JKMS

Original Article
Otorhinolaryngology

Check for updates

Development of Novel Musical Stimuli to Investigate the Perception of Musical Emotions in Individuals With Hearing Loss

Jihyun Lee [ID],[1,2] Ji-Hye Han [ID],[1,2] and Hyo-Jeong Lee [ID] [1,2,3]

[1]Laboratory of Brain & Cognitive Sciences for Convergence Medicine, Hallym University College of Medicine, Anyang, Korea
[2]Ear and Interaction Center, Doheun Institute for Digital Innovation in Medicine (D.I.D.I.M.), Hallym University Medical Center, Anyang, Korea
[3]Department of Otorhinolaryngology, Hallym University College of Medicine, Chuncheon, Korea

OPEN ACCESS

Address for Correspondence:
Hyo-Jeong Lee, MD, PhD
Department of Otorhinolaryngology-Head and Neck Surgery, Hallym University Sacred Heart Hospital, Hallym University College of Medicine, 22 Gwanpyeong-ro 170-beon-gil, Dongan-gu, Anyang 14068, Republic of Korea.
Email: hyojlee@hallym.ac.kr

ORCID iDs
Jihyun Lee [ID]
https://orcid.org/0000-0002-8693-2217
Ji-Hye Han [ID]
https://orcid.org/0000-0001-9886-7590
Hyo-Jeong Lee [ID]
https://orcid.org/0000-0003-2258-0803

ABSTRACT

Background: Many studies have examined the perception of musical emotion using excerpts from familiar music that includes highly expressed emotions to classify emotional choices. However, using familiar music to study musical emotions in people with acquired hearing loss could produce ambiguous results as to whether the emotional perception is due to previous experiences or listening to the current musical stimuli. To overcome this limitation, we developed new musical stimuli to study emotional perception without the effects of episodic memory.

Methods: A musician was instructed to compose five melodies with evenly distributed pitches around 1 kHz. The melodies were created to express the emotions of *happy*, *sad*, *angry*, *tender*, and *neutral*. To evaluate whether these melodies expressed the intended emotions, two methods were applied. First, we classified the expressed emotions of melodies with selected musical features from 60 features using genetic algorithm-based *k*-nearest neighbors. Second, forty-four people with normal hearing participated in an online survey regarding the emotional perception of music based on dimensional and discrete approaches to evaluate the musical stimuli set.

Results: Twenty-four selected musical features produced classification for intended emotions with an accuracy of 76%. The results of the online survey in the normal hearing (NH) group showed that the intended emotions were selected significantly more often than the others. K-means clustering analysis revealed that melodies with arousal and valence ratings corresponded to representative quadrants of interest. Additionally, the applicability of the stimuli was tested in 4 individuals with high-frequency hearing loss.

Conclusion: By applying the individuals with NH, the musical stimuli were shown to classify emotions with high accuracy, as expressed. These results confirm that the set of musical stimuli can be used to study the perceived emotion in music, demonstrating the validity of the musical stimuli, independent of innate musical bias such as due to episodic memory. Furthermore, musical stimuli could be helpful for further studying perceived musical emotion in people with hearing loss because of the controlled pitch for each emotion.

Keywords: Musical Stimuli; Development; Hearing Loss

Generated by xmlinkpress

JKMS

INTRODUCTION

Perception of musical emotion is one of the main perceptual issues in persons with hearing loss. The nature of the perception of musical emotion remains obscure, however, since little is known about the fundamentals of musical perception in hearing-impaired people. Researchers have attempted to determine the perceptual foundations of musical emotion in clinical populations for decades.[1-6] For example, pitch perception in older adults with hearing loss is strongly correlated with discrimination in prosodic emotion, regardless of the type of hearing loss.[7] In individuals with cochlear implants (CIs), tempo cues have been used to discriminate between happy and sad music rather than mode cues.[3] In these studies, the applied stimuli were tones or excerpts from classical music to focus on the relationship between acoustical features such as the pitches of tones, tempo, and mode and the perception of musical emotion. The consideration of frequency range for musical stimuli is crucial for auditory information in hearing-impaired people; however, many studies have used existing melodies without considering the frequency range.

Episodic memory is one of the key processes in music-evoked emotions rather than in the perception of musical emotion.[8] In individuals with hearing-impairment, excerpts from familiar music may bias emotional perception based on memory of the specific melody. As long as most adults' hearing loss is not congenital, they may have experience listening to music from movies, television commercials, and famous singers in their daily lives.[9] Nonetheless, most studies (75%) investigating musical emotions have used excerpts of commercially recorded music as stimuli according to Eerola and Vuoskoski's review,[10] even in studies on people with hearing loss.[3,4,6]

Although a number of studies have contributed to the increasing knowledge of musical emotions using excerpts from familiar music,[11-13] the limitations of the existing stimuli have led few groups to develop a set of musical stimuli for research on emotions to fully understand musical emotions.[14,15] However, there are still limitations to exploring the perception of musical emotions in people with hearing loss as mentioned above.

In this study, we developed new musical stimuli to control the pitch of the melody and avoid biasing the participants with hearing loss, as may have occurred previously. We selected four categories of emotions, including happy, sad, angry, and tender, and no emotion as neutral, with these distributed in terms of valence and arousal. These emotions were chosen based on the two representative models for emotion classification, discrete and dimensional models. The former suggests that emotions can be expressed in terms of a limited number of descriptors, such as happiness, anger, sadness, disgust, surprise, and fear,[16] whereas the latter describes all emotions in two dimensions (arousal and valence).[17] Once the musical stimuli were developed, the expression of the intended emotion for each stimulus was evaluated in two ways: 1) the classification of intended musical emotion with musical features and 2) the perception of emotion in musical stimuli by an online survey in individuals with normal hearing. We hypothesized that the developed stimuli expressed the intended emotions as classified by the models. In addition, the musical stimuli represented musical emotions in people with various types of hearing loss. Furthermore, listeners with high-frequency hearing loss participated testing the applicability of the musical stimuli to study the effect of hearing loss on the perceived emotions of music in future studies.

METHODS

Stimuli

Five melodies in a monophonic piece were constructed by a professional musician as the music stimuli (**Fig. 1**, **Table 1**). The pitches of the melodies were evenly distributed around 1 kHz and are intended to express five target emotions: *happy*, *sad*, *angry*, *tender*, and *neutral*. All of these emotions, except *neutral*, were selected according to the folksonomy representation by clustering social data.[18] The musician was asked to maintain the pitches in the piece and was otherwise free to vary the musical features such as tempo, dynamics, and articulation to express the intended emotion. Five melodies expressing the 5 intended emotions were recorded using a Yamaha S90ES synthesizer (Yamaha, Japan).

Evaluation method 1: Musical features

Musical feature extraction

MIRtoolbox was employed to extract features from each musical stimulus.[19] A total of 60 features were extracted from six types of measures: full length, dynamics, rhythm, spectral, pitch, and harmony using the frame-based approach.[20] The length of each frame was 46 ms, with 50% overlap for most of the features. Features including low-energy ratio, fluctuation, tone-related features, tempo, and pulse clarity, which require longer frame length, were extracted with 2 seconds frames and 50% overlap (**Table 2**). The entire length of each musical

Fig. 1. Notations of musical melodies A to E.

Table 1. Stimuli properties of the melodies employed in the current study

ID of melody	Intended emotions	Lengths (sec)	Tempo (bpm)
A1	Happy	8.5	113
A2	Sad	17.5	112
A3	Angry	8.3	90
A4	Tender	13.1	116
A5	Neutral	9.1	150
B1	Happy	8	116
B2	Sad	18.6	116
B3	Angry	7.2	127
B4	Tender	13	120
B5	Neutral	8	119
C1	Happy	7.9	137
C2	Sad	17.3	102
C3	Angry	9.3	111
C4	Tender	10.6	126
C5	Neutral	6.5	148
D1	Happy	8.2	93
D2	Sad	17.6	115
D3	Angry	7.9	120
D4	Tender	13.2	122
D5	Neutral	11	149
E1	Happy	9.2	105
E2	Sad	19.1	131
E3	Angry	9.5	109
E4	Tender	12.6	121
E5	Neutral	10.5	109

stimulus was the first feature (F1). The dynamic features included the means and standard deviations of the RMS energy, low-energy ratio, attack time, attack slope, decay time, decay leap, and decay slope (F2–F14), whereas the rhythm features were event density, fluctuation strength, tempo, and pulse clarity (F15–F22). Brightness, centroid, zero-cross, roll-off, spread, skewness, kurtosis, and flatness (F23–F38) were constructed from the spectral features. The means and standard deviations of the first 7 Mel-frequency cepstral coefficients (MFCCs) with or without delta were also included in the timbre features (F39–F52). The pitch features included the mean and standard deviation of pitch and chromagram (F53–F56) while the final harmony features contained those of mode and entropy (F57–F60).

Musical feature selection and classification
A genetic algorithm (GA)-based feature selection method was applied to increase the efficiency of standard classifiers and find the optimal features from a set of given features, which is guided by the principles of evolution and natural genetics.[21] We used a MATLAB-based toolbox to adopt the GA method.[22]

A GA method began with a set of ten initial chromosomes, called a population. Each chromosome consisted of a randomized number of musical features. Each chromosome was assessed with the k-nearest neighbor (KNN) as a fitness function and a classifier for fitness values. We adopted the leave-one-out cross-validation technique to evaluate the classification accuracy. In each generation, two chromosomes were selected based on the best two fitness values as the parents and reproduced the next generation with the following procedures: 1) crossover with 60% probability to combine two parents, 2) mutation with 5% probability to change randomly to the individual parent to choose the next generation.[23] The above process was iterated 1,000 times with an evolved population that consisted of two parents and the remaining chromosomes.

Table 2. The extracted feature set

Category	No.	Feature	Acronym
Length	1	Total length	**Leng**
Dynamics	2,3	RMS energy	mRMS, sdRMS
	4	Low energy ratio	mLE
	5,6	Attack time	**mAT, sdAT**
	7,8	Attack slope	mAS, sdAS
	9,10	Decay time	mDT, sdDT
	11,12	Decay leap	**mDL**, sdDL
	13,14	Decay slope	mDS, **sdDS**
Rhythm	15,16	Event density	**mED**, sdED
	17,18	Fluctuation strength	**mFS, sdFS**
	19,20	Tempo	**mT, sdT**
	21,22	Pulse clarity	mPC, **sdPC**
Spectral	23,24	Brightness	**mB, sdB**
	25,26	Centroid	mC, sdC
	27,28	Zerocross	mZ, sdZ
	29,30	Rolloff	**mR, sdR**
	31,32	Spread	**mSP**, sdSP
	33,34	Skewness	mSK, **sdSK**
	35,36	Kurtosis	mK, sdK
	37,38	Flatness	**mF**, sdF
	39,40	1st MFCC + delta	**mM1, mD1**
	41,42	2nd MFCC + delta	**mM2, mD2**
	43,44	3rd MFCC + delta	**mM3**, mD3
	45,46	4th MFCC + delta	mM4, mD4
	47,48	5th MFCC + delta	mM5, **mD5**
	49,50	6th MFCC + delta	mM6, mD6
	51,52	7th MFCC + delta	mM7, mD7
Pitch	53,54	Pitch	mP, sdP
	55,56	Chromagram	mCh, sdCh
Harmony	57,58	Mode	mMo, sdMo
	59,60	Entropy	mE, sdE

Bold characters in the acronym represent features selected by a genetic algorithm method. Characters with the underbar in the feature column mean that extracted with a 2 seconds frame and 50% overlapping. Note that 'm' and 'sd' in the acronym indicate the mean and standard deviation, respectively.
MFCC = Mel-frequency cepstral coefficient.

Evaluation method 2: Behavioral experiment

Participants
Forty-four native Korean speakers with normal hearing (NH; 22 females and 22 males) aged 19–36 years (mean age 24.9 ± 4.3 years) participated in the evaluation procedure of musical stimuli. The NH group participants had normal pure-tone average thresholds (lower than 25 dB) in both ears and no neurological or cognitive issues. They were recruited through social network platforms and help-wanted advertisement sites. In addition to the NH subjects, 4 individuals with high-frequency hearing loss (HFHL; 3 females and 1 male) participated in the evaluation procedure to test the applicability of the stimuli. The pure tone audiogram of the HFHL participants showed a slope to the right, with averaged thresholds of 72.5 dB at 4 and 8 kHz (**Supplementary Table 1**). The average 4-tone (0.5, 1, 2, and 3 skHz) pure-tone audiometry thresholds of the HFHL participants were equal to or higher than 25 dB HL, and the types of hearing loss were symmetrical (interaural difference ≤ 20 dB). All participants reported no emotional or anxiety disorders and no neuropsychiatric medication during the last six months.

Procedure
The evaluation procedure for the musical stimuli was administered by an online survey which was developed with the online questionnaire tool (ko.surveymonkey.com).

For the procedure, a total of 125 stimuli (25 auditory stimuli × 5 repetitions) were created. During the survey, three kinds of questions based on the dimensional and discrete models were used for each stimulus, and they were randomly presented. Two out of each three questions (dimensional questions) were evaluated based on a 9-point Likert scale to rate the levels of arousal (1 = very calm and 9 = very excited) and valence of the perceived emotion (1 = negative and 9 = positive). To clarify the scales, we added a self-assessment manikin as a picture-oriented instrument.[24] Last, in the third question, the participants were asked to select one of the expressed emotions among five choices: *happy*, *sad*, *angry*, *tender*, and *neutral*. For the third question, the participants earned a point for a correct answer only if the answer coincided with the emotion expressed by the melody, and they were informed whether they could receive a reward after finishing the survey. The participants were also informed that they could earn additional financial rewards when they received scores higher than 70 points to help them maintain their attention during the survey. There were no time constraints on the responses. The participants completed the survey individually in a quiet environment.

k-means clustering analysis

k-means clustering analysis was conducted to classify the participants' responses to the perceived musical emotions based on the two dimensions of arousal and valence. The individual 25 objects (stimuli) were averaged across the arousal and valence ratings of all participants. Then, the average rated arousal and valence values were normalized by a scale ranging between −1 and 1. The average value for each cluster was computed, and all objects were allocated to the closest cluster. The distance between the object and each cluster was calculated using the squared Euclidean distance with the number of clusters (k) as 5. This clustering was replicated 10 times to return the solution with the lowest total sum of distances among all the replicates. The process was performed using the function "kmeans" in MATLAB 2018b (The MathWorks Inc., Natick, MA, USA).

Statistical analysis

A two-way repeated-measures analysis of variance (rANOVA) was used to assess the effect of the choice of emotions and types of melodies on both the categorical selection and dimensional ratings of the participants. This analysis was performed using the functions "fitrm" and "ranova" in MATLAB. Post hoc testing was conducted using Tukey's honestly significant difference tests, and a *P* value < 0.05 was used as the cutoff for statistically significant differences. To examine group differences using a discrete approach, the percent of correct choices of the intended emotion was calculated for each emotion and compared between groups. The Mann-Whitney *U* test was applied to account for unequal group sizes. All data are expressed as the mean ± standard error unless otherwise stated.

Ethics statement

The study protocol was approved by the Institutional Review Board of Hallym University Sacred Hospital (File No. 2021-03-003-001) and written informed consent was obtained from each participant.

RESULTS

Evaluation method 1: Musical features

Feature selection was performed to choose a proper subset of the original feature set and find the maximal prediction accuracy[25] using the machine learning technique and genetic
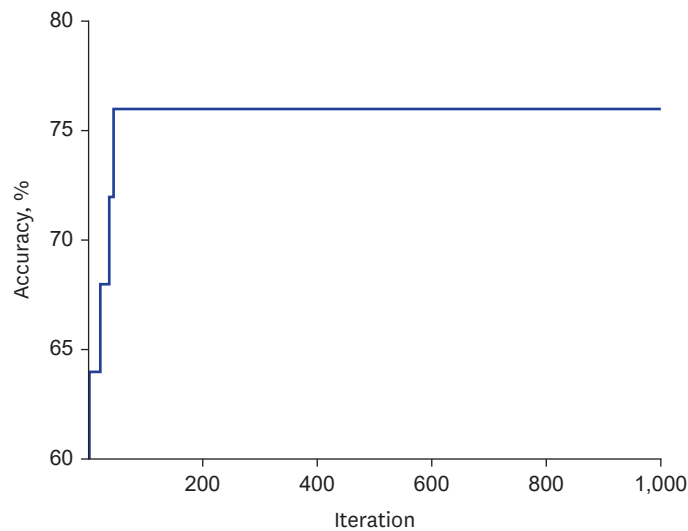
Fig. 2. Genetic algorithm evolution. The accuracy is defined by the maximum fitness of each iteration. The genetic algorithm was repeated 1,000 times to obtain the best accuracy.

algorithm. **Fig. 2** shows the accuracy against the number of iterations. Twenty-four of the 60 features were selected as the best solutions for their 76% accuracy. The selected features included Length, mAttacktime, sdAttacktime, mDecayleap, sdDecayslope, mEventdensity, mFluctuation, sdFluctuation, mTempo, sdTempo, sdPulseclarity, mBrightness, sdBrightness, mRolloff, sdRolloff, mSpread, sdSkewness, mFlatness, mMFCC1, mMFCC1d, mMFCC2, mMFCC2d, mMFCC3, and mMFCC5d (bold characters of acronym in **Table 2**).

The results for KNN classification after feature selection are shown in **Table 3**. Note that we added 'i' in front of the intended emotions, as in iHappy, to distinguish them from the emotions predicted by classification or the emotions selected by the participants. iAngry was predicted only as *angry*, while iHappy and iSad were recognized as *happy* and *sad,* respectively, with 80% accuracy. iTender was confused with *neutral* (40%), and iNeutral was confused with *happy* and *tender* (20%).

Evaluation method 2: Behavioral experiment
Dimensional approach
We examined the rating scales for consistency with the stimuli as the unit of measurement to justify averaging the subjective ratings over the participants. The arousal and valence ratings proved to be interindividual consistent (Cronbach's α: arousal = 0.97, valence = 0.91) based on the widely used criterion threshold of 0.70.[26] Since each auditory stimulus was presented five

Table 3. The predicted emotion classified by KNN

Intended emotion	Predicted emotion				
	Happy	Sad	Angry	Tender	Neutral
iHappy	**80**	0	0	0	20
iSad	0	**80**	0	20	0
iAngry	0	0	**100**	0	0
iTender	0	0	0	**60**	40
iNeutral	20	0	0	20	**60**

Emotions with "i" indicate the emotion intended to distinguish them from the predicted emotion. The unit of predicted emotion is a percentage (%). The bolded numbers indicate the percentage of predicted emotion classified by KNN corresponding to the intended emotion.

times during the survey, averaging across the participants' ratings, including five trials per auditory stimulus, is justified.

Fig. 3A shows the averaged ratings of arousal and valence by each participant. The rated arousal and valence were normalized in the range of −1 to 1 since the arousal ratings exhibited more extensive ranges than the valence ratings (1.7–6.4; arousal, 4.1–7.3; valence).

Next, we examined whether the intended emotions were clearly distinguished according to the normalized ratings of arousal and valence. An rANOVA was performed to determine the main effect of musical stimuli and the intended emotions for each normalized rating of arousal and valence. The rANOVA results showed significant interaction effects between type of melodies and intended emotions for the rated arousal (F $_{[16,688]}$ = 25.26; $P < 0.001$; partial η^2 = 0.37). We performed post hoc comparisons to determine which components were significantly different from others. The arousal ratings between all pairs of intended emotions were significantly distinguished from each other in melody D, whereas the ratings between iNeutral and iTender were not significantly different in melody A. No differences in arousal ratings between iAngry and iHappy were shown among melodies B, C, and E.

For the rated valence, rANOVA showed significant interaction effects between types of melodies and types of emotions (F $_{[16,688]}$ = 3.77; $P < 0.001$; partial η^2 = 0.08). Post-hoc tests revealed that most valence ratings were significantly different from those of other intended emotions, with several exceptions. The valence ratings of iAngry were not distinguishable from those of the other intended emotions except for iHappy. More specifically, the valence ratings of iAngry were not significantly different from those of iSad in melodies C and E. The valence of iAngry was also not distinguishable from that of iTender in all melodies. For iNeutral, a post hoc test revealed that the valence ratings were significantly different from those for iAngry in only melody C ($P = 0.019$). In addition, the valence ratings of iNeutral were not different from those of iTender for any melodies.
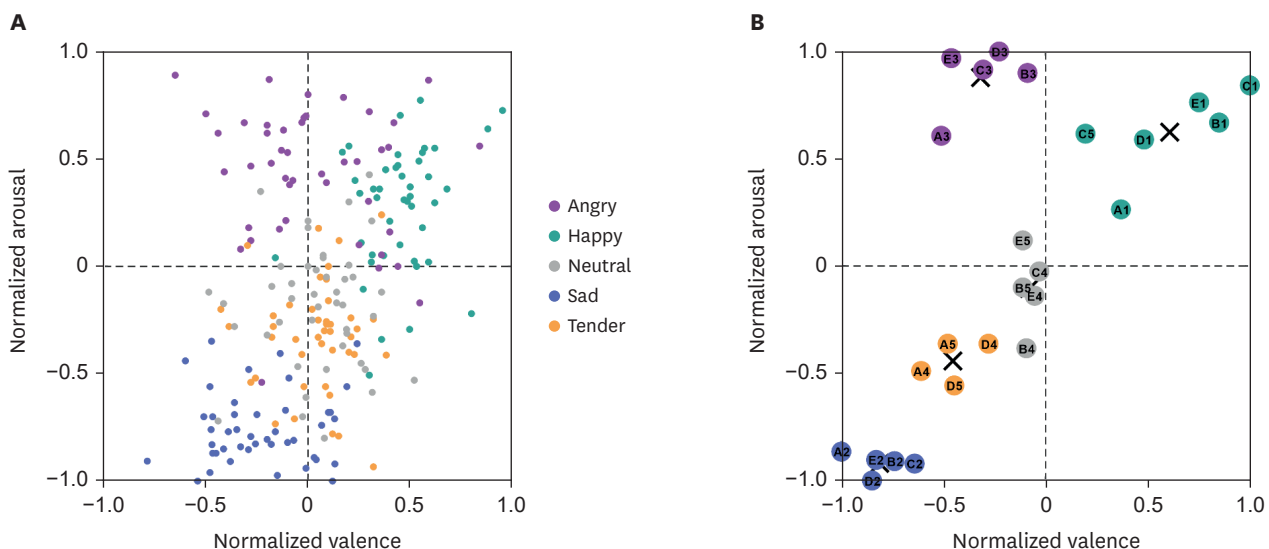

Fig. 3. Arousal and valence ratings. (**A**) The subject ratings of intended emotions on the normalized arousal and valence dimensions. Each dot represents the individual subject's mean ratings of each intended emotion for all melodies with five repeated trials. Five colors indicate five intended emotions: green – happy, purple – angry, blue – sad, yellow – tender, and gray – neutral. (**B**) *k*-means clustering analysis of normalized arousal and valence ratings. The stimuli were classified into five clusters. Each dot represents an individual stimulus. The X of each cluster indicates the centroid location. Each color denotes the same emotion as those of (**A**).

In Russell's circumplex dimensional model, each quadrant on the arousal and valence plane corresponds to basic emotions, including *happy*, *sad*, *angry*, and *tender*. Additionally, *neutral* was added by Ekman[27] in the discrete model. To determine whether the ratings of arousal and valence were placed in the representative of the quadrants of interest, we conducted a *k*-means clustering analysis (**Fig. 3B**). The melodies were classified into five clusters. The clustering accuracy for the arousal and valence scales corresponded to our classification of musical features with 76% accuracy. All melodies expressing *happy*, *angry*, and *sad* were in the first, second and, third quadrants, respectively. However, melodies that were intended to be *tender* and *neutral* were indistinguishable.

Discrete approach

The selected perceived emotions of the musical stimuli in the NH and HFHL groups are shown in **Fig 4A**. The percent corrects for each emotion in the NH group were 77.6 ± 2.4% for *happy*, 48.8 ± 3.4% for *sad*, 38.6 ± 4.6% for *angry*, 43.2 ± 3.1% for *tender*, and 54.5 ± 3.9% for *neutral*. To examine the perceptual effects of HFHL on the musical emotion for those stimuli, the percent correct on the perceived emotion for the musical stimuli were obtained in 4 subjects with HFHL (**Fig. 4A**). The rANOVA results showed no significant differences between the NH and HFHL groups. However, the overall percent correct of the intended emotions in the NH group tended to be higher than that in the HFHL group, except for iTender and the percent correct averaged across all intended emotions (**Fig. 4B**).

Table 4 shows the percentage of choice for the intended emotions as a function of stimulus type. An rANOVA was performed to examine the main effect of melody type and the choice of emotions on each intended emotion. For all intended emotions, the interaction between type of musical stimuli and choice of emotions was significantly different (iHappy: $F_{[16,688]} = 10.71$, $P < 0.001$; iSad: $F_{[16,688]} = 2.88$, $P < 0.001$; iAngry: $F_{[16,688]} = 4.20$, $P < 0.001$; iTender: $F_{[16,688]} = 4.77$, $P < 0.001$; and iNeutral: $F_{[16,688]} = 16.16$, $P < 0.001$). Post-hoc tests revealed that the intended emotions were chosen significantly more often than the other emotions for each
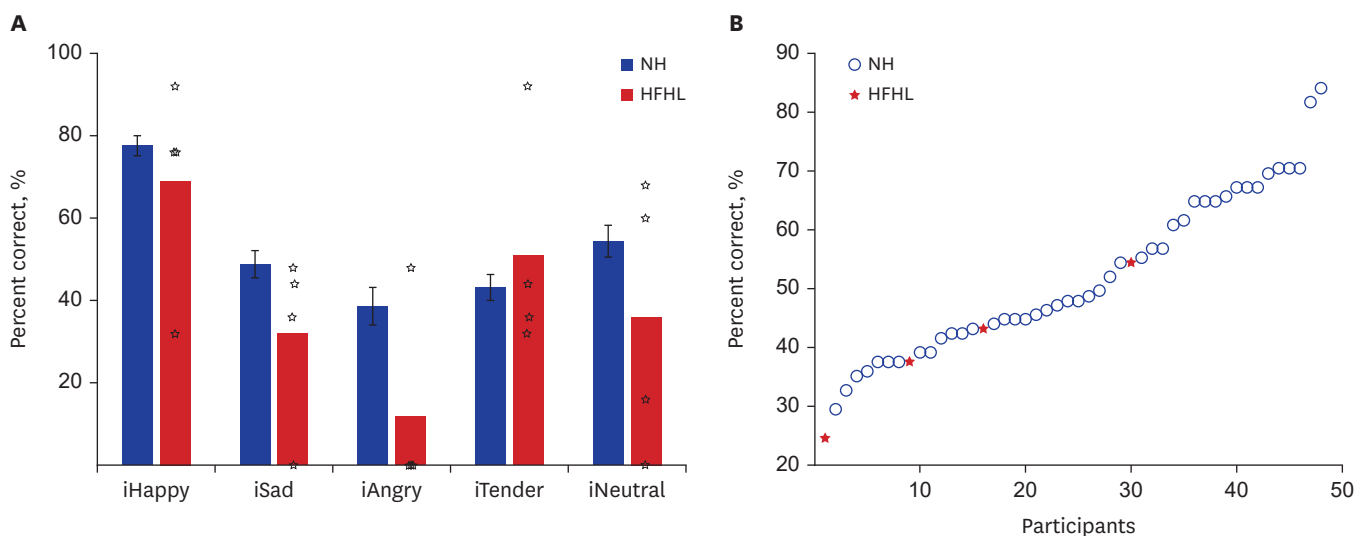
Fig. 4. The percent correct for intended emotions in the discrete approach. (**A**) The bar graphs show the average percent correct of each chosen intended emotion in all participants, including the NH group and the HFHL group. The error bars of the NH group indicate the standard error of the mean, and the pentagon shapes represent individuals with HFHL. (**B**) The sorted percent correct for the NH group (open circles) and HFHL groups (filled pentagon shapes) averaged over all intended emotions.
NH = normal hearing, HFHL = high-frequency hearing loss.

Table 4. The percentage of chosen intended emotion in each melody

Intended emotion	Melody	Choice				
		Happy	Sad	Angry	Tender	Neutral
iHappy	A	**57.7**	3.6***	2.7***	15.5***	20.5***
	B	**86.4**	1.4***	1.8***	5.5***	5.0***
	C	**91.8**	0.5***	2.3***	2.3***	3.2***
	D	**70**	0.9***	8.6***	5.9***	14.6***
	E	**82.3**	1.8***	5.0***	4.6***	6.4***
iSad	A	0.9***	**59.1**	1.4***	32.3*	6.4***
	B	5.9***	**48.6**	0***	42.7	2.7***
	C	3.6***	**46.4**	0.5***	45.5	4.1***
	D	3.2***	**43.6**	0.9***	48.6	3.6***
	E	1.8***	**46.4**	1.4***	46.4	4.1***
iAngry	A	25.5	2.7***	**37.3**	5.0***	29.6
	B	49.1	0.5***	**33.6**	2.7***	14.1*
	C	34.1	1.8***	**41.8**	3.6***	18.6*
	D	39.6	0***	**36.8**	1.4***	22.3
	E	31.4	2.7***	**43.6**	2.7***	19.6*
iTender	A	1.8***	29.1	2.3***	**44.6**	22.3
	B	14.1***	14.6***	2.3***	**51.8**	17.3***
	C	24.6	13.2**	5.9***	**38.6**	17.7*
	D	9.6***	21.8*	1.4***	**43.1**	24.1
	E	19.1*	11.3**	8.6***	**37.7**	23.2
iNeutral	A	2.3***	3.6***	6.4***	18.6***	**69.1**
	B	16.8***	2.3***	3.2***	27.3	**50.5**
	C	43.2	1.4***	14.1	8.6**	**32.7**
	D	6.4***	8.2***	1.8***	23.2***	**60.5**
	E	16.8***	0.9***	5.5***	17.3***	**59.6**

$^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.
The bolded numbers indicate the percentage of chosen emotion corresponding to the intended emotion.

type of musical stimulus. Specifically, iHappy was selected more often than other emotions for all types of stimuli ($P < 0.001$ in all), and iSad was also selected more often than other emotions for all stimuli except for *tender* (A only; $P < 0.05$). In addition, iAngry was selected significantly more often than *sad* and *tender* for all types of stimuli ($P < 0.001$). However, there were no significant differences for all stimuli with iHappy. There were also no significant differences compared to iNeutral for stimuli A and D. iTender was selected significantly more often than *angry, happy* (except for C), and *sad* (except for A), but was confused with *neutral* in stimuli A, D, and E. iNeutral was chosen significantly more often than other emotions for most stimuli (except for C with *happy* and *angry* and for B with *tender*).

DISCUSSION

This study aimed to develop new musical stimuli and to evaluate whether these stimuli accurately express the intended emotions. The properties of the developed melodies were evaluated using musical features for emotional classification and self-reports of the perceived emotions to stimuli, to support their validity and reliability. Twenty-four of the 60 acoustic features were selected by a GA method from the musical stimuli, which were used to classify emotions by a KNN method with 76% accuracy. The *k*-means clustering analysis with arousal and valence ratings revealed that the musical stimuli were classified into quadrants of interest in the plane of the dimensional model. In addition, the NH subjects selected the expressed emotions the most in the categorical choice. The stimuli were also applicable in the participants with HFHL since they also perceived the intended emotions of the musical stimuli, although the percent correct was lower than that of the NH subjects.

This study developed new musical stimuli based on two considerations: the pitches of the melodies and musical timbre. Pitch perception in hearing-impaired listeners is strongly related to emotional recognition.[7] Thus, musical stimuli were developed with evenly distributed pitches around 1 kHz for future studies with various types of HL.

A previous report showed that infants can differentiate and memorize various types of timbre even for stimuli they have never heard before.[28] A close relationship between timbre and emotion is further evident in a neurophysiological study that suggests significant correlations between the intended emotions and timbral differences.[29] Pitch and timbre were considered the main features of music in our study.

A digital piano was chosen based on these two considerations. It covers the full range of pitch, even over 2 kHz, and is emotionally neutral. It has also been suggested that the timbral features of pianos are emotionally neutral which indicates the minimal effect on the judgment of musical emotion.[30] Furthermore, using only one instrument may implicate no bias to the auditory memory incurred by timbre.

Over the past decade, studies on the perceived emotions of music have demonstrated that a computational perspective can explain musical emotions.[31] The main goal of the computational approach is to explore the relationships between emotions and acoustic features extracted from music. To this end, research groups have attempted to predict emotional categories using machine learning techniques such as KNN.[32] The KNN with feature selection precludes redundant and irrelevant features, thereby vastly improving classification performance.[32] In our study, the classification accuracy with KNN was enhanced using GA (76%; **Fig. 2**) compared to classification without feature selection (56%, data not shown).

The result of the classification of the selected musical features indicates that iAngry, iHappy, and iSad were mostly recognized as intended (**Table 3**). However, iTender was confused with *neutral*, and iNeutral was frequently selected as *tender* and *happy*. Previous meta-analysis results revealed similar decoded accuracy patterns, where anger and sadness were decoded with the highest accuracy, followed by fear and happiness, and tenderness had the lowest accuracy.[33] Considering that the meta-analysis results are in line with our classification data, we infer that the musical stimuli used in this study have similar patterns of acoustic features.

For the rANOVA results of both arousal and valence ratings, most intended emotions were significantly distinguished from other emotions in all melodies. The arousal ratings for iAngry were not significantly different from those for iHappy, whereas the valence ratings for iAngry were significantly different from those for only iHappy. Regardless of the nonsignificant differences in iAngry from others, the k-means clustering analysis depicted that the melodies expressing *angry* separated from those expressing other emotions (**Fig. 3B**). This classification between iAngry and other emotions was due to contrast arousal and valence levels.

The participants could not discriminate between iTender and iNeutral in melody A in terms of arousal ratings but did differentiate iTender from iNeutral in terms of the valence ratings for all melodies. The k-means clustering analysis also showed that melodies expressing *tender* and *neutral* were indistinguishable from each other (**Fig. 3B**). iTender might be confused with *neutral* since the two were rated mid-level for both arousal and valence whereas other emotions were rated higher or lower than those.

iSad was exceedingly confused with *tender*. A previous study that assessed the effect of pitch variation on arousal and valence ratings revealed that pitch level was positively correlated with valence ratings but negatively correlated with arousal ratings.[34]

In this study, iSad, which has a negative valence, was considered *tender*. Meanwhile, *tender* has a positive valence with a similar level of arousal to that of *sad*, while *sad* was not selected in response to iTender. Considering that higher pitches were distributed around 1 kHz in our study, these findings lead us to regard the possibility that high pitches might cause biased choices toward positive emotions as mentioned above.

In this study, *angry* was confused with *happy* in that the *happy* was not differentiated from *angry* in the rating of iAngry for all melodies. Nonetheless, the *angry* was selected significantly less often than *happy* in iHappy. Such distinction in selection between iHappy and iAngry would be related to the tendency toward positive valence. Another discrepancy was shown between the result of the discrete approach (cf. **Fig. 4**, *angry*) and the classification of iAngry using musical features (cf. **Table 3**). In terms of musical features, the musician in this study might have created five angry musical stimuli using similar acoustic features. This led to the iAngry stimuli being recognized as angry completely in the emotional classification using acoustic features, while the participants perceived the iAngry stimuli as positive because of their high pitch. Interestingly, iNeutral was mainly perceived as *neutral*, which was in line with a previous study suggesting that *no expression* seemed easy to convey.[35]

The musical stimuli were developed for our main motivation to assess differences between hearing loss and normal-hearing individuals in perceiving the emotions of music. To this end, we performed a pilot study with a small subgroup of HFHL participants, aiming to test whether they perceived musical emotion less than people with NH because of their impairment in specific frequency areas, without musical bias such as episodic memory. The average percentage at which the intended was chosen emotion was lower in the HFHL group than in the NH group (**Fig. 4A**). Therefore, the average percent correct across all emotions was lower in HFHL group than in the NH group (**Fig. 4B**). Considering the small size of the HFHL group, the group comparisons in this study are difficult to generalize. Nevertheless, the tendency of the results was similar to those of previous study regarding the perception of musical emotion in hearing-impaired children.[36] In addition, all of the individuals with HFHL participated in the test without claiming any difficulty until the end.

Taken together, our data suggest that the newly developed musical stimuli have the potential to be applicable for future studies regarding the perception of musical emotions in people with hearing loss.

In conclusion, novel musical melodies were developed and evaluated for their accessibility as ecologically valid stimuli to investigate the behavioral and neural processing of perceived musical emotion. The findings indicate that the intended emotion can be delivered by listening to the stimuli, yet some variability exists among the listener responses. With only a few subjects with hearing loss being recruited, we showed that these musical stimuli were able to reflect the intended emotions in this clinical population, revealing their suitability for use in future studies. This work provides an example for developing musical stimuli for researchers who study musical emotional perception in hearing-impaired people.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

Supplementary Table 1
Pure Tone Thresholds of Hearing loss group

Click here to view

REFERENCES

1. Singh G, Liskovoi L, Launer S, Russo F. The Emotional Communication in Hearing Questionnaire (EMO-CHeQ): development and evaluation. *Ear Hear* 2019;40(2):260-71.
   PUBMED | CROSSREF

2. Picou EM, Singh G, Goy H, Russo F, Hickson L, Oxenham AJ, et al. Hearing, emotion, amplification, research, and training workshop: current understanding of hearing loss and emotion perception and priorities for future research. *Trends Hear* 2018;22:2331216518803215.
   PUBMED | CROSSREF

3. Hopyan T, Manno FA 3rd, Papsin BC, Gordon KA. Sad and happy emotion discrimination in music by children with cochlear implants. *Child Neuropsychol* 2016;22(3):366-80.
   PUBMED | CROSSREF

4. Pralus A, Belfi A, Hirel C, Lévêque Y, Fornoni L, Bigand E, et al. Recognition of musical emotions and their perceived intensity after unilateral brain damage. *Cortex* 2020;130:78-93.
   PUBMED | CROSSREF

5. Ambert-Dahan E, Giraud AL, Sterkers O, Samson S. Judgment of musical emotions after cochlear implantation in adults with progressive deafness. *Front Psychol* 2015;6:181.
   PUBMED | CROSSREF

6. Giannantonio S, Polonenko MJ, Papsin BC, Paludetti G, Gordon KA. Experience changes how emotion in music is judged: Evidence from children listening with bilateral cochlear implants, bimodal devices, and normal hearing. *PLoS One* 2015;10(8):e0136685.
   PUBMED | CROSSREF

7. Mitchell RL, Kingston RA. Age-related decline in emotional prosody discrimination: acoustic correlates. *Exp Psychol* 2014;61(3):215-23.
   PUBMED | CROSSREF

8. Juslin PN, Västfjäll D. Emotional responses to music: the need to consider underlying mechanisms. *Behav Brain Sci* 2008;31(5):559-75.
   PUBMED | CROSSREF

9. Feltner C, Wallace IF, Kistler CE, Coker-Schwimmer M, Jonas DE. Screening for hearing loss in older adults: updated evidence report and systematic review for the US Preventive Services Task Force. *JAMA* 2021;325(12):1202-15.
   PUBMED | CROSSREF

10. Eerola T, Vuoskoski JK. A review of music and emotion studies: approaches, emotion modelss, and stimuli. *Music Percept* 2013;30(3):307-40.
    CROSSREF

11. Schellenberg GE, Krysciak AM, Campbell RJ. Perceiving emotion in melody: interactive effects of pitch and rhythm. *Music Percept* 2000;18(2):155-71.
    CROSSREF

12. Lundqvist LO, Carlsson F, Hilmersson P, Juslin PN. Emotional responses to music: experience, expression, and physiology. *Psychol Music* 2009;37(1):61-90.
    CROSSREF

13. Sammler D, Grigutsch M, Fritz T, Koelsch S. Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology* 2007;44(2):293-304.
    PUBMED | **CROSSREF**

14. Vieillard S, Peretz I, Gosselin N, Khalfa S, Gagnon L, Bouchard B. Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cogn Emotion* 2008;22(4):720-52.
    CROSSREF

15. Paquette S, Peretz I, Belin P. The "Musical Emotional Bursts": a validated set of musical affect bursts to investigate auditory affective processing. *Front Psychol* 2013;4:509.
    PUBMED | **CROSSREF**

16. Ekman P, Friesen WV. Constants across cultures in the face and emotion. *J Pers Soc Psychol* 1971;17(2):124-9.
    PUBMED | **CROSSREF**

17. Russell JA. A circumplex model of affect. *J Pers Soc Psychol* 1980;39(6):1161-78.
    CROSSREF

18. Laurier C, Sordo M, Serra J, Herrera P. Music mood representations from social tags. In: Proceedings of the 10th International Society for Music Information Retrieval Conference, ISMIR 2009; October 26-30, 2009; Kobe, Japan. Montreal, Canada: ISMIR; 2009, 381-6.

19. Lartillot O, Toiviainen P. MIR in Matlab (II): A toolbox for musical feature extraction from audio. In: Proceedings of the 8th International Conference on Music Information Retrieval, ISMIR 2007; September 23-27, 2007; Vienna, Austria. Montreal, Canada: ISMIR; 2007, 127-30.

20. Tzanetakis G, Cook P. Musical genre classification of audio signals using geometric methods. In: Proceedings of the 18th European Signal Processing Conference; August 23-27, 2010; Aalborg, Denmark. Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers (IEEE); 2010, 497-501.

21. Girsang AS, Manalu AS, Huang KW. Feature selection for musical genre classification using a genetic algorithm. *Adv Sci Technol Eng Syst J* 2019;4(2):162-9.
    CROSSREF

22. Farzaneh M. MATLAB Central File Exchange. Feature selection in classification using genetic algorithm. https://www.mathworks.com/matlabcentral/fileexchange/74105-feature-selection-in-classification-using-genetic-algorithm. Accessed September 13, 2021.

23. Goldberg D. *Genetic Algorithms in Optimization, Search and Machine Learning*. Reading, MA, USA: Addison-Wesley Publishing Company; 1988.

24. Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry* 1994;25(1):49-59.
    PUBMED | **CROSSREF**

25. Yang YH, Chen HH. *Music Emotion Recognition*. Boca Raton, FL, USA: CRC Press; 2011.

26. Bland JM, Altman DG. Cronbach's alpha. *BMJ* 1997;314(7080):572.
    PUBMED | **CROSSREF**

27. Ekman P. Are there basic emotions? *Psychol Rev* 1992;99(3):550-3.
    PUBMED | **CROSSREF**

28. Trehub SE, Endman MW, Thorpe LA. Infants' perception of timbre: classification of complex tones by spectral structure. *J Exp Child Psychol* 1990;49(2):300-13.
    PUBMED | **CROSSREF**

29. Goydke KN, Altenmüller E, Möller J, Münte TF. Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. *Brain Res Cogn Brain Res* 2004;21(3):351-9.
    PUBMED | **CROSSREF**

30. Chau CJ, Wu B, Horner A. Timbre features and music emotion in plucked string, mallet percussion, and keyboard tones. In: Proceedings of the 40th International Computer Music Conference, ICMC 2014 and 11th Sound and Music Computing Conference, SMC 2014 - Music Technology Meets Philosophy; September 14-20, 2014; Athens, Greece. Ann Arbor, MI, USA: Michigan Publishing; 2014, 982-9.

31. Laurier C, Lartillot O, Eerola T, Toiviaine P. Exploring relationships between audio features and emotion in music. In: Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music; August 12-16, 2009; Jyväskylä, Finland. Jyväskylä, Finland: European Society for the Cognitive Sciences of Music; 2009, 260-4.
    CROSSREF

32. Saari P, Eerola T, Lartillot O. Generalizability and simplicity as criteria in feature selection: application to mood classification in music. *IEEE Trans Audio Speech Lang Process* 2011;19(6):1802-12.
    CROSSREF

33. Juslin PN, Laukka P. Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol Bull* 2003;129(5):770-814.
    PUBMED | **CROSSREF**

34. Jaquet L, Danuser B, Gomez P. Music and felt emotions: How systematic pitch level variations affect the experience of pleasantness and arousal. *Psychol Music* 2012;42(1):51-70.
   CROSSREF

35. Gabrielsson A, Juslin PN. Exmotional expression in music performance: between the performer's intention and the listener's experience. *Psychol Music* 1996;24(1):68-91.
   CROSSREF

36. Darrow AA. The role of music in deaf culture: deaf students' perception of emotion in music. *J Music Ther* 2006;43(1):2-15.
   PUBMED | **CROSSREF**