The Role of Listening Expertise, Attention, and Musical Style in the Perception of Clash of Keys

Reinhard Kopiez and Friedrich Platz
Hanover University of Music and Drama,
Hanover, Germany

This explorative study investigates the perception of clash of keys in music. It is a replication and extension of an earlier study by Wolpert (2000) on the perception of a harmonic (bitonal) manipulation of a melody and accompaniment. We investigated (a) how reliable results were, (b) how results would change if listeners' attention changed from nondirected (NDL) to directed listening (DL), and (c) whether the perception of clash of keys is influenced by the musical style of the particular composition. Participants included 101 expert listeners and 147 nonexpert listeners who evaluated music of four different styles in two versions each (original and with a pitch difference of 200 cents between melody and accompaniment). On the whole, expert listeners noticed the clash of keys significantly more often than did nonexperts (NDL: 49.30% vs. 9.30%; DL: 78.00% vs. 46.90%). For NDL, the perception of clash of keys differed between musical styles and decreased from classical to rock 'n' roll and from pop to jazz. Differences in responses are mainly explained by acculturation effects (listening expertise, attention, musical style, and familiarity with the particular piece).

Received May 19, 2008, accepted November 5, 2008.

Key words: attention, sensory pleasantness, perception, tonality, listening

The organization of sounds into connected patterns or groups of elements seems to be a fundamental feature of human music perception and listening experience, whereby a particular musical style leads to internal structural regularities within that idiom (Krumhansl, 1990). In Western music, one of these structural regularities is the internal processing system for pitch structures, which is characterized by a tendency to organize sequences of pitches or harmonies in a hierarchical way around a central reference pitch (the so-called tonic). In music theory, sophisticated systems for the description of hierarchical pitch and chord relationships have been developed over the last 4 centuries (for an historical overview see Dahlhaus, 1990; Thomson, 1999). These concepts are centered on the terms of *harmony* and *tonality*. Dahlhaus (2001) described harmony as the "simultaneous combination of notes . . . to produce chord progressions . . . and also prescriptively to denote a system of structural principles governing their combination" (p. 858). As a result, sequences of chord or pitch progressions can generate the impression of a referential pitch class or tonic. As Dahlhaus (1990) emphasized, the perception of tonality or tonal harmony rests on at least two assumptions: "first that a triad constitutes a primary, direct unit; and second, that the progression of chordal roots establishes the key" (p. 3). This harmonic organization of music is "one of the main conceptual categories in Western musical thought" (Hyer, 2001, p. 583) and results in the construction of tonality which is often used as a synonym for *key*. At the beginning of the 20th century, composers such as Ives, Stravinsky, or Milhaud showed an increased interest in the combination of conflict tonalities and started using superimposed voices with distinct tonalities (Whittall, 2001).

The first experimental approaches to the question of perceptual relevance of bitonality or clash of keys was conducted by Theodor Billroth (1894). The famous surgeon, enthusiastic musician, and close friend of Brahms used the clash of keys paradigm as a test for tonal perception in his informal "Bridesmaids' test" (Billroth, 1894). In his treatise on musical ability, he assumed that deficits in the perception of harmony are the result of attentional deficits and are not indicators of amusia. To test this hypothesis, he used a participant with presumed tonal deafness but a strong interest in attending concerts and operas. Billroth played the beginning of the bridesmaids' song from the opera Der Freischütz (Act III) by C. M. von Weber on a piano in 2 versions: (a) with the melody in the original key of F# major and the accompaniment in F major; (b) with

Music Perception volume 26, issue 4, pp. 321–334, issn 0730-7829, electronic issn 1533-8312 DOI:10.1525/mp.2009.26.4.321

the melody in G major and the accompaniment in F major. The participant recognized the tune but did not perceive the clash of keys and preferred the bitonal first version over the second. Based on this result, Billroth concluded that there might be a large proportion in every audience that is resigned to out of tune performances.

From the perspective of cognitive psychology, in the last 20 years there also has been a large amount of studies centered on the question of the perceptual reality of pitch and key organization (for an overview see Krumhansl, 1990; Thomson, 1999). All studies have been based on the attentive and directed listening of participants. However, support for the assumption of different levels of information processing in music was already given by the writings of the German philosopher Gottfried Wilhelm Leibniz. In his Monadologie, Leibniz (1714/1991) distinguished between two different pathways from sensation to consciousness: The first way is called "perception," a non or preattentive and involuntary sensation characterized by unconscious information processing; the second way, called "apperception," denotes the attentive and conscious processing of information. Arguments for the importance of both nonattentive and attentive processes in human auditory perception also have been given by recent studies on auditory information processing. For example, Vitevich (2003) investigated "change deafness" for voices (the inability to detect changes between two speaker voices). He found that 40% of the participants did not detect the change in speaker. He assumed that the allocation of attention may influence the detection of changes. In an extensive study on listening behavior, Behne (1990) found that "hearing with shared attention" (so-called "diffuse hearing") is widely distributed among young people and is still increasing in everyday life. This finding was confirmed by recent studies that showed that music often is used for the accompaniment of other activities (Sloboda, O'Neill, & Ivaldi, 2001) and is widely perceived in terms of nonanalytical categories of surface features, such as volume, sound, and so forth (Schramm, 2005). Neurophysiological studies also have given clear evidence for the effects of auditory attention on the activation of different brain networks: Jäncke, Buchanan, Lutz, and Shah (2001) found that the degree of activation in Heschl's gyrus and planum polare depends on the direction of attention in a dichotic listening task. To summarize, these results give support for the hypothesis that participants are highly sensitive to listening instructions given in an experimental task and will change their listening strategy when directed instructions are given.

Wolpert's (2000) Study of Nondirected Listening to Key in Music

The original study by Wolpert (2000) was motivated by the question of which implicit musical schemata would be used when listeners were given no explicit instructions on key. In other words, the focus was not "what people can hear, given specific directions, but what they do hear" (p. 225). Wolpert's interest in the question of nondirected listening can be traced back to the 1990s. In a forced-choice experiment based on a matching task for melodies varying in harmonization, instrumentation, and accompaniment, Wolpert asked participants to compare a model melody with two modifications and to decide, among other things, which of the two modifications was most like the model (Wolpert, 1990). The participants' attention was not directed to any one aspect of the music. When given the option to choose between (a) a version with the same melody and harmonic accompaniment as the model, but with different instrumentation, and (b) a version with the same melody and instrumentation, but the accompaniment in C and melody in G (bitonal version), 95% of the nonmusicians did not notice the clash of keys in the bitonal version and decided it was identical to the model. This was the case for only 5% of the musicians. Wolpert concluded that "musicans and nonmusicians do not . . . listen by the same rules" (p. 103). However, it remained open as to whether harmonic accompaniments less consonant with the tonic (as, for example, with a difference in key of a major second) would result in a more salient clash of keys. Wolpert followed with a second study (Wolpert, 2000) in which musicians ($n = 10$, professional performers, arrangers, or composers) and nonmusicians ($n = 40$, undergraduate students of psychology; 7 participants with instrumental experience of 6 years or more) listened to three versions of the jazz standard, "You Make Me Feel So Young." In the first of two additionally modified versions, the accompaniment was transposed up by a major second; in the second version the accompaniment was transposed down by the same interval. A nondirected listening paradigm was used, and participants were asked to write down in free comments "what differences, if any, they heard among the excerpts" (p. 227). Answers were coded by means of content analysis into three categories: a response was considered a *match* if words such as "out of tune," or "too high" were used; a response was considered a *possible match* if words such as "lower" or "higher" were used, even if the participant mentioned other differences that did not exist; a response was considered a *mismatch* only if differences that did not exist (e.g., surface features such as changes in tempo or length) were mentioned. Wolpert scored 100% of the musicians' answers as matches compared to only 20% of the nonmusicians.

Critical Appreciation of Wolpert's Study

While these findings are spectacular (the listening task could have been too easy for the experts, or the participants had exceptional hearing skills), a number of methodological issues also limit the interpretation of Wolpert's (2000) results. First, there was a bias in her use of more nonmusicians (n = 40) than musicians (n = 10) as participants. Second, the sound quality of the versions differed and contained technical problems (i.e., exaggerated sound level in the low frequency range of the version transposed down and crackles on the accompaniment track, which was transferred from LP). Third, there was no group with a directed listening condition. Fourth, no hypothesis was given for an effect of the direction of transposition (major second up or down), and, finally, musical style was limited to one genre (jazz). However, as can be derived from Pascall (2001), "style manifests itself in characteristic usages of form, texture, harmony, melody, rhythm and ethos" (p. 638) and thus, could be an influential factor for the perception of clash of keys.

Rationale and Research Questions

Based on these critical points, our exploratory study tried to replicate the findings of Wolpert (2000) while extending her study by using stimuli from a broader stylistic range. We also included directed listening as an additional listening condition.

The main research questions were as follows: The first question was in accordance with Wolpert (2000) and asked whether it can be confirmed that expert listeners would notice the clash of keys more frequently compared to nonexperts in the condition of nondirected listening. The second question asked whether the rate of perception of clash of keys would increase from nondirected listening to directed listening for both expert and nonexpert listeners. The third question considered whether there are differences in the perception of clash of keys between musical styles. It was assumed that the perception of clash of keys is more easily perceived in music whose language has tonal coherence as its basic principle, such as in 18th century music. In jazz music, it might be more difficult to detect tonal transposition due to its altered harmonic style. Pop music and rock 'n' roll would lie within the range of these two styles. The fourth question focused on whether familiarity with a particular piece increases the rate of perception of clash of keys. It was assumed that familiarity would increase perception of key conflicts.

Method

A 2 × 2 × 4 × 2 study design was used with *listening expertise* (experts/nonexperts) and *focus of attention* (nondirected/directed listening) as between-participant variables and *musical style* (4 styles) and *version* (2 versions) as within-participant variables. Familiarity with the particular piece was used as a control variable.

Participants

A total of 248 participants attended the experiment. The biographical data of the four groups of participants were as follows: Group 1 (experts, nondirected listening, n = 69; age range = 18-25 years; mean instrumental experience = 14.0 years, SD = 4.1; instruments: 21 woodwind, 20 string, 9 piano, 8 brass, 5 guitar, 4 percussion, 2 voice); Group 2 (experts, directed listening, n = 32, age range = 18-25 years; mean instrumental experience = 10.7 years, SD = 4.3; instruments: 13 piano, 9 woodwind, 3 string, 1 brass, 2 percussion, 2 guitar, 2 accordion); Group 3 (nonexperts, nondirected listening, n = 115, age range = 14-18 years, 26.8% with instrumental training of more than 5 years); Group 4 (nonexperts, directed listening, n = 32, age range = 14-18 years, 30.4% with instrumental training of more than 5 years). All expert listeners were music students from the Hanover University of Music and Drama, Germany.

Materials

SOUND EXAMPLES

First, the original sound examples from Wolpert's (2000) study were used. Wolpert used the audio files from a "Music Minus One" recording with an additional track recorded from a professional jazz singer. The bitonal versions in her study were produced by pitching the accompaniment a major second up (+200 cents) or down (−200 cents). We eliminated the deficiencies in recording quality of the three versions of "You Make Me Feel So Young" with the aid of audio restoration software (Dart Pro XP). Sound differences between the original and transposed versions were minimized by equalization. Second, audio files of three new pieces representing the musical styles of rock 'n' roll, pop, and classical

TABLE 1. Musical Examples Used in the Listening Experiment.

Title/Composer/Performer	Versions	Modification		Duration (s)	Style
You Make Me Feel So Young	Original*			34	
Music: J. J. Mylow	Accompaniment higher*	Accompaniment	200 c ↑		Jazz
Lyrics: M. Gordon	Accompaniment lower*	Accompaniment	200 c ↓		
You've Got a Friend	Original			44	
Music & Lyrics: C. King	200 c	Vocals	100 c ↓		Pop
Performer: Soul Control		Accompaniment	100 c ↑		
Walking on Sunshine	Original			37	
Music & Lyrics: K. C. Rew	200 c	Vocals	100 c ↓		Rock 'n'
Performer: Soul Control		Accompaniment	100 c ↑		Roll
Trumpet Concerto, by J. Haydn, 2nd	Original			38	
Movement, "Andante" (HOB. VIIe:1.)	200 c	Trumpet	100 c ↓		Classical
[Piano reduction]		Accompaniment	100 c ↑		
Performer: Friedrich Platz					

Note. *Versions used in the study by Wolpert (2000); pitch of the vocal track was kept constant. For all other modifications direction of pitch shift for each track is indicated by arrows. Key difference between melody and accompaniment was always a major second (200 cents).

were recorded as multitrack versions in a professional recording studio. Pieces were performed by students of the Hanover University of Music and Drama (see Table 1 for the list of pieces and the Appendix for leadsheets and scores). Transposed versions of the audio files were produced by using the pitch shift function of the software Adobe Audition (V 1.5). To avoid changes of timbre when transposing only the accompaniment by a major second, we divided the total amount of pitch difference between the accompaniment and solo voice tracks. Thus, the mistuned version ("200 c") the vocal track was pitched down 100 cents and the accompaniment transposed up 100 cents, resulting in a key difference of a major second (200 cents). Due to the limited pitch shift range, the resulting audio files did not show noticeable artefacts. In Wolpert's study no hypothesis was indicated for differences in the perception of clash of keys due to direction of transposition (accompaniment transposed up or down). Thus, we decided to focus on the absolute pitch difference of a major second between melody and accompaniment, but to keep the total number of sound examples constant; additional mistuned versions of the rock 'n' roll, pop, and classical examples were produced (version "100 c") using a key difference of a minor second. However, to make analysis comparable to due to Wolpert's original study, the 100 c data were not analyzed. Finally, recording levels of all tracks were normalized.[1]

[1]Sound examples and supplementary online materials can be obtained from the website http://musicweb.hmt-hannover.de/ndl

EVALUATION SHEET

The evaluation sheet asked participants to indicate musical expertise (years of instrumental training), absolute pitch, preferred musical style, and familiarity with the particular stimulus.

Procedure

All sound examples were presented in a different random order for each experimental group. Additionally, as a distracter task for short-term pitch memory, an excerpt from a Mahler song (see Appendix, Ex. 5) was presented for evaluation after each second test example. A one minute break was inserted after each sound example, during which participants wrote their comments. All participants listened to all stimuli in the respective listening condition (group). To test the nondirected listening (NDL) condition, we gave the following instructions to the participants: "You will listen to four pieces of music presented in three different interpretations each. These pieces will be presented in random order. Please write what differences, if any, you hear among the excerpts." For the directed listening (DL) condition, the following instruction was given: "In the performance of music the same piece can be played in different ways, which is called 'interpretation.' For example, interpretations can differ in the balance between melody and accompaniment. You will listen to four pieces of music presented in three different interpretations each. These pieces will be presented in random order. Please describe in a few phrases your impression of the fit between melody and accompaniment."

Participants were not directed to the bitonal clash of keys. The entire procedure lasted about 30 minutes.

Results

Scoring of Answers

Verbal descriptions of sound examples were coded by means of content analysis into two categories: (a) a response was considered a "match" (correct key error identification) if the participant described any difference that included key words such as "out of tune," "lower," "higher," "sounds strange," "in tune," and so forth; (b) a response was coded a "mismatch" if a participant mentioned that there was no difference or described only differences that did not exist. The proportion of incorrect comments (mismatches) thus equaled 100 minus the percentage of matches. In Wolpert's (2000) study a response was considered a "possible match" if the participant mentioned that the music was lower or higher but also emphasized differences that did not exist. However, in contrast to Wolpert's (2000) study, only a very low number of responses in the category of possible matches were observed, which was not sufficient for statistical analysis. Additionally, from our point of view, this category is ambiguous in coding: frequencies of matches and possible matches were pooled to matches. In our study, in the group of expert listeners there were 15 participants with self-reported absolute pitch. Due to missing statistical differences in the correct identification of clash of keys between relative and absolute pitch possessors, $\chi^2(1, N = 403) = 0.72$, *n.s.*, absolute pitch possessors were not excluded from data analysis. Answers were coded by the investigators ($n = 2$), and the interrater reliability was $r(38) = .90, p < .01$.

Differences Between Expert and Nonexpert Listeners

As Table 2 also shows, the overall difference in percentage of matches between expert and nonexpert listeners was about 41.0% with 58.3% for experts, and 17.5% for nonexperts, $\chi^2(1, N = 991) = 21.0, p < .01$.

Nondirected and Directed Listening

The proportion of participants who noticed the clash of keys in all modified versions is shown in Table 2. Compared to the nondirected listening (NDL) condition, in the directed listening (DL) condition there was a difference of about 38.0% in the perception of clash of keys, with a percentage of 24.3% for NDL and 62.4% for DL, $\chi^2(1, N = 991) = 821.52, p < .01$.

Differences Between Expert and Nonexpert Listeners in Nondirected and Directed Listening

Table 2 shows clearly that attention had a strong influence in both groups: nonexpert listeners increased the proportion of matches from 9.3% to 46.9% when listening with guided attention, $\chi^2(1, N = 588) = 205.47$, $p < .01$). The same strong effect of directed listening also was found in expert listeners: proportion of answers related to the clash of keys increased from 49.3% in the NDL condition of to 78.0% in the DL condition, $\chi^2(1, N = 403) = 62.59, p < .01$. This means that the nonexperts nearly reached the same proportion of matches in directed listening as did the group of experts in nondirected listening, $\chi^2(1, N = 404) = 0.15$, *n.s.* Due to the potential confound between age and expertise, we tested whether there was an influence of age on the identification of key errors. No significant relationship between age and the number of matches was found for older and younger nonexperts, $\chi^2(1, N = 588) = 1.13$, *n.s.*, and for older and younger experts, $\chi^2(1, N = 403) = 0.96$, *n.s.*

The Influence of Musical Style on the Perception of Clash of Keys

A comparison of percentage of matches between expert and nonexpert listeners in each musical style (see Table 2) revealed a consistent pattern: for both experts and nonexperts, the percentage of matches decreased as they moved from the classical to jazz compositions. However, percentage of matches in nonexpert listeners (both NDL and DL) started from 32.0% for the classical composition, then decreased to 15.6% for jazz, and always showed a much smaller percentage of matches than in the group of expert listeners. This distance between experts and nonexperts remained roughly constant across styles: classical—$\chi^2(1, N = 247) = 17.80$, $p < .01$; rock 'n' roll—$\chi^2(1, N = 248) = 32.30$, $p < .01$; pop—$\chi^2(1, N = 248) = 37.50$, $p < .01$; jazz—$\chi^2(1, N = 248) = 7.20$, $p < .05$.

The Influence of Listening Expertise, Attention, and Musical Style on the Perception of Clash of Keys

Previous findings were differentiated by a comparison between all manipulated versions used in the four styles for expert and nonexpert listeners and for both listening conditions. For the group of experts (see Table 2) listening in a nondirected way, the overall tendency of a decreasing percentage of matches from classical to jazz was confirmed. However, this sequence of styles was

TABLE 2. Percentage of Responses to Clash of Keys for all Listening Conditions, Groups of Listening Expertise, and Musical Styles.

		Experts ($n = 101$)	Nonexperts ($n = 147$)	DL ($n = 64$)	NDL ($n = 184$)	Experts ($n = 101$)		Nonexperts ($n = 147$)		Total ($n = 248$)
						DL ($n = 32$)	NDL ($n = 69$)	DL ($n = 32$)	NDL ($n = 115$)	
Jazz*	Unfamiliar	45.2% (19/42)**	32.2% (19/59)	60.0% (33/55)	10.9% (5/46)	51.9% (14/27)	33.3% (5/15)	67.9% (19/28)	0% (0/31)	37.6% (38/101)
	Familiar	27.6% (16/58)	4.7% (4/85)	37.5% (3/8)	12.6% (17/135)	40.0% (2/5)	26.4% (14/53)	33.3% (1/3)	3.7% (3/82)	14.0% (20/143)
	Total	34.7% (35/101)	15.6% (23/147)	56.2% (36/64)	12.0% (22/184)	50.0% (16/32)	27.5% (19/69)	62.5% (20/32)	2.6% (3/115)	23.4% (58/248)
Pop**	Unfamiliar	75.6% (34/45)	10.2% (6/59)	50.0% (25/50)	27.8% (15/54)	90.9% (20/225)	60.9% (14/32)	17.9% (5/28)	3.2% (1/31)	38.5% (40/104)
	Familiar	41.1% (23/56)	5.7% (5/88)	64.3% (9/14)	14.6% (19/130)	80.0% (8/10)	32.6% (15/46)	25.0% (1/4)	4.8% (4/84)	19.4% (28/144)
	Total	56.4% (57/101)	7.5% (11/147)	53.1% (34/64)	18.5% (34/184)	87.5% (28/32)	42.0% (29/69)	18.8% (6/32)	4.3% (5/115)	27.4% (68/248)
Rock 'n' Roll**	Unfamiliar	63.0% 29/46	10.2% (10/98)	50.0% (5/10)	25.4% (34/134)	83.3% (5/6)	60.0% (24/40)	0% (0/4)	10.6% (10/94)	27.1% (39/144)
	Familiar	71.2% (37/52)	23.4% (11/47)	63.5% (33/52)	31.9% (15/47)	88.5% (23/26)	53.8% (14/26)	38.5% (10/26)	4.8% (1/21)	48.5% (48/99)
	Total	66.3% (67/101)	15.0% (22/147)	60.9% (39/64)	22.7% (50/184)	87.5% (28/32)	56.5% (39/69)	34.4% (11/32)	9.6% (11/115)	35.9% (89/248)
Classical**	Unfamiliar	71.1% (27/38)	43.3% (29/67)	80.0% (40/50)	29.1% (16/55)	81.0% (17/21)	58.8% (10/17)	79.3% (23/29)	15.8% (6/38)	53.3% (56/105)
	Familiar	75.0% (36/48)	22.8% (18/79)	0% (0/2)	43.2% (54/125)	—	75.0% (36/48)	0% (0/2)	23.4% (18/77)	42.5% (54/127)
	Total	76.0% (76/100)	32.0% (47/147)	79.4% (50/63)	39.7% (73/184)	87.1% (27/31)	71.0% (49/69)	71.9% (23/32)	20.9% (24/115)	49.8% (123/247)
All pieces***	Unfamiliar	63.7% (109/171)	22.6% (64/283)	62.4% (103/165)	24.2% (70/289)	73.7% (56/76)	55.8% (53/95)	52.8% (47/89)	8.8% (17/194)	38.1% (173/454)
	Familiar	52.3% (112/214)	12.7% (38/299)	59.2% (45/76)	24.0% (105/437)	80.5% (33/41)	45.7% (79/173)	43.3% (12/35)	9.8% (26/264)	29.2% (150/513)
	Total	58.3% (235/403)	17.5% (103/588)	62.4% (159/255)	24.3% (179/736)	78.0% (99/127)	49.3% (136/276)	46.9% (60/128)	9.3% (43/460)	34.1% (338/991)

Note. *% of matches of all coded responses for manipulated jazz version with pitch difference between accompaniment and vocal track of 200 c; **Number of responses coded as matches/total number of coded responses (matches + mismatches); ***One manipulated piece per style (total = 4 pieces)

only based on the direction of decrease of matches and does not suggest a scale-like variable. The control condition of directed listening (see Table 2) showed clearly that expert listeners did not exceed 87.5% of matches. Only the jazz stimulus showed a much lower proportion of matches of only 50.0% in expert listeners in the DL condition. In the group of nonexpert listeners we could observe the same overall tendency of a decrease of percentage of matches from classical to jazz. However, in all versions the percentage of matches in the NDL condition never exceeded 20.9% of answers. On the other hand, nonexperts strongly benefited from directed listening instructions and increased match percentages for all musical examples.

As Table 3 shows, except for the decrease of matches between classical and rock 'n' roll and from pop to jazz in the group of expert listeners, all differences became statistically significant. In the group of nonexpert listeners, the only nonsignificant match decrease was found between the styles of pop and jazz and those of rock 'n' roll and pop. Surprisingly, the group of expert listeners never reached 100% in the perception of clash

TABLE 3. Differences in the Perception of Clash of Keys for Four Musical Styles (Chi-square) in the Nondirected Listening (NDL) Condition for Expert and Nonexpert Listeners.

	Expert listeners			
	Classical	Rock 'n' Roll	Pop	Jazz
Classical				
Rock 'n' Roll	1.13			
Pop	5.13*	7.35**		
Jazz	13.23**	6.89*	2.08	

Note. $^*p < .05$; $^{**}p < .01$ (two-tailed)

	Nonexpert listeners			
	Classical	Rock 'n' Roll	Pop	Jazz
Classical				
Rock 'n' Roll	4.82*			
Pop	12.45**	2.25		
Jazz	16.33**	4.57*	0.05	

Note. $^*p < .05$; $^{**}p < .01$ (two-tailed)

of keys. In other words, a significant proportion (29%) of the expert listeners in nondirected listening to a mistuned piece of classical music did not notice the clash of keys. However, the questions remains open as to whether the expert listeners' high rate of key error identification in the classical style resulted from their music training and familiarity with the respective piece.

The Influence of Familiarity

Unexpectedly, familiarity with a musical piece had an overall negative effect on the perception of clash of keys: matches for familiar pieces were 29.2%, and for unfamiliar pieces, 38.1%. This effect of a smaller proportion of matches for familiar pieces could be observed in experts (52.3% of matches for familiar and 63.7% for unfamiliar pieces: $\chi^2(1, N = 385) = 5.06$, $p < .03$) as well as in nonexperts (12.7% of matches for familiar and 22.6% for unfamiliar pieces; $\chi^2(1, N = 582) = 9.87$, $p < .01$). Although familiarity with the particular piece did not show a significant overall effect for all pieces, $\chi^2(1, N = 967) = 1.17$, *n.s.*, Table 2 shows significant effects for some pieces: for jazz—$\chi^2(1, N = 244) = 10.60$, $p < .01$, pop—$\chi^2(1, N = 248) = 6.20$, $p < .05$, and rock 'n' roll—$\chi^2(1, N = 243) = 6.00$, $p < .05$. However, in two out of three cases (jazz and pop) familiarity with the piece resulted in a lower percentage or matches.

TABLE 4. Comparison of Percentage of "Matches" in Wolpert's (2000) and Our Studies.

	Wolpert (2000) Matches (%)	Kopiez & Platz Matches (%) 'Jazz' only
Experts	100	27.5
Nonexperts	40.0	2.6

Note. For better comparability, frequencies in the categories of "matches" and "possible matches" from Wolpert's (2000) study have been pooled and recalculated to the category of "matches."

Comparison with the Results from Wolpert's Study

The comparison of our results with those in Wolpert's (2000) study (see Table 4) showed that in our study neither experts nor nonexperts reached the percentages of matches as found by Wolpert. If we compare our results for the jazz piece to those reported by Wolpert, we can see a large difference of about 77.0% for expert listeners and of about 36.0% for nonexperts. Differences between both studies also remain important for the comparison between Wolpert's and our percentage of matches for all stimuli used. Averaged across all manipulated pieces, the experts' percentage of matches for the NDL condition never exceeded about 49.3% in our study.

Discussion

Three of our four research questions were confirmed: expert listeners showed a higher percentage of matches in the perception of clash of keys than nonexperts, the percentage of matches increased from nondirected to directed listening, and the percentage of matches was dependent on musical style. However, familiarity with a particular piece decreased the perception of clash of keys. The first finding is in accordance with Wolpert's (2000) study. However, we could not confirm her finding of 100% of perception of clash of keys in expert listeners—neither for her jazz example nor for the examples from other musical styles. We would like to propose two explanations for her findings: First, the "thin" sound of the +200 c version and the "booming" sound of the −200 c version could have given acoustical cues to manipulation of stimuli for the expert listeners in her experiment. Second, the low number of participants in the expert group ($n = 10$) could have resulted in a highly selected and biased sample.

In the following sections we will discuss three explanations for the findings of our study: first, the cultural explanation that emphasizes the role of training, stimulus familiarity, and attention for perceptual skills; second, the psychoacoustic explanation, which concentrates on the role of roughness for the perception of clash of keys; third, the role of musical style, which tries to find style-specific reasons for the perception of key conflicts.

The Role of Listening Expertise, Familiarity, and Attention

The general benefit of listening expertise in the perception of clash of keys was confirmed in our study. This seems to be a general effect of training that is independent from the modality of perception. For example, in a study on visual "change blindness," Werner and Thies (2000) showed that expertise in a domain (e.g., American football) can increase observers' sensitivity to changes in domain-related images (e.g., football scene). Concerning sensitivity to out-of-tune notes, there is clear evidence for superior performance of musicians over nonmusicians. For example, Elliot, Platt, and Racine (1987) found that nonmusicians performed poorly in tasks in which intonation was evaluated in comparison with an internalized standard (e.g., interval adjustment tasks), but they showed that the same achievement in out-of-tune note discrimination tasks did not require such a comparison. Münte, Kohlmetz, Nager, and Altenmüller (2001) confirmed superior sensitivity to auditory localization in conductors, and Kopiez (2003) observed highly trained adaptability to equal tempered intonation in the performance of professional trumpet players.

The Role of Familiarity with the Particular Piece

A surprising finding was the negative effect of familiarity on the perception of clash of keys. This was the case for the styles of jazz, pop, and rock 'n' roll, but not for classical. At first sight this effect is counterintuitive; results from reading research did in fact predict a positive effect of familiarity with a text passage on the improvement in writing error detection (Pilotti, Chodorow, & Thornton, 2004, 2005). Familiarity with musical pieces seems to make one less sensitive to mistunings. However, currently we only find anecdotal support for this hypothesis: during a concert by the rock band Van Halen, the pitch of the playback of the last piece, the famous hit "Jump," differed from the band's pitch by about 146 cents due to a technical fault.[2] The audience did not seem to be disturbed by this clash of keys. In other words, listening to a familiar piece of music seems to be largely guided by expectancy and less conducive to close attention than is listening to a novel piece.

Psychoacoustics of Sensory Consonance (Roughness)

The question of whether the perception of clash of keys could be driven by psychoacoustic features such as sensory roughness remains open. Auditory roughness is important for our study because it could correlate with clash of key perception in terms of an implicit cue that self-report does not access. Support for the importance of sensory processes in the perception of consonance and harmony can be found in psychoacoustic theories of consonance. The main group of theories focuses on the work by Helmholtz. In his theory of sensory consonance (Konsonanz) and harmony (Klangverwandtschaft), he described the sensory character of dissonant dyads as follows: "It is apparent to the simplest natural observation that the essence of dissonance consists merely in very rapid beats . . . The cause of the unpleasantness of dissonance we attribute to this roughness and entanglement. The meaning of this distinction may be thus briefly stated: Consonance is a continuous, dissonance an intermittent sensation of tone." (Helmholtz, 1863/1954, p. 226). The main reason for perceived roughness is the number of beats caused by interference between adjacent partials. This means that the frequency of fluctuation degree influences auditory roughness. For example, as Helmholtz's roughness curves for violin dyads over two octaves showed, roughness is highest for the minor and major seconds (with a maximum at a frequency difference of about 33 Hz for all registers) and lowest for the fifth and octave. Thus, the fifth and octave are generally perceived as highly consonant.

Auditory roughness also is part of the concept of sensory pleasantness (Aures, 1985a; Zwicker & Fastl, 1999). The more pleasant a sound is evaluated, the lower its sharpness, roughness, and loudness, and the higher its "tonality" rating. For the description of the degree of sensory consonance on our stimuli, auditory roughness was calculated according to the procedure of Aures (1985b). Following his model, the average specific roughness, measured in centi asper (cA) per Bark, was calculated. In this procedure, envelope fluctuations of the excitation in 24 critical bands are calculated. The partial roughness per critical band is evaluated, and partial roughnesses are summarized after being balanced by the correlation between envelope fluctuations of adjacent critical bands. This was done by means of the software dBSONIC (2004). A sensory roughness of 1 asper is defined as that

[2]The Van Halen concert took place in Greensboro, NC, on September 29, 2007 and is documented in the internet on http://www.youtube.com/watch?v=Mjx_GjyXCs4 (retrieved August 12, 2008).

TABLE 5. Stimuli Roughness Calculation Following Aures's (1985b) Model of Sensory Roughness.

Example	Version	Sensory roughness [centi Asper] Mean	*SD*
You Make Me Feel So Young [Jazz]*	Original	12.28 (11.75)	6.20 (6.44)
	200 c ↑	10.63 (11.73)	6.21 (4.52)
	200 c ↓	10.54 (12.03)	5.87 (4.72)
You've Got a Friend [Pop]	Original	18.01	8.18
	200 c	19.39	7.90
Walking on Sunshine [Rock 'n' Roll]	Original	18.75	9.23
	200 c	20.54	10.86
Haydn: *Trumpet Concerto* [Classical]	Original	7.34	4.07
	200 c	6.90	4.14

Note. *Due to significant signal noise and a low frequency bias, the original sound files from Wolpert's (2000) study had to be restored. Roughness values for the unrestored sound files are indicated in brackets. Time series analyses of sensory roughness of all sound files can be obtained from the supporting online material (http://musicweb.hmt-hannover.de/ndl).

of a 100% amplitude modulated 60 dB, 1 kHz tone modulated at a frequency of 70 Hz (Zwicker & Fastl, 1999). Recording levels of soundfiles were normalized before analysis, and roughness was calculated for time windows of 100 ms (roughness sequence length: 500 ms). Table 5 shows the roughness values for the stimuli used.

Mean values of roughness for the out of key and original versions show only small differences between the two versions of each piece. This can be explained by the concept of sensory roughness: roughness is a feature of more time-invariant musical events, such as a consonant or dissonant final chord. However, in the case of time-varying events such as music, the transposition of a voice not only makes consonant intervals dissonant, but dissonant intervals also consonant. This compensation effect could be the reason for the small mean differences between in key and out of key versions.

Thus, we assume: (a) that models of sensory roughness are valid for the explanation of consonance in isolated dyads, but are only of limited value if applied to complex chords or time-varying signals such as music; and (b) that both "roughness and tonal fusion determine the sensation of consonance and dissonance" (Ebeling, 2008). In other words, just as vision is not "inverse optics" (Mausfeld, 2002), we can say that hearing is not simply "inverse acoustics" or "inverse music theory." Thus, we will try to give alternative explanations as to why the surprisingly low perception rate of clash of keys in expert and nonexpert listeners results not from a perceptual deficit, but from the perceptual system's adequate mechanisms for the particular music stimulus used. There is evidence from other studies on auditory perception that the evaluation of harmonic intonation is characterized by a relatively large tolerance towards pitch deviations. Fricke (1988) showed, for example, that music performance always is characterized by a certain pitch width (Klangbreite), such as that produced (deliberately or randomly) in instrumental or vocal ensemble intonation and in string vibrato. These pitch deviations make a sound natural and pleasant. A pitch variation of up to 80 cents between ensemble players does not seem to disturb the overall pleasantness of listening to music. This explanation concurs with the theory of categorical perception, which states that a change in some variable along a continuum is perceived not as gradual, but rather as instances of discrete categories. From this point of view, listening is more a kind of "harmonizing in hearing" (Zurechthören) than "inverse acoustics." This assumption goes along with the studies by Enders (1981) on the perception of chord intonation, in which listening experiments demonstrated a relationship between chord structure and intonation discrimination. Distortions of the pitch of individual tones within a consonant chord (e.g., major/minor chord) were easier to identify than were altered pitches within dissonant chords (e.g., dominant seventh chord). This effect can result in an acceptable chord intonation width of up to ±60 cents. In a very recent published neurophysiological theory of consonance, Ebeling (2008) demonstrated that deviations from exact intonation in dyads are congruent with the subjective impression of consonance. In his mathematical formulation of a periodicity-based theory of consonance, the author showed that intonational uncertainty is consistent with the overall impression of "tonal fusion" (Verschmelzung) as the main feature of tonal consonance. The calculation of neural consonance by the underlying general coincidence function is consistent with deviations in signal peridocity (pitch). In other words, key divergence doesn't alter the sense of chord progressions: it just tends to add confusing acoustic baggage on the whole.

The Role of Musical Style

Differences in the perception of clash of keys between pieces from different musical styles are an important finding of this study. We will try to give an explanation for these differences. The perception rate of clash of keys was highest in the classical and lowest in the jazz piece. The classical piece (*Trumpet Concerto* by Haydn) has a simple harmonic structure, and its tonal coherence is a basic feature. The harmonic style of this music corresponds to eighteenth century European music. Listeners of this music expect that the relationships between melody and accompaniment will be perfectly consistent from the tonal point of view. Nonmusicians also may have had innumerable occasions to listen to eighteenth century music without being concert-goers. In other styles this kind of consistency is not necessary: some aspects of tonal relationships in jazz music were similar to examples of avant-garde music of those years, and this could make it more difficult to notice the tonal transpositions in the accompaniment. In other words, the perception of clash of keys is assumed to be easier in music in which key clashes normally are not present. The harmonic style of pop music is closer to classical than to jazz music, and the sound of electronic instruments in rock 'n' roll music also implies some inharmonic features (e.g., distorted guitar sounds). However, we should bear in mind that only one piece per style was tested, and the questions of (a) generalizability of findings to other pieces of the same style and (b) representativeness of the selected pieces continue to be of general interest for empirical music research (see Kopiez & Lehmann, 2008).

General Discussion

In our study we found that the perception of clash of keys is dependent on listening expertise, attentional set, and musical style. In both groups of expert and nonexpert listeners there was a high proportion of participants who did not notice the crucial point of key conflict. Even when directed to listen to the relationship between melody and accompaniment, expert listeners achieved only 78.0% correct, averaged across all four musical styles. Without guided attention, only 49.3% of the expert listener group noticed the clash of keys. This means that the expert listeners' behavior was not a result of low-level perceptual processing, but of acquired musical and cultural knowledge. On the other hand, results indicate that nonexpert listeners have a high potential to focus on musical features and to improve listening achievement when they receive instructions. In the case of directed listening, this group increased the percentage of matches from 9.3% to 46.9%.

If the overall impression of tonal coherence is not the result of mere sensory information (such as sensory roughness), we have to find alternative explanations for the participants' behavior. One explanation could be derived from experiments related to the perception of tonal closure in cadences. As Eberlein and Fricke (1992) showed, the impression of tonal closure of a cadence is not the result of relationships between the roots of chords (which can differ from the bass notes), but of voice-leading patterns, such as a leading note in the upper voice or a falling fourth and fifth in the bass. This relative importance of interval size and direction also was observed in studies of tonality perception in melodic sequences by Auhagen (1994). He found that the final tone of a short melodic sequence is rated as its tonal centre if the tone is reached by a rising fourth, falling fifth, or rising minor second (leading note). Against this background, we assume that participants listening to music in conflicting keys focus only on salient structural features of the music, such as melodic interval patterns or cadences, to generate the impression of tonal coherence. We also have to consider that dissonance could influence other categories of music perception that are more relevant for music listening than the detection of key conflicts. For example, as Bigand, Parncutt, and Lerdahl (1996) showed, sensory roughness of dissonant chords contributes to ratings of chord tension in expert listeners. We conclude that, due to the widespread diffuse listening common in everyday life (Behne, 1990), most instances of clash of keys (e.g., out–of tune singing in popular TV shows) may remain unperceived by the majority of the audience.

Author Note

We would like to thank Rita S. Wolpert for her musical examples, Sebastian Seuring for data collection and recording of stimuli, and Wolfgang Auhagen, Martin Ebeling, Marco Lehmann, Emery Schubert, and two anonymous reviewers for helpful comments on an earlier version of this paper.

Correspondence concerning this article should be addressed to Reinhard Kopiez, Hanover University of Music and Drama, Emmichplatz 1, 30175 Hanover, Germany. E-MAIL: kopiez@hmt-hannover.de

References

AUHAGEN, W. (1994). Experimentelle Untersuchungen zur auditiven Tonalitätsbestimmung in Melodien [Experimental studies on the determination of tonality in melodies]. *Kölner Beiträge zur Musikforschung, Vol. 180*. Kassel: Bosse.

AURES, W. (1985a). Der sensorische Wohlklang als Funktion psychoakustischer Empfindungsgrößen [Sensory pleasantness as a function of psychoacoustic sensation]. *Acustica, 58*, 282-290.

Aures, W. (1985b). Ein Berechnungsverfahren der Rauhigkeit [A procedure for calculating roughness]. *Acustica, 58*, 268-281.

Behne, K.-E. (1990). *Hörertypologien. Zur Psychologie des jugendlichen Musikgeschmacks* [Typologies of listeners. On the psychology of musical taste in youths]. Regensburg: Bosse.

Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception and Psychophysics, 58*, 125-141.

Billroth, T. (1894). *Wer ist musikalisch?* [Who is musical?]. Berlin: Gebrüder Paetel.

Dahlhaus, C. (1990). *Studies on the origin of harmonic tonality* (R. O. Gjerdingen, Trans.). Princeton, NJ: Princeton University Press.

Dahlhaus, C. (2001). Harmony. In S. Sadie (Ed.), *The new Grove dictionary of music and musicians* (2nd ed., Vol. 10, pp. 858-877). London: Macmillan.

dBSONIC (2004). dBSONIC (Version 4.4) [Computer software]. Limonest, France: 01dB-Metravib.

Ebeling, M. (2008). Neural periodicity detection as a basis for the perception of consonance: A mathematical model of tonal fusion. *Journal of the Acoustical Society of America, 124*, 2320-2329.

Eberlein, R., & Fricke, J. (1992). *Kadenzwahrnehmung und Kadenzgeschichte: Ein Beitrag zu einer Grammatik der Musik* [Cadence perception and the history of the cadence: A contribution to a grammar of music]. Frankfurt, Germany: Peter Lang.

Elliot, J., Platt, J. R., & Racine, R. J. (1987). Adjustment of successive and simultaneous intervals by musically experienced and inexperienced subjects. *Perception and Psychophysics, 42*, 594-598.

Enders, B. (1981). *Studien zur Durchhörbarkeit und Intonationsbeurteilung von Akkorden* [Studies on the aural comprehension of chords and the evaluation of chordal intonation]. Regensburg: Gustav Bosse.

Fricke, J. (1988). Klangbreite und Tonempfindung. Bedingungen kategorialer Wahrnehmung aufgrund experimenteller Untersuchung der Intonation [Sound width and tone sensation. Conditions for categorical perception, based on the experimental exploration of intonation]. *Musikpsychologie, 5*, 67-87.

Helmholtz, H. (1954). *On the sensations of tone as a physiological basis for the theory of music* (A. J. Ellis, Trans.). New York: Dover Publications. (Original work published 1863)

Hyer, B. (2001). Tonality. In S. Sadie (Ed.), *The new Grove dictionary of music and musicians* (2nd ed., Vol. 25, pp. 583-594). London: Macmillan.

Jäncke, L., Buchanan, T. W., Lutz, K., & Shah, N. J. (2001). Focused and nonfocused attention in verbal and emotional dichotic listening: An FMRI study. *Brain and Language, 78*, 349-363.

Kopiez, R. (2003). Intonation of harmonic intervals: Adaptability of expert musicians to equal temperament and just intonation. *Music Perception, 20*, 383-410.

Kopiez, R., & Lehmann, M. (2008). The 'open-earedness' hypothesis and the development of age-related aesthetic reactions to music in elementary school children. *British Journal of Music Education, 25*, 121-138.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Leibniz, G. W. (1991). *The monadology. An edition for students* (N. Rascher, Trans.). Pittsburgh, PA: University of Pittburgh Press. (Original work published 1714)

Mausfeld, R. (2002). The physicalistic trap in perception theory. In D. Heyer & R. Mausfeld (Eds.), *Perception and the physical world* (pp. 75-112). Chichester: Wiley.

Münte, T. F., Kohlmetz, C., Nager, W., & Altenmüller, E. (2001). Superior auditory spatial tuning in conductors. *Nature, 409*, 580.

Pascall, R. (2001). Style. In S. Sadie (Ed.), *The new Grove dictionary of music and musicians* (2nd ed., Vol. 24, pp. 638-642). London: Macmillan.

Pilotti, M., Chodorow, M., & Thornton, K. C. (2004). Error detection in text: Do feedback and familiarity help? *Journal of General Psychology, 131*, 242-267.

Pilotti, M., Chodorow, M., & Thornton, K. C. (2005). Effects of familiarity and type of encoding on proofreading of text. *Reading and Writing: An Interdisciplinary Journal, 18*, 325-341.

Schramm, H. (2005). *Mood Management durch Musik* [Mood management by music]. Köln: Herbert von Halem Verlag.

Sloboda, J. A., O'Neill, S., & Ivaldi, A. (2001). Functions of music in everyday life: An exploratory study using the Experience Sampling Method. *Musicae Scientiae, 5*, 9-32.

Thomson, W. (1999). *Tonality in music*. San Marino, CA: Everett Books.

Vitevich, M. S. (2003). Change deafness: The inability to detect changes between two voices. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 333-342.

Werner, S., & Thies, B. (2000). Is "change blindness" attenuated by domain-specific expertise? An expert-novice's comparison of change detection in football images. *Visual Cognition, 7*, 163-173.

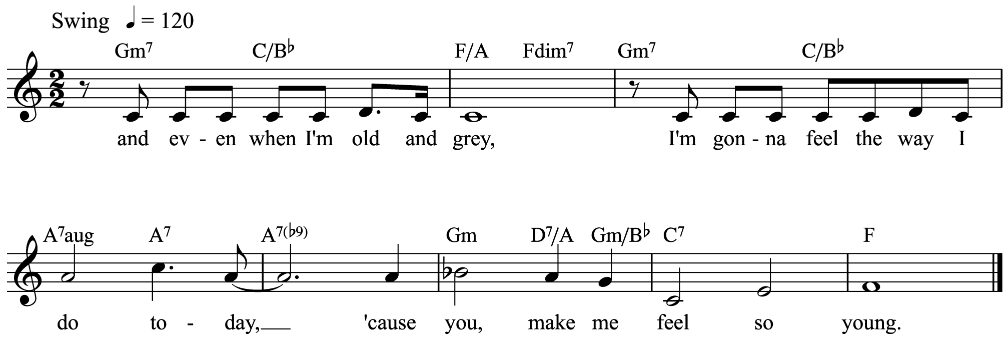
Whittall, A. (2001). Bitonality. In S. Sadie (Ed.), *The new Grove dictionary of music and musicians* (2nd ed., Vol. 3, pp. 637). London: Macmillan.

Wolpert, R. S. (1990). Recognition of melody, harmonic accompaniment, and instrumentation: Musicians vs. nonmusicians. *Music Perception, 8*, 95-106.

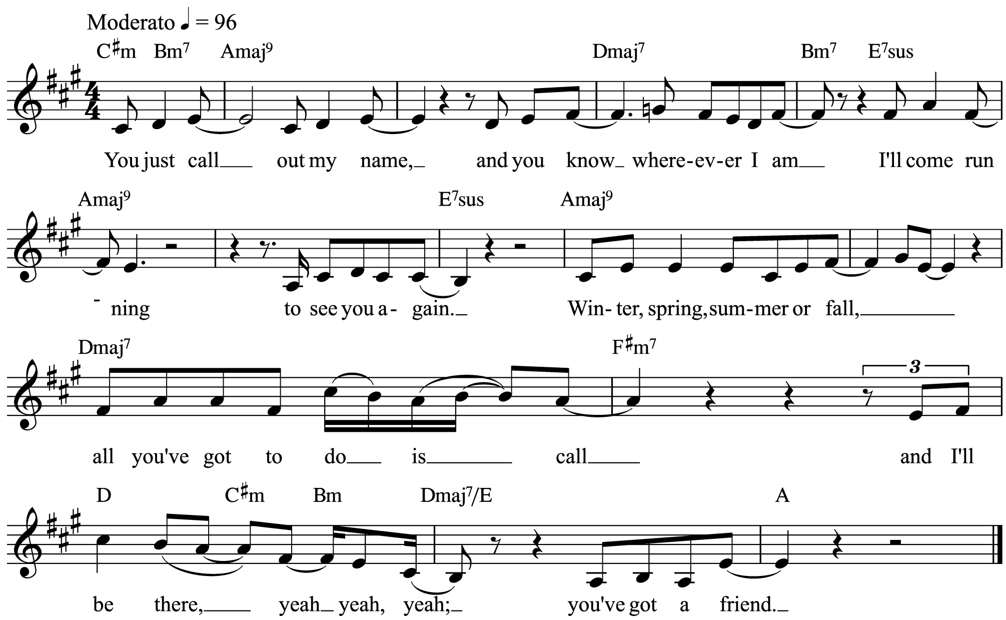
Wolpert, R. S. (2000). Attention to key in a nondirected music listening task. *Music Perception, 18*, 225-230.

Zwicker, E., & Fastl, H. (1999). *Psychoacoustics. Facts and models* (2nd ed.). Berlin: Springer.

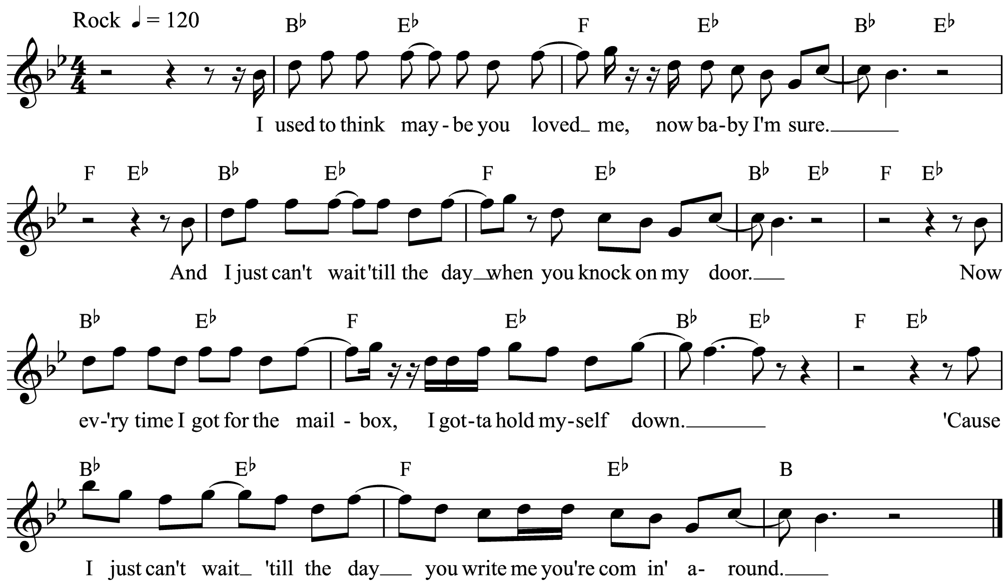
Appendix



EXAMPLE 1. Lead sheet of the test piece used for the style of "Jazz" (*You Make Me Feel So Young*).



EXAMPLE 2. Lead sheet of the test piece used for the style of "Pop" (*You've Got a Friend*).



EXAMPLE 3. Lead sheet of the test piece used for the style of "Rock 'n' Roll" (*Walking on Sunshine*).

EXAMPLE 4. Piano reduction of the test piece for the style of "Classical" (Haydn: *Trumpet Concerto,* 2nd movement).

EXAMPLE 5. Piano reduction of the distractor piece (Mahler: No. 4 from *5 Lieder nach Rückert*).