Crossmodal Correspondences Between Music, Odor, and Emotion," Proc. 37th Annu. Meet. Cogn. Sci. Soc. CogSci 2015, no. 1, pp. 1326–1331, 2015.

[14] S. E. Palmer, K. B. Schloss, Z. Xu, and L. R. Prado-León, "Music-color associations are mediated by emotion," Proc. Natl. Acad. Sci. U. S. A., vol. 110, no. 22, pp. 8836–8841, 2013, doi: 10.1073/pnas.1212562110.

[15] A. Bhatara, X. Dartis, D. Morizet, and S. Romagny, "Designing Emotional and Intuitive Sounds for Tech: Insights From Psychoacoustics," Hum. Behav. Emerg. Technol., vol. 2025, no. 1, p. 5925146, 2025, doi: 10.1155/hbe2/5925146.

[16] Y.-C. Chen, P.-C. Huang, A. Woods, and C. Spence, "When 'Bouba' equals 'Kiki': Cultural commonalities and cultural differences in sound-shape correspondences," Sci. Rep., vol. 6, no. 1, p. 26681, May 2016, doi: 10.1038/srep26681.

[17] Y. Qi, F. Huang, Z. Li, and X. Wan, "Crossmodal Correspondences in the Sounds of Chinese Instruments," Perception, vol. 49, no. 1, pp. 81–97, Jan. 2020, doi: 10.1177/0301006619888992.

[18] E. G. Schellenberg and S. E. Trehub, "Frequency ratios and the perception of tone patterns," *Psychon. Bull. Rev.*, vol. 1, no. 2, pp. 191–201, Jun. 1994, doi: 10.3758/BF03200773.

[19] R Core Team, R: A language and environment for statistical computing. (2026). R Foundation for Statistical Computing, Vienna, Austria. [Online]. Available: https://www.r-project.org/

[20] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," J. Stat. Softw., vol. 67, no. 1, 2015, doi: 10.18637/jss.v067.i01.

[21] C. Spence and O. Deroy, "Crossmodal Correspondences: Innate or Learned?," i-Percept., vol. 3, no. 5, pp. 316–318, 2012, doi: 10.1068/i0526ic.

[22] J. G. Neuhoff, "Adaptive biases in Visual and Auditory Looming Perception," in Spatial Biases in Perception and Cognition, T. L. Hubbard, Ed., Cambridge University Press, 2018, pp. 180–192.

[23] J. Liu, C. B. Hilton, E. Bergelson, and S. A. Mehr, "Language experience predicts music processing in a half-million speakers of fifty-four languages," Curr. Biol. CB, vol. 33, no. 10, pp. 1916-1925.e4, May 2023, doi: 10.1016/j.cub.2023.03.067.

[24] Y. Li, C. Tang, J. Lu, J. Wu, and E. F. Chang, "Human cortical encoding of pitch in tonal and non-tonal languages," Nat. Commun., vol. 12, p. 1161, Feb. 2021, doi: 10.1038/s41467-021-21430-x.

[25] W. Hao, "A Comparative Study of Chinese and Western Music," Highlights Art Des., vol. 3, no. 1, 2023, https://pdfs.semanticscholar.org/1de5/4e977027aa0a1a7532d3f64c3012766a8ef1.pdf

[26] E. Guerdoux, R. Trouillet, and D. Brouillet, "Olfactory-Visual Congruence Effects Stable Across Ages: Yellow Is Warmer When It Is Pleasantly Lemony," *Atten. Percept. Psychophys.*, vol. 76, no. 5, pp. 1280–1286, 2014, doi: 10.3758/s13414-014-0703-6.

[27] L. Doucé, "The Effect of High, Partial, and Low Multisensory Congruity between Light and Scent on Consumer Evaluations and Approach Behavior," Sustainability, vol. 14, no. 9, May 2022, doi: 10.3390/su14095495.

[28] K. Velenis, A. Zacharakis, and E. Cambouropoulos, "Beyond fundamental frequency: Pitch height judgments in Indefinite-Pitch sounds," JASA Express Lett., vol. 6, no. 4, p. 044406, Apr. 2026, doi: 10.1121/10.0043569.