CHORUS-PLAYLIST: EXPLORING THE IMPACT OF LISTENING TO ONLY CHORUSES IN A PLAYLIST

Kosetsu Tsukuda **Masahiro Hamasaki** **Masataka Goto**
National Institute of Advanced Industrial Science and Technology (AIST), Japan
{k.tsukuda, masahiro.hamasaki, m.goto}@aist.go.jp

ABSTRACT

When people listen to playlists on a music streaming service, they typically listen to each song from start to end in order. However, what if it were possible to use a function to listen to only the choruses of each song in a playlist one after another? In this paper, we call this music listening concept "chorus-playlist," and we investigate its potential impact from various perspectives such as the demand and the objectives for listening to music with chorus-playlist. To this end, we conducted a questionnaire-based online user survey involving 214 participants. Our analysis results suggest reusable insights, including the following: (1) We show a high demand for listening to existing playlists with the chorus-playlist approach. We also reveal preferred options for chorus playback, such as adding crossfade transitions between choruses. (2) People listen to playlists with chorus-playlist for various objectives. For example, when they listen to their own self-made playlists, they want to boost a mood or listen to music in a specific context such as work or driving. (3) There is also a high demand for playlist creation on the premise of continuous listening to only the choruses of the songs in a playlist. The diversities of artists, genres, and moods are more important when creating such a playlist than when creating a usual playlist.

1. INTRODUCTION

The chorus of a song is one of the most distinctive parts in the song. In terms of acoustic aspects, it has been reported that the chorus tends to have louder sound, contain heavier instrumentation and additional vocals, and include the highest-pitch vocal note in a song [1–3]. In terms of cognitive aspects, the chorus tends to be the catchiest, most memorable, and most salient part of a song for emotional expression [4–7]. Moreover, the chorus is often characterized by the property of being a song's most repeated section [8, 9]. Because of these characteristics, the chorus has attracted academic attention. For example, research has been conducted on music structure analysis including chorus detection [8–35] and its use for music summary generation [36, 37]. In addition, music datasets specializing in choruses have been made publicly available [38].

Attribution: K. Tsukuda, M. Hamasaki, and M. Goto, "Chorus-Playlist: Exploring the Impact of Listening to Only Choruses in a Playlist", in *Proc. of the 24th Int. Society for Music Information Retrieval Conf.*, Milan, Italy, 2023.

As for general music listening habits beyond choruses, the amount of time spent listening to music through playlists on music streaming services has increased [39–42]. It has also been reported that this listening time is longer than the time of listening to music via albums [43]. On the services, both playlists that are created by general users and those that are created by professional curators or automatically generated are widely available [42, 44–46]. As a result, Spotify has over 4 billion playlists, for example [42]. With the popularity of playlists, many studies have been conducted on playlist recommendation, generation, and analysis [40, 42, 44, 47–76].

Given the importance of choruses and playlists, we focus on a listening approach in which only the choruses of the songs in a playlist are played one after another. Certain smartphone music player applications such as Vocacolle App by DWANGO Co., Ltd., MIXTRAX App by Pioneer Corporation, and KENWOOD Music Control have provided a function to continuously play the choruses of songs or parts including the choruses. However, there has been no academic discussion on the impact of this listening approach. In this paper, we refer to the concept of continuous listening to only the choruses of songs in a playlist as "chorus-playlist." This concept can be applied not only when a user listens to playlists that she created but also when she listens to playlists created by other users. Under this concept, a user could create a playlist on the premise of continuous listening to only the choruses of the playlist's songs. Hence, the goal of this paper is to reveal the usefulness of chorus-playlist and provide reusable insights.

To achieve this goal, we conducted a questionnaire-based online user survey involving 214 participants. Our contributions can be summarized as follows.

- To our knowledge, this is the first study investigating the impact of continuous listening to only the choruses of songs in a playlist.
- We reveal user preferences for chorus playback in a playlist (*e.g.*, users prefer to add crossfade transitions between choruses). We also show a high demand and certain user objectives for listening to playlists with chorus-playlist. For example, people often want to listen to their own self-made playlists to boost a mood or in a specific context such as work or driving.
- We show that people tend to be willing to create a playlist for continuously listening to only the choruses of the songs in the playlist. We also reveal important properties in creating such playlists (*e.g.*, the diversity of moods, genres, and artists).

- According to the survey results, we suggest new research topics for the music information retrieval (MIR) community (*e.g.*, song recommendation for users who listen to music with chorus-playlist). We also make several proposals for music streaming platforms to attract users (*e.g.*, when using the chorus-playlist approach, people would be willing to listen to playlists containing hit songs to efficiently check them out).
- We have made a portion of the survey results publicly available on the web to support future studies [1].

2. RELATED WORK

2.1 Chorus Analysis and Detection

The chorus is distinctive compared to other sections of a song in terms of acoustic, structural, and cognitive aspects. In terms of acoustic aspects, it has been known that choruses tend to be louder, to contain heavier instrumentation and more vocals, and to have the highest-pitch vocal note in a song [1–3]. More recently, Balen et al. [77] revealed that choruses have a smaller dynamic range and a greater variety of MFCC-measurable timbres, as compared to other sections. Regarding structural aspects, choruses are usually repeated more than other sections such as intros and verses [8, 9]. As for cognitive aspects, choruses are the catchiest and most memorable sections for listeners [4, 5, 7]. For artists, too, choruses are distinctive in that they are the most salient sections for emotional expression [6]. Given these characteristics, we focus here on choruses as song excerpts, rather than other sections. As we will show in section 7, the chorus is more preferred than other sections for continuous listening in a playlist.

Because of the importance of choruses, many studies have addressed musical structure analysis, including chorus detection, based on music audio signals [8–15, 17–24, 26,27,29–32,34,35] or lyrics [16,25,28,33]. The accuracy of identifying the chorus section is approximately 80-90% in terms of the F-measure [12,22,29,33]. Accordingly, the feasibility of implementing the chorus-playlist concept on music streaming services is sufficiently high. For songs in which the correct chorus cannot be detected, it is also possible to have users on a service manually correct choruses through a collective intelligence approach [78].

2.2 Playlist Analysis and Recommendation

On today's music streaming services, playlists have become a central way to listen to music [40–42]. The majority of playlists on services are created by general users rather than professional music enthusiasts [44]. Users create playlists not only for their own listening but also to share their musical preferences with other users such as friends and followers [44, 53, 70]. It has also been reported that playlists can be characterized by certain properties [40, 62, 68, 69, 74, 76] such as song order [42] and low diversity in terms of both artists and genres [70]. In this paper, too, we report objectives for listening to music with chorus-playlist (section 5) and important properties for creating a playlist to continuously listen to only the choruses of songs in the playlist (section 6).

Although playlists are actively created by users, it is time consuming to manually create a playlist [44]. To ease the process, two approaches have been studied: assisted playlist creation [49–52, 54, 55, 71] and automatic playlist generation [47, 48, 56, 58–61, 63–67, 72, 73, 75]. In song recommendation for a playlist or generation of a playlist, it is typical to consider the song order and/or the audio similarity between songs. Furthermore, there are several studies that automatically extract prominent sections (not only limited to choruses) from individual songs and generate DJ mixes [79] or medleys [80] by connecting them. Music streaming services such as Spotify and Deezer provide functions for automatic playlist generation to promote song discovery by users [45, 46]. In this paper, according to our survey results, we discuss not only new approaches for these research topics but also how to encourage users on music streaming services to more actively interact with playlists and discover novel songs.

3. PARTICIPANTS

We recruited participants for our survey via an online research company in Japan. We limited the participants to those who are Japanese, listened to music an average of at least one day per week via any music streaming service, and had created at least 10 playlists on the service. We paid 51.6 USD (7,000 JPY) to each participant. Although 222 participants answered the questionnaire in sections 4, 5, 6, and 7 through a web browser, to make the analysis results more reliable, we removed the answers from eight participants who submitted improper responses to a free-response question. The remaining 214 participants were diverse in both gender and age range: 89 males (10s: 1; 20s: 29; 30s: 35; 40s: 17; 50s: 7), and 125 females (10s: 1; 20s: 45; 30s: 38; 40s: 25; 50s: 16).

4. PREFERENCE FOR CHORUS PLAYBACK

4.1 Chorus Playback Choices

As explained in section 1, the concept of chorus-playlist enables a user to continuously listen to only the choruses of songs in a playlist. However, some users may prefer to add crossfade transitions between choruses. In this section, we investigate the preferences for chorus playback in chorus-playlist in terms of the following three choices.

- TimePreChorus: the playback time before the chorus. The options are "no playback," "5 seconds," and "10 seconds." "No playback" means that only the choruses are continuously played, without any part of the song before the chorus.
- Crossfade: whether 1-second crossfade transitions are added between songs. The options are "on" and "off."
- TimeChorus: the playback time for the chorus. The options are "15 seconds," "30 seconds," and "adaptive." In the case of "15 (resp. 30) seconds," 15 (resp. 30) seconds on a song is played from the beginning of the

[1] The data can be downloaded from https://github.com/ktsukuda/chorus_playlist.

first chorus [2]. In the case of "adaptive," the first chorus is played from beginning to end regardless of its length.

Hereafter, we refer to a combination of these three choices in the form of a triplet such as (TimePreChorus, Crossfade, TimeChorus) = (5 seconds, off, adaptive).

4.2 Dataset

In this survey, the participants listened to playlists that we provided. To reduce bias due to the played songs, we used 120 songs created by professional musicians that we commissioned. That is, we guaranteed that the participants had never listened to any of the 120 songs. Instead of using chorus detection methods introduced in section 2.1, a music expert manually labeled the start and end times of each song's first chorus to prevent detection errors. For 26 songs that started with the chorus, the expert labeled the start and end times of the second chorus, because the mood and/or beat of such leading choruses are sometimes different from those of the second and subsequent choruses. The average and standard deviation of chorus lengths for the 120 songs were 29.0 and 7.66 seconds, respectively, and 52 songs had choruses longer than 30 seconds.

As the participants had various music preferences, we created diverse playlists by sampling the songs to be included in playlists as follows. First, the 120 songs were plotted in the valence-arousal (VA) space according to their audio features. Next, we applied the k-means algorithm and classified the 120 songs in the VA space into three clusters. We created three playlists by sampling five songs at random for each playlist from one cluster. Similarly, we created three more playlists from another cluster. Finally, we created three playlists containing diverse songs in terms of their moods by randomly selecting one song from each of the two previous clusters and three songs from the remaining cluster. Each playlist's song order was also determined randomly. In total, we created nine playlists that each comprised five songs. Note that there were no song overlaps between any pairs of playlists.

For each playlist, there were 18 total option combinations (3 options for TimePreChorus $\times$ 2 options for Crossfade $\times$ 3 options for TimeChorus). For example, for (5 seconds, on, adaptive), we first created an MP3 file by cutting each song in a playlist from five seconds before the first chorus to the end of the chorus and then connecting the songs with crossfade transitions. We then created an MP4 file to simulate the participants' experience of listening to the playlist on a smartphone application. Specifically, given an MP3 file, we created an MP4 file in which the MP3's playlist was played and images changed at the same time the song changed in the playlist. Each image contained a song's title and artist name like in the music playback screen of a music player application (Figure 1) [3].

4.3 Procedure

First, we investigated the participants' preferred option combinations. To this end, three dropdown lists with the options for each choice were presented to the participants. For each participant, three playlists were selected at random from the set of playlists described in section 4.2. Once the participant had selected an option for each choice, the three playlists (MP4 files) satisfying that option combination were displayed (Figure 1). When the participant changed the combination, the displayed playlists were also changed to those satisfying the new option combination [4]. After listening to playlists for any option combination, the participants reported their most preferred combination such as (10 seconds, on, 30 seconds), by selecting those options from the dropdown lists [5].

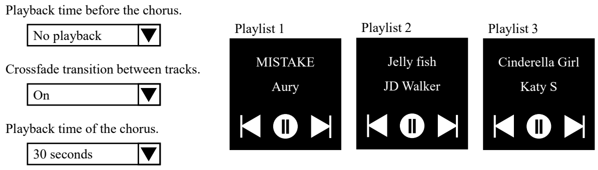

Figure 1. Interface example from our user survey. In this example, when a participant selects the options (no playback, on, 30 seconds), three playlists satisfying this combination are displayed.

Even if a participant chose "adaptive" as the preferred option for TimeChorus, she may have liked "30 seconds" almost as much. Accordingly, after the above investigation, we investigated the option preferences including such subtle differences. To this end, we displayed each of the 18 options with a six-point Likert scale ranging from "not preferred at all" to "very preferred," and we asked the participants to rate their preferences for each option.

4.4 Results

Table 1 and Figure 2 indicate the results for the first and second investigations, respectively. In Table 1, we can see that the most popular combination was (5 seconds, on, adaptive). Even participants who chose other combinations also tended to prefer each of these options. In fact, for the results shown in Figure 2, paired Wilcoxon signed-rank tests with Bonferroni correction revealed that the median of "5 seconds" was statistically higher than the other two options for TimePreChorus at $p < 0.01$ [6]. Similarly, "on" for Crossfade and "adaptive" for TimeChorus were statistically higher than the other options at $p < 0.01$. In particular, it was not obvious that "5 seconds" was the most preferred option for TimePreChorus, making this a useful, reusable insight for realizing chorus-playlist.

A music streaming service could offer the concept of chorus-playlist by implementing a function that enables users to listen to only choruses for all existing playlists on the service. If a service provided this function, it would be ideal to enable users to play playlists with arbitrary option combinations, as we did, to reflect users' preferences. If it

[2] Thus, if the first chorus is less than 15 (resp. 30) seconds long, the song continues play after the chorus until the total playback time reaches 15 (resp. 30) seconds.

[3] The images also include icons of pause, next, and previous buttons.

[4] Note that the songs contained in the three playlists did not change; only the options for playing them were changed.

[5] It was not mandatory to listen to the playlists for all 18 option combinations. In fact, most participants reported their most preferred combination by narrowing down their preferences while switching options and listening to the corresponding playlists.

[6] Throughout this paper, ** (*) in a figure denotes a statistical difference at $p < 0.01$ ($p < 0.05$).

Table 1. Preference distribution for option combinations.

TimePreChorus	Crossfade	TimeChorus	# participnats
No playback	On	15 seconds	10 (4.67%)
		30 seconds	3 (1.40%)
		adaptive	38 (17.76%)
	Off	15 seconds	8 (3.74%)
		30 seconds	0
		adaptive	8 (3.74%)
5 seconds	On	15 seconds	26 (12.15%)
		30 seconds	14 (6.54%)
		adaptive	**49 (22.90%)**
	Off	15 seconds	1 (0.47%)
		30 seconds	1 (0.47%)
		adaptive	16 (7.48%)
10 seconds	On	15 seconds	3 (1.40%)
		30 seconds	6 (2.80%)
		adaptive	22 (10.28%)
	Off	15 seconds	1 (0.47%)
		30 seconds	1 (0.47%)
		adaptive	7 (3.27%)

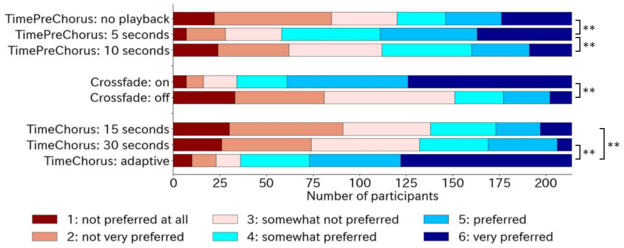

Figure 2. Preference distributions for each option.

is difficult to implement such a flexible function, chorus-playlist should be provided with the option combination (5 seconds, on, adaptive), which should maximize the average user satisfaction according to this section's results.

5. DEMAND FOR CHORUS-PLAYLIST

We described above how to implement a chorus-playlist function. In this section, we investigate the demand for listening to existing playlists with such a function.

5.1 Procedure

First, we showed the following description: "Suppose that a function to play only the choruses of the songs in a playlist has become available on the music streaming service that you usually use. Please rate on a scale from 1 (unwilling) to 6 (very willing) how much you would like to use this function to listen to existing playlists that you created." When the answer was "unwilling," they were asked to respond freely on why he/she did not want to use it. Otherwise, when the answer was one of the remaining five items, they were asked to respond freely with at least one objective for listening to playlists with the function. We did not set a cap on the number of responses.

Next, in a similar way, we asked the participants to indicate their willingness on a 6-point scale to use the function to listen to existing playlists created by other users. According to their willingness, they were asked to provide free responses as they did for the first question.

5.2 Results

Figure 3 shows the answer distribution. We refer to playlists created by the participant and by other users "self-made playlists" and "others' playlists." Only 11 and 3 participants answered "unwilling" for self-made playlists and

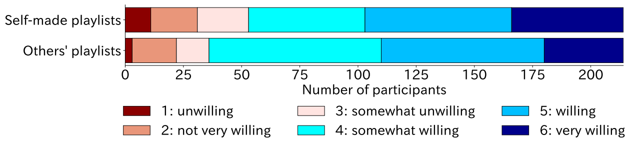

Figure 3. Distribution of the willingness to listen to playlists with the chorus-playlist function.

others' playlists, respectively. The most popular reason for unwillingness was "I believe that there is value in listening to the entire song, including parts other than the chorus." On the other hand, because 75.2 % (161) and 82.7 % (177) participants for the two respective types of playlists answered "very willing," "willing," or "somewhat willing," we conclude that there is a sufficiently high demand for chorus-playlist.

We manually grouped the free responses on their objectives. When a response included multiple objectives corresponding to different groups, it was assigned to multiple groups. Table 2 lists the top 10 objectives in terms of the group sizes for each of the two kinds of playlists. Each number in parentheses indicates the number of participants who gave that objective. Below, we discuss the results.

In the case of self-made playlists, the top three objectives could be achieved just by continuously listening to the choruses of songs in a playlist. For example, the first objective was "boost a mood." As self-made playlists usually contain songs that match the user's own music preferences, the user's mood would be boosted even when listening to playlists in the usual way [68]. Nevertheless, it is interesting that the participants answered that they wanted to further boost their mood by listening to only choruses. It would be beneficial to recommend songs for a playlist that are suitable for boosting a user's mood when the user listens to only the choruses in the playlist. For the second objective, the participants answered with various contexts such as "work" and "driving." When listening to a playlist with the chorus-playlist function in a specific context, there could be various reasons such as "increasing concentration" and "relaxing." It would be an interesting future work to conduct a more detailed analysis of the contexts in which the conventional playlist listening approach or the chorus-playlist approach are preferred. As for the fourth objective, it is known that people consider it valuable to listen to music with others and let others listen to their favorite songs [81–84]. However, because it takes much time for others to listen to all the songs in a playlist, people may hesitate to introduce their favorite songs. Thus, the participants answered that they wanted to efficiently introduce others such as friends or family to their favorite songs when they listen to music together in person. That is, chorus-playlist could encourage people to interact with others through music in the real world.

On the other hand, regarding the top six objectives for others' playlists, although those objectives could be achieved by conventional playlist listening, the participants wanted to use the chorus-playlist function to achieve the objectives more efficiently in a shorter time. In particular, as seen from the first, second, fourth, and fifth objectives, there is a strong demand for efficiently discovering and lis-

Table 2. Top 10 free-response objectives for listening to playlists with the chorus-playlist function.

Rank	Self-made playlist	Others' playlist
1	Boost a mood (79)	Find unfamiliar songs that suit my preference (109)
2	Listen to a playlist in a specific context (68)	Listen to hit songs (48)
3	Listen to a playlist within a limited time (46)	Learn other people's music preferences (44)
4	Recommend my favorite songs to others (36)	Listen to songs by unfamiliar artists (37)
5	Explore desired songs (35)	Listen to songs in unfamiliar genres (26)
6	Listen to many songs (23)	Preview a playlist (18)
7	Listen to various songs (20)	Listen to a playlist in a specific context (13)
8	Recall songs listened to in the past (18)	Refer to a playlist for creating my playlists (12)
9	Listen to only the choruses of my favorite songs (19)	Listen to a playlist for a change of pace (9)
10	Sing songs in a playlist (12)	Boost a mood (6)

tening to unfamiliar songs. Accordingly, if a music streaming service provides playlists consisting of hit songs of the past week, songs by a specific artist, or songs in a specific genre with the chorus-playlist function, many users would likely listen to those playlists by using the function. This would enable users to find more new favorite songs and help increase their music listening activity.

In summary, we have revealed that chorus-playlist can generate new interactions between people and music especially when they listen to self-made playlists. Moreover, as discussed above, the results in Table 2 can provide useful insights for both researchers and music streaming services.

6. IMPORTANT PROPERTIES FOR CREATING PLAYLISTS IN CHORUS-PLAYLIST CONCEPT

6.1 Procedure

In section 5, we assumed application of the chorus-playlist concept to existing playlists. However, by taking the concept a step further, a user could create a playlist on the premise of continuous listening to only the choruses in the playlist's songs (for simplicity, we refer to creating such a playlist as "creating a chorus-playlist"). Therefore, we asked the participants how much they would like to create chorus-playlists on the music streaming service that they used regularly. They answered with their willingness on a 6-point scale from "unwilling" to "very willing." We also told the participants that they could also listen to each song from beginning to end through, not just the chorus.

When users create playlists, they consider certain properties such as the diversity of artists and the song order. Hence, we wondered whether there are any differences regarding the importance of these properties when creating a chorus-playlist as compared to creating a usual playlist. To answer this question, we considered the following 11 properties derived from past studies [42, 70]. (1) SongHit: including songs with high popularity. (2) SongNew: including new songs in terms of the release dates. (3) ArtistSame: including as many songs by the same artist as possible. (4) ArtistDiv: including songs by as many different artists as possible. (5) GenreSame: including as many songs in the same genre as possible. (6) GenreDiv: including songs in as many different genres as possible. (7) MoodSame: including as many songs with the same musical mood as possible. (8) MoodDiv: including songs with as many different moods as possible. (9) SongOrder: the song order in the playlist. (10) SongTop: the first song in the playlist. (11) SongLast: the last song in the playlist.

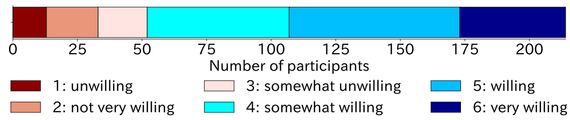

Figure 4. Willingness to create a chorus-playlist.

The participants who answered the first question with anything other than "unwilling" rated the importance of each property in creating a chorus-playlist and in creating a usual playlist on a 6-point scale from "not at all important" to "very important." The 11 properties were displayed in a random order to each participant.

6.2 Results

Figure 4 shows that chorus-playlist creation has the potential to be a new way of enjoying music, because 75.7% (162) participants answered "very willing," "willing," or "somewhat willing," while only 6.07% (13) participants answered "unwilling." Next, as shown in Figure 5, paired Wilcoxon signed-rank tests indicated that statistical differences between the two playlist types were observed for nine properties [7]. Hence, we can say that people tended to emphasize different properties when creating a chorus-playlist as compared to creating a usual playlist. Existing studies on song recommendation for playlists or playlist generation have proposed various methods focusing on the song order in a playlist [54, 61, 64, 73]. However, for chorus-playlist, the SongOrder, FirstSong, and LastSong properties were relatively less important. In contrast, hit songs and new songs were more important for chorus-playlist. Furthermore, the results revealed the importance of diversity in terms of artists, genres, and moods for chorus-playlist. It has been reported that the diversities of artists and genres tend to be low in usual playlists [70]; however, to support users creating chorus-playlists, it would be important to recommend songs to diversify such properties. These results thus open up new recommendation approaches in the MIR community.

7. PLAYBACK METHOD COMPARISON

We have revealed a high demand to try chorus-playlist. In this section, we investigate whether the chorus-playlist playback method is preferred to other playback methods.

7.1 Procedure

For comparison, we used the following two types of playlists. (1) Head-playlist: a user continuously listens to

[7] Figure 5 shows the results for the 201 participants besides the 13 participants who answered "unwilling." The same statistical differences were obtained even with only the aforementioned 162 participants.

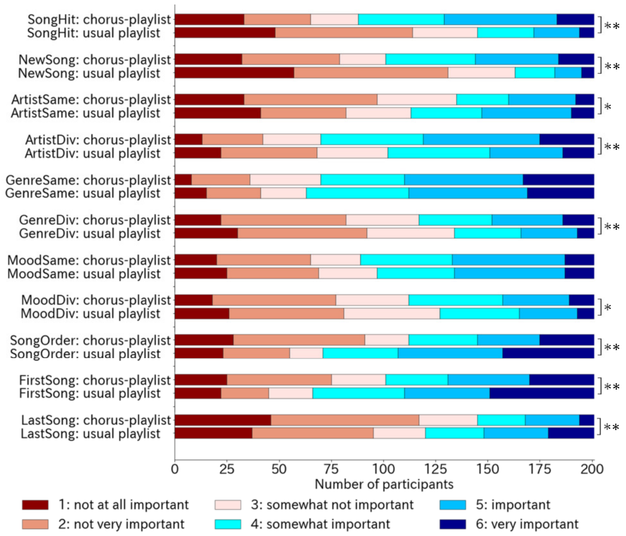

Figure 5. Property importance for playlist creation.

only the head sections of each song in a playlist. (2) 30sec-playlist: a user continuously listens to only the parts after the first 30 seconds of each song in a playlist. We adopted 30 seconds according to the preview samples on a music streaming service (Deezer) [85].

Similarly to the investigation described in section 4, we had the participants listen to each type of playlist by selecting options for two choices. In the head-playlist case, the two choices were 1-second crossfade (options: "on" and "off") and the playing time from the head of each song (options: "15 seconds" and "30 seconds"). In the case of 30sec-playlist, the two choices were 1-second crossfade (options: "on" and "off") and the playing time after the first 30 seconds of each song (options: "15 seconds" and "30 seconds"). The participants listened to playlists with each playback method by using their favorite option combinations. Here, each participant was assigned three playlists containing the same songs as those in section 4. Then, they were asked to rate their willingness to listen to self-made playlists with each method on a 6-point scale.

7.2 Results

Figure 6 shows the results. Note that the chorus-playlist results are repeated from "self-made playlists" in Figure 3. Paired Wilcoxon signed-rank tests with Bonferroni correction revealed that the median for chorus-playlist was statistically higher than the medians for head-playlist and 30sec-playlist. It thus became clear that it was not enough to simply play any part of the songs in a playlist continuously, but that it was important for users to be able to play choruses continuously.

8. DISCUSSION AND CONCLUSION

In this paper, we have studied the concept of chorus-playlist. The reusable insights obtained from our user survey can be summarized as follows.

- We showed that there is a high demand for chorus-playlist. When the participants listened to songs in a playlist with chorus-playlist, they tended to prefer to listen to 5 seconds before the chorus, add crossfade transitions between songs, and listen to the chorus from beginning to end. As discussed in section 7.2, it is more important to play choruses continuously than to play other sections continuously.
- As seen in Table 2, the objectives for listening to music with chorus-playlist were largely different between self-made playlists and others' playlists. In particular, certain objectives for self-made playlists were unexpected, in that people wanted to enjoy music in a new way with chorus-playlist for objectives such as boosting their mood. These results could provide guidelines for researchers and services to consider new research topics and activate user interaction, respectively.
- We revealed a high demand for creating a chorus-playlist. As in Figure 5, hit songs, new songs, and the diversities of artists, genres, and moods are more important when creating a chorus-playlist than when creating a usual playlist. These results also provide new viewpoints for studies on assisted playlist creation.

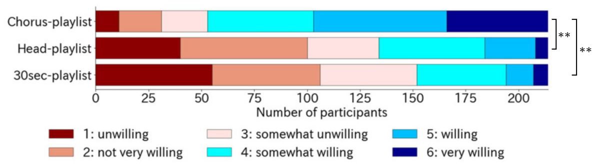

Figure 6. Willingness for three playlist types.

We acknowledge a limitation of this paper in that all the participants in our user survey were Japanese. Because peoples' music preferences, listening behaviors, and music itself vary widely from country to country [86–90], not all of the findings reported here can be generalized. Nevertheless, we believe that this study provides a worthwhile contribution as a first step toward understanding the impact of the chorus-playlist concept. At the same time, this limitation indicates further possibilities such as investigating the differences among countries and cultures. The publicly available dataset of results from our user survey will enable researchers to perform such comparisons.

Another limitation is that the participants did not experience chorus-playlist on the music streaming services they usually used. However, because they answered the survey after experiencing the chorus-playlist concept by listening to the playlists that we provided, we think that they could sufficiently imagine the situation of listening to self-made playlists and others' playlists with the chorus-playlist approach. We currently provide the chorus-playlist function in a music-related smartphone application (Vocacolle App) and web service (Kiite [8]). In the future, we will investigate the function's usage in those more realistic environments.

Finally, although we considered only the first chorus of a song (except when a song started with the chorus), the final chorus tends to be longer and contain heavier instrumentation than other choruses [1]. Therefore, it would also be an interesting future work to investigate the impact of differences between choruses in chorus-playlist listening. Moreover, the concept of listening to only choruses can be applied not only to playlists but also to other song lists such as album track lists, which could further enrich and diversify people's music listening experience.

[8] `https://kiite.jp`

9. ACKNOWLEDGMENTS

This work was supported in part by JST CREST Grant Number JPMJCR20D4 and JSPS KAKENHI Grant Number JP21H04917, Japan.

10. REFERENCES

[1] K. Stephenson, *What to listen for in rock: A stylistic analysis*. Yale University Press, 2002.

[2] J. Covach, "Form in rock music," *Engaging music: Essays in music analysis*, pp. 65–76, 2005.

[3] W. Everett, *The foundations of rock: from "Blue suede shoes" to "Suite: Judy blue eyes"*. Oxford University Press, 2008.

[4] R. Middleton, "Song form," *The Continuum Encyclopedia of Popular Music of the World*, vol. 2, pp. 513–519, 2003.

[5] A. Eronen, "Chorus detection with combined use of MFCC and chroma features and image processing filters," in *Proceedings of the 10th International Conference on Digital Audio Effects*, ser. DAFx 2007, 2007, pp. 229–236.

[6] S. Hult, L. B. Kreiberg, S. S. Brandt, and B. T. Jónsson, "Analysis of the effect of dataset construction methodology on transferability of music emotion recognition models," in *Proceedings of the 2020 International Conference on Multimedia Retrieval*, ser. ICMR 2020, 2020, pp. 316–320.

[7] I. Salakka, A. Pitkäniemi, E. Pentikäinen, K. Mikkonen, P. Saari, P. Toiviainen, and T. Särkämö, "What makes music memorable? Relationships between acoustic musical features and music-evoked emotions and memories in older adults," *PlOS ONE*, vol. 16, no. 5, pp. 1–18, 2021.

[8] J. Paulus, M. Müller, and A. Klapuri, "State of the art report: Audio-based music structure analysis," in *Proceedings of the 11th International Society for Music Information Retrieval Conference*, ser. ISMIR 2010, 2010, pp. 625–636.

[9] M. Müller, P. Grosche, and N. Jiang, "A segment-based fitness measure for capturing repetitive structures of music recordings," in *Proceedings of the 12th International Society for Music Information Retrieval Conference*, ser. ISMIR 2011, 2011, pp. 615–620.

[10] J. Foote, "Automatic audio segmentation using a measure of audio novelty," in *Proceedings of the 2000 IEEE International Conference on Multimedia and Expo*, ser. ICME 2000, 2000, pp. 452–455.

[11] M. Cooper and J. Foote, "Summarizing popular music via structural similarity analysis," in *Proceedings of the 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, ser. WASPAA 2003, 2003, pp. 127–130.

[12] M. Goto, "A chorus section detection method for musical audio signals and its application to a music listening station," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 5, pp. 1783–1794, 2006.

[13] M. Müller and S. Ewert, "Joint structure analysis with applications to music annotation and synchronization," in *Proceedings of the 9th International Conference on Music Information Retrieval*, ser. ISMIR 2008, 2008, pp. 389–394.

[14] J. Paulus and A. Klapuri, "Music structure analysis using a probabilistic fitness measure and a greedy search algorithm," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 6, pp. 1159–1170, 2009.

[15] J. Serrà, M. Müller, P. Grosche, and J. L. Arcos, "Unsupervised detection of music boundaries by time series structure features," in *Proceedings of the 26th AAAI Conference on Artificial Intelligence*, ser. AAAI 2012, 2012, pp. 1613–1619.

[16] A. Baratè, L. A. Ludovico, and E. Santucci, "A semantics-driven approach to lyrics segmentation," in *Proceedings of the 8th International Workshop on Semantic and Social Media Adaptation and Personalization*, ser. SMAP 2013, 2013, pp. 73–79.

[17] O. Nieto and T. Jehan, "Convex non-negative matrix factorization for automatic music structure identification," in *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, ser. ICASSP 2013, 2013, pp. 236–240.

[18] F. Kaiser and G. Peeters, "A simple fusion method of state and sequence segmentation for music structure discovery," in *Proceedings of the 14th International Society for Music Information Retrieval Conference*, ser. ISMIR 2013, 2013, pp. 257–262.

[19] H. Grohganz, M. Clausen, N. Jiang, and M. Müller, "Converting path structures into block structures using eigenvalue decompositions of self-similarity matrices," in *Proceedings of the 14th International Society for Music Information Retrieval Conference*, ser. ISMIR 2013, 2013, pp. 209–214.

[20] G. Peeters and V. Bisot, "Improving music structure segmentation using lag-priors," in *Proceedings of the 15th International Society for Music Information Retrieval Conference*, ser. ISMIR 2014, 2014, pp. 337–342.

[21] B. McFee and D. Ellis, "Analyzing song structure with spectral clustering," in *Proceedings of the 15th International Society for Music Information Retrieval Conference*, ser. ISMIR 2014, 2014, pp. 405–410.

[22] C.-H. Yeh, W.-Y. Tseng, C.-Y. Chen, Y.-D. Lin, Y.-R. Tsai, H.-I. Bi, Y.-C. Lin, and H.-Y. Lin, "Popular music

representation: Chorus detection & emotion recognition," *Multimedia Tools and Applications*, vol. 73, pp. 2103–2128, 2014.

[23] T. Grill and J. Schlüter, "Music boundary detection using neural networks on combined features and two-level annotations," in *Proceedings of the 16th International Society for Music Information Retrieval Conference*, ser. ISMIR 2015, 2015, pp. 531–537.

[24] J. B. L. Smith and M. Goto, "Using priors to improve estimates of music structure," in *Proceedings of the 17th International Society for Music Information Retrieval Conference*, ser. ISMIR 2016, 2016, pp. 554–560.

[25] K. Watanabe, Y. Matsubayashi, N. Orita, N. Okazaki, K. Inui, S. Fukayama, T. Nakano, J. B. L. Smith, and M. Goto, "Modeling discourse segments in lyrics using repeated patterns," in *Proceedings of the 26th International Conference on Computational Linguistics*, ser. COLING 2016, 2016, pp. 1959–1969.

[26] G. Sargent, F. Bimbot, and E. Vincent, "Estimating the structural segmentation of popular music pieces under regularity constraints," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 2, pp. 344–358, 2017.

[27] T. Cheng, J. B. L. Smith, and M. Goto, "Music structure boundary detection and labelling by a deconvolution of path-enhanced self-similarity matrix," in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing*, ser. ICASSP 2018, 2018, pp. 106–110.

[28] M. Fell, Y. Nechaev, E. Cabrio, and F. Gandon, "Lyrics segmentation: Textual macrostructure detection using convolutions," in *Proceedings of the 27th International Conference on Computational Linguistics*, ser. COLING 2018, 2018, pp. 2044–2054.

[29] Y. Huang, S. Chou, and Y. Yang, "Pop Music Highlighter: Marking the emotion keypoints," *Transactions of the International Society for Music Information Retrieval*, vol. 1, no. 1, pp. 68–78, 2018.

[30] A. Maezawa, "Music boundary detection based on a hybrid deep model of novelty, homogeneity, repetition and duration," in *Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing*, ser. ICASSP 2019, 2019, pp. 206–210.

[31] G. Shibata, R. Nishikimi, E. Nakamura, and K. Yoshii, "Statistical music structure analysis based on a homogeneity-, repetitiveness-, and regularity-aware hierarchical hidden semi-markov model," in *Proceedings of the 20th International Society for Music Information Retrieval Conference*, ser. ISMIR 2019, 2019, pp. 268–275.

[32] G. Shibata, R. Nishikimi, and K. Yoshii, "Music structure analysis based on an LSTM-HSMM hybrid model," in *Proceedings of the 21th International Society for Music Information Retrieval Conference*, ser. ISMIR 2020, 2020, pp. 23–29.

[33] K. Watanabe and M. Goto, "A chorus-section detection method for lyrics text," in *Proceedings of the 21st International Society for Music Information Retrieval Conference*, ser. ISMIR 2020, 2020, pp. 351–359.

[34] J. Wang, J. B. L. Smith, J. Chen, X. Song, and Y. Wang, "Supervised chorus detection for popular music using convolutional neural network and multi-task learning," in *Proceedings of the 2021 IEEE International Conference on Acoustics, Speech and Signal Processing*, ser. ICASSP 2021, 2021, pp. 566–570.

[35] J. Wang, Y. Hung, and J. B. L. Smith, "To catch a chorus, verse, intro, or anything else: Analyzing a song with structural functions," in *Proceedings of the 2022 IEEE International Conference on Acoustics, Speech and Signal Processing*, ser. ICASSP 2022, 2022, pp. 416–420.

[36] S. Gao and H. Li, "Popular song summarization using chorus section detection from audio signal," in *Proceedings of the IEEE 17th International Workshop on Multimedia Signal Processing*, ser. MMSP 2015, 2015, pp. 1–6.

[37] Y. Gao, Y. Shen, X. Zhang, S. Yu, and W. Li, "Music summary detection with state space embedding and recurrence plot," in *Proceedings of the 6th Conference on Sound and Music Technology*, ser. CSMT 2019, 2019, pp. 41–51.

[38] K. Zhang, H. Zhang, S. Li, C. Yang, and L. Sun, "The PMEmo dataset for music emotion recognition," in *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval*, ser. ICMR 2018, 2018, pp. 135–142.

[39] J. H. Lee and N. M. Waterman, "Understanding user requirements for music information services," in *Proceedings of the 13th International Society for Music Information Retrieval Conference*, ser. ISMIR 2012, 2012, pp. 253–258.

[40] L. Porcaro and E. Gomez, "20 years of playlists: A statistical analysis on popularity and diversity," in *Proceedings of the 20th International Society for Music Information Retrieval Conference*, ser. ISMIR 2019, 2019, pp. 130–136.

[41] I. Kamehkhosh, G. Bonnin, and D. Jannach, "Effects of recommendations on the playlist creation behavior of users," *User Modeling and User-Adapted Interaction*, vol. 30, no. 2, pp. 285–322, 2020.

[42] H. V. Schweiger, E. Parada-Cabaleiro, and M. Schedl, "Does track sequence in user-generated playlists matter?," in *Proceedings of the 22nd International Society*

for Music Information Retrieval Conference, ser. ISMIR 2021, 2021, pp. 618–625.

[43] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, "Controllable music playlist generation based on knowledge graph and reinforcement learning," *Sensors*, vol. 22, no. 10, pp. 135–142, 2022.

[44] G. Bonnin and D. Jannach, "Automated generation of music playlists: Survey and experiments," *ACM Computing Surveys*, vol. 47, no. 2, pp. 1–35, 2014.

[45] K. Jacobson, V. Murali, E. Newett, B. Whitman, and R. Yon, "Music personalization at Spotify," in *Proceedings of the 10th ACM Conference on Recommender Systems*, ser. RecSys 2016, 2016, p. 373.

[46] T. Bontempelli, B. Chapus, F. Rigaud, M. Morlon, M. Lorant, and G. Salha-Galvan, "Flow Moods: Recommending music by moods on deezer," in *Proceedings of the 16th ACM Conference on Recommender Systems*, ser. RecSys 2022, 2022, pp. 452–455.

[47] J. J. Aucouturier and F. Pachet, "Scaling up music playlist generation," in *Proceedings of the 2002 IEEE International Conference on Multimedia and Expo*, ser. ICME 2002, 2002, pp. 105–108.

[48] B. Logan, "Content-based playlist generation: Exploratory experiments," in *Proceedings of the 3rd International Conference on Music Information Retrieval*, ser. ISMIR 2002, 2002, pp. 295–296.

[49] S. Pauws and B. Eggen, "PATS: Realization and user evaluation of an automatic playlist generator," in *Proceedings of the 3rd International Conference on Music Information Retrieval*, ser. ISMIR 2002, 2002, pp. 222–230.

[50] R. V. Gulik and F. Vignoli, "Visual playlist generation on the artist map," in *Proceedings of the 6th International Conference on Music Information Retrieval*, ser. ISMIR 2005, 2005, pp. 520–523.

[51] E. Pampalk, T. Pohle, and G. Widmer, "Dynamic playlist generation based on skipping behavior," in *Proceedings of the 6th International Conference on Music Information Retrieval*, ser. ISMIR 2005, 2005, pp. 634–637.

[52] S. Pauws and S. V. D. Wijdeven, "User evaluation of a new interactive playlist generation concept," in *Proceedings of the 6th International Conference on Music Information Retrieval*, ser. ISMIR 2005, 2005, pp. 638–643.

[53] S. J. Cunningham, D. Bainbridge, and A. Falconer, ""more of an art than a science": Supporting the creation of playlists and mixes," in *Proceedings of the 7th International Conference on Music Information Retrieval*, ser. ISMIR 2006, 2006, pp. 240–245.

[54] E. Pampalk and M. Gasser, "An implementation of a simple playlist generator based on audio similarity measures and user feedback," in *Proceedings of the 7th International Conference on Music Information Retrieval*, ser. ISMIR 2006, 2006, pp. 389–390.

[55] N. Oliver and L. Kreger-Stickles, "PAPA: Physiology and purpose-aware automatic playlist generation," in *Proceedings of the 7th International Conference on Music Information Retrieval*, ser. ISMIR 2006, 2006, pp. 250–253.

[56] S. Pauws, W. Verhaegh, and M. Vossen, "Fast generation of optimal music playlists using local search," in *Proceedings of the 7th International Conference on Music Information Retrieval*, ser. ISMIR 2006, 2006, pp. 138–143.

[57] B. Fields, C. Rhodes, M. A. Casey, and K. Jacobson, "Social playlists and bottleneck measurements: Exploiting musician social graphs using content-based dissimilarity and pairwise maximum flow values," in *Proceedings of the 9th International Conference on Music Information Retrieval*, ser. ISMIR 2008, 2008, pp. 559–564.

[58] A. Flexer, D. Schnitzer, M. Gasser, and G. Widmer, "Playlist generation using start and end songs," in *Proceedings of the 9th International Conference on Music Information Retrieval*, ser. ISMIR 2008, 2008, pp. 173–178.

[59] K. Bosteels, E. Pampalk, and E. E. Kerre, "Evaluating and analysing dynamic playlist generation heuristics using radio logs and fuzzy set theory," in *Proceedings of the 10th International Society for Music Information Retrieval Conference*, ser. ISMIR 2009, 2009, pp. 351–356.

[60] F. Maillet, D. Eck, G. Desjardins, and P. Lamere, "Steerable playlist generation by learning song similarity from radio station playlists," in *Proceedings of the 10th International Society for Music Information Retrieval Conference*, ser. ISMIR 2009, 2009, pp. 345–350.

[61] B. McFee and G. R. G. Lanckriet, "The natural language of playlists," in *Proceedings of the 12th International Society for Music Information Retrieval Conference*, ser. ISMIR 2011, 2011, pp. 537–542.

[62] J. H. Lee, "How similar is too similar?: Exploring users' perceptions of similarity in playlist evaluation," in *Proceedings of the 12th International Society for Music Information Retrieval Conference*, ser. ISMIR 2011, 2011, pp. 109–114.

[63] B. McFee and G. R. G. Lanckriet, "Hypergraph models of playlist dialects," in *Proceedings of the 13th International Society for Music Information Retrieval Conference*, ser. ISMIR 2012, 2012, pp. 343–348.

[64] J. L. Moore, S. Chen, T. Joachims, and D. Turnbull, "Learning to embed songs and tags for playlist prediction," in *Proceedings of the 13th International Society for Music Information Retrieval Conference*, ser. ISMIR 2012, 2012, pp. 349–354.

[65] G. Bonnin and D. Jannach, "Evaluating the quality of generated playlists based on hand-crafted samples," in *Proceedings of the 14th International Society for Music Information Retrieval Conference*, ser. ISMIR 2013, 2013, pp. 263–268.

[66] D. Jannach, L. Lerche, and I. Kamehkhosh, "Beyond "hitting the hits": Generating coherent music playlist continuations with the right tracks," in *Proceedings of the 9th ACM Conference on Recommender Systems*, ser. RecSys '15, 2015, pp. 187–194.

[67] T. Nakano, J. Kato, M. Hamasaki, and M. Goto, "PlaylistPlayer: An interface using multiple criteria to change the playback order of a music playlist," in *Proceedings of the 21st International Conference on Intelligent User Interfaces*, ser. IUI 2016, 2016, pp. 186–190.

[68] M. Pichl, E. Zangerle, and G. Specht, "Understanding playlist creation on music streaming platforms," in *Proceedings of the 2016 IEEE International Symposium on Multimedia*, ser. ISM 2016, 2016, pp. 475–480.

[69] C. Chung, Y. Chen, and H. H. Chen, "Exploiting playlists for representation of songs and words for text-based music retrieval," in *Proceedings of the 18th International Society for Music Information Retrieval Conference*, ser. ISMIR 2017, 2017, pp. 478–485.

[70] R. Dias, D. Gonçalves, and M. J. Fonseca, "From manual to assisted playlist creation: A survey," *Multimedia Tools and Applications*, vol. 76, no. 12, pp. 14 375–14 403, 2017.

[71] L. F. Pontello, P. H. F. Holanda, B. Guilherme, J. P. V. Cardoso, O. Goussevskaia, and A. P. C. D. Silva, "Mixtape: Using real-time user feedback to navigate large media collections," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 13, no. 4, pp. 1–22, 2017.

[72] R. M. Bittner, M. Gu, G. Hernandez, E. J. Humphrey, T. Jehan, H. McCurry, and N. Montecchio, "Automatic playlist sequencing and transitions," in *Proceedings of the 18th International Society for Music Information Retrieval Conference*, ser. ISMIR 2017, 2017, pp. 442–448.

[73] S. Shih and H. Chi, "Automatic, personalized, and flexible playlist generation using reinforcement learning," in *Proceedings of the 19th International Society for Music Information Retrieval Conference*, ser. ISMIR 2018, 2018, pp. 168–174.

[74] S. Y. Park, A. Laplante, J. H. Lee, and B. Kaneshiro, "Tunes together: Perception and experience of collaborative playlists," in *Proceedings of the 20th International Society for Music Information Retrieval Conference*, ser. ISMIR 2019, 2019, pp. 723–730.

[75] A. Patwari, N. Kong, J. Wang, U. Gargi, M. Covell, and A. Jansen, "Semantically meaningful attributes from co-listen embeddings for playlist exploration and expansion," in *Proceedings of the 21st International Society for Music Information Retrieval Conference*, ser. ISMIR 2020, 2020, pp. 527–533.

[76] Z. Li, M. Song, S. Duan, and Z. Wang, "Are users attracted by playlist titles and covers? Understanding playlist selection behavior on a music streaming platform," *Journal of Innovation & Knowledge*, vol. 7, no. 3, pp. 1–14, 2022.

[77] J. V. Balen, J. A. Burgoyne, F. Wiering, and R. C. Veltkamp, "An analysis of chorus features in popular song," in *Proceedings of the 14th Society of Music Information Retrieval Conference*, ser. ISMIR 2013, 2013, pp. 107–112.

[78] M. Goto, K. Yoshii, H. Fujihara, M. Mauch, and T. Nakano, "Songle: A web service for active music listening improved by user contributions," in *Proceedings of the 12th International Society for Music Information Retrieval Conference*, ser. ISMIR 2011, 2011, pp. 311–316.

[79] A. Kim, S. Park, J. Park, J.-W. Ha, T. Kwon, and J. Nam, "Automatic DJ mix generation using highlight detection," in *Proceedings of the 18th International Society for Music Information Retrieval Conference*, ser. ISMIR 2017, 2017, pp. 1–2.

[80] Y.-S. Huang, S.-Y. Chou, and Y.-H. Yang, "Generating music medleys via playing music puzzle games," in *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, ser. AAAI 2018, 2018, pp. 2281–2288.

[81] A. Bassoli, J. Moore, S. Agamanolis, and H. C. Group, "tunA: Local music sharing with handheld Wi-Fi devices," in *Proceedings of the 5th Wireless World Conference 2004*, ser. WWC 2004, 2004, pp. 1–23.

[82] M. Håkansson, M. Rost, and L. E. Holmquist, "Gifts from friends and strangers: A study of mobile music sharing," in *Proceedings of the 10th European Conference on Computer-Supported Cooperative Work*, ser. ECSCW 2007, 2007, pp. 311–330.

[83] M. Håkansson, M. Rost, M. Jacobsson, and L. E. Holmquist, "Facilitating mobile music sharing and social interaction with Push!Music," in *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, ser. HICSS 2007, 2007, pp. 87–96.

[84] K. Tsukuda, K. Ishida, M. Hamasaki, and M. Goto, "Kiite Cafe: A web service for getting together virtually to listen to music." in *Proceedings of the 22nd*

International Society for Music Information Retrieval Conference, ser. ISMIR 2021, 2021, pp. 697–704.

[85] K. M. Ibrahim, J. Royo-Letelier, E. V. Epure, G. Peeters, and G. Richard, "Audio-based auto-tagging with contextual tags for music," in *Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing*, ser. ICASSP 2020, 2020, pp. 16–20.

[86] X. Hu and J. H. Lee, "A cross-cultural study of music mood perception between American and Chinese listeners," in *Proceedings of the 13th International Society for Music Information Retrieval Conference*, ser. ISMIR 2012, 2012, pp. 535–540.

[87] Y. Yang and X. Hu, "Cross-cultural music mood classification: A comparison on English and Chinese songs," in *Proceedings of the 13th International Society for Music Information Retrieval Conference*, ser. ISMIR 2012, 2012, pp. 19–24.

[88] X. Hu, J. H. Lee, K. Choi, and J. S. Downie, "A cross-cultural study on the mood of K-POP songs," in *Proceedings of the 15th International Society for Music Information Retrieval Conference*, ser. ISMIR 2014, 2014, pp. 385–390.

[89] M. Liu, X. Hu, and M. Schedl, "Artist preferences and cultural, socio-economic distances across countries: A big data perspective," in *Proceedings of the 18th International Society for Music Information Retrieval Conference*, ser. ISMIR 2017, 2017, pp. 103–111.

[90] C. Bauer and M. Schedl, "Global and country-specific mainstreaminess measures: Definitions, analysis, and usage for improving personalized music recommendation systems," *PLOS ONE*, vol. 14, no. 6, pp. 1–36, 2019.