IEICE TRANS. INF. & SYST., VOL.E106–D, NO.4 APRIL 2023

PAPER

Why and How People View Lyrics While Listening to Music on a Smartphone*

Kosetsu TSUKUDA[†a)], **Masahiro HAMASAKI**[†b)], ***Nonmembers***, ***and*** **Masataka GOTO**[†c)], ***Fellow***

SUMMARY Why and how do people view lyrics? Although various lyrics-based music systems have been proposed, this fundamental question remains unexplored. Better understanding of lyrics viewing behavior would be beneficial for both researchers and music streaming platforms to improve their lyrics-based systems. Therefore, in this paper, we investigate why and how people view lyrics, especially when they listen to music on a smartphone. To answer "why," we conduct a questionnaire-based online user survey involving 206 participants. To answer "how," we analyze over 23 million lyrics request logs sent from the smartphone application of a music streaming service. Our analysis results suggest several reusable insights, including the following: (1) People have high demand for viewing lyrics to confirm what the artist sings, more deeply understand the lyrics, sing the song, and figure out the structure such as verse and chorus. (2) People like to view lyrics after returning home at night and before going to sleep rather than during the daytime. (3) People usually view the same lyrics repeatedly over time. Applying these insights, we also discuss application examples that could enable people to more actively view lyrics and listen to new songs, which would not only diversify and enrich people's music listening experiences but also be beneficial especially for music streaming platforms.

key words: *lyrics, smartphone, user behavior analysis, log analysis*

1. Introduction

When people seek help in identifying a particular song that they have listened to, they often provide words in the song's lyrics as a clue for identification [2], [3]. In other situations when people listen to music, it has been reported that they choose songs according to not only the musical audio content, such as the music genre, mood, melody, vocal timbre, and rhythm, but also the topics of lyrics [4], [5]. To meet these demands, in the field of Music Information Retrieval (MIR), researchers have proposed systems for identifying a song by using the words in lyrics as a query [6]–[9] and systems for exploring songs according to the topics estimated from lyrics [10]–[13]. As illustrated here, lyrics are an essential element of music for both listeners and MIR researchers.

Despite the importance of lyrics, more fundamental investigation of lyrics remains an under-addressed topic: why and how do people view lyrics? In this paper, we aim to answer these questions. Investigating people's lyrics-viewing behavior and revealing reusable insights would be beneficial for researchers and music streaming platforms to implement lyrics-related systems and functions, such as viewing support for lyrics and song recommendation based on lyrics. With regard to music listening, researchers have investigated why and how people listen to music [14]–[21], and the obtained insights have contributed to later studies in the MIR community. Although listening to music includes listening to sung lyrics, our study differs from these studies in that we focus on lyrics-viewing behavior.

Users can view lyrics in various ways, such as a lyrics sheet included with a compact disc (CD), a web service for lyrics search, and a YouTube video with lyrics overlaid [22]. Recently, some smartphone applications for online music services (*e.g.*, Spotify and Apple Music) have provided a function that enables a user to view song's lyrics while listening to the song. Such a function will become one of the main means for viewing lyrics, given the current situation in which music streaming services on smartphones have become a mainstream format for listening to music [16]. In light of the above, we investigate the behavior of viewing lyrics on a smartphone while listening to music, because we can make the obtained insights more reusable for future work.

Our main contributions can be summarized as follows:

- To our knowledge, this is the first study on the interactions between users and lyrics in terms of why and how users view lyrics when they listen to music.
- To investigate why users view lyrics, we conducted a large-scale questionnaire-based online survey involving 206 participants. In the survey, more than 75% of the participants answered that they often view lyrics to confirm what an artist sings or more deeply understand lyrics. Moreover, over 50% of the participants often view lyrics to sing a song or figure out the structure of the lyrics (verse, chorus, *etc.*). These results are beneficial for both MIR researchers and music streaming platforms to implement their systems or functions. In fact, in this paper, we suggest examples of functions to support users according to their reasons for viewing lyrics, such as a function that displays tips to sing each part of the song's lyrics for users who want to sing.
- We investigated how users view lyrics by analyzing over 23 million lyrics request logs for over 600 thousand smartphone users for a year on a music stream-

Manuscript received September 30, 2022.
Manuscript publicized January 18, 2023.
[†]The authors are with National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba-shi, 305–8568 Japan.
[*]The original version of this paper was published at a conference [1].
a) E-mail: k.tsukuda@aist.go.jp
b) E-mail: masahiro.hamasaki@aist.go.jp
c) E-mail: m.goto@aist.go.jp
DOI: 10.1587/transinf.2022EDP7177

ing service. The data shows that people tend to view more lyrics after coming back home at night and before going to bed. In addition, an average of 37.8% of user's viewed lyrics have already been viewed by the user, and eventually the user gets bored with viewing the same lyrics. Considering these findings, we make several proposals for music streaming platforms to attract users (*e.g.*, when a user gets bored with the lyrics of a song, the platform could suggest related lyrics in terms of the topic).

2. Related Work

2.1 User Behavior in Music Listening

One approach to analyze user behavior in music listening is conducting user studies based on questionnaires and interviews. Typical questions about music listening ask why people listen to music [17]–[19], [21] and how they use music websites, services, and applications [2], [16], [23]. Regarding the former question, the main reasons include emotional reasons such as relaxation [18] (even at work [19]) and relief [17]. People also listen to music to concentrate and to pass time [21]. Regarding the latter question, Lee and Waterman [16] revealed that people use music websites and applications for various reasons such as discovering new music and learning about the artists. They also compared their results with those in 2004 [2] and showed increases in the popularity of music streaming and mobile music consumption. A more recent work conducted a survey on the use of cloud music services and considered the future design of such services [23]. Moreover, Lee and Price [15] conducted interviews with music listeners and revealed seven typical personas, such as a user who enjoys curating music that is already familiar and a user who enjoys serendipitous music discovery.

Another approach is analyzing users' play logs. These logs are typically collected from (1) APIs provided by online music services [24], [25] or (2) Twitter, where tweets related to music listening are gathered via specific tags such as "#nowplaying" and "#itunes" [26]–[28]. Logs have been analyzed in terms of various aspects, including the long tail distribution of listening events per user, track, and artist [24], [25], [28], the popularity of genres, moods, and tags [26]–[28], and the temporal distribution (hour of day and day of week) [24], [27], *etc.*. One characteristic of music listening behavior is *repeat consumption* [29]. Reports have indicated that, in a user's music play logs, about 70% of played songs have already been played before, and this percentage is much higher than for other domains such as viewing videos and visiting restaurants [29], [30]. In repeat consumption, the number of times a song is played is heavy-tailed (*i.e.*, a user repeatedly listens to a small proportion of songs again and again). Benson et al. [30] reported that each song has its own *lifetime* for a user: at the beginning of the lifetime, the temporal gap between listening events is small; but at the end of the lifetime, the gap becomes large, and eventually the user becomes bored with the song.

Although listening to sung lyrics is one factor in listening to music, our study differs from the above studies in that we particularly focus on lyrics viewing behavior. Focusing on a particular element of music is beneficial to suggest new possibilities for future research as was indicated by Demetriou et al. [5] who focused on vocals. Research on *why* people listen to music has tended to involve user studies, because they have the advantage of enabling researchers to ask questions to analyze people's intent. In contrast, research on *how* people listen to music has often analyzed large log data to take advantage of statistical processing. Applying both of these advantages, in this paper, we investigate why and how people view lyrics by using questionnaires and logs, respectively.

2.2 Lyrics in MIR

Researchers have considered lyrics in various studies, including lyrics-to-audio alignment [31]–[37], analysis of lyrics characteristics [38]–[43], accurate lyrics retrieval [44]–[46], and genre and mood classification [47]–[52]. Below, we review more related studies that aim to support user activity by using lyrics.

One major approach is enabling users to search for songs by words in lyrics, in which a query can be text [6], [7] or user's sung lyrics [8], [9]. Systems have also been proposed for exploring songs according to topics estimated from lyrics [10]–[13]. Fujihara et al. [53] proposed the concept of a "Music Web" in which songs are hyperlinked to each other based on phrases of lyrics. Visualization is also a useful approach to browse a music collection. *SongWords* [54] displays a music collection on a two-dimensional canvas based on self-organizing maps for lyrics and tags. *Lyricon* [55] is a system for displaying icons that match the word sequences of lyrics so that users can intuitively understand the lyrics. Moreover, Funasawa et al. [56] implemented a system that automatically generates slideshows for music by generating queries from lyrics and searching for images. O'Hara et al. [57] demonstrated how to learn the meanings of chord sequences from lyrics annotated with chords. Ibraham et al. [58] proposed a method for estimating the intelligibility of lyrics in a given song to help users learn a second language.

In this paper, we investigate more fundamental questions about lyrics: why and how people view them. For researchers, the insights of our analysis can be used in implementing lyrics-based systems. For example, when researchers propose systems to support understanding lyrics, they can claim these systems' importance based on the high demand for deeply understanding lyrics, as we will report in Sect. 3.2.1. In Sects. 3 and 4, we also suggest application examples such as recommending songs according to lyrics and supporting lyrics viewing. We believe that our suggestions are also beneficial for music streaming platforms to make their smartphone applications more attractive to users.

3. Why People View Lyrics

In this section, we report why people view lyrics by conducting an online survey involving 206 participants.

3.1 Participants

We recruited participants for our survey via an online research company. We limited the participants to those who listen to music on average at least one day per week on a smartphone application via any online music service and have viewed lyrics on the application while listening to music at least 10 times in their lifetime. In addition, to align with the user nationality in the lyrics viewing log data, as described in Sect. 4.1.1, all participants were Japanese. The participants answered our questionnaire through a web browser. We paid about 15.73 USD (2,275 JPY) to each participant. Although 297 participants joined the survey, to make the analysis results more reliable, we removed the answers from 91 participants: 14 of them gave the same answers to all questions (*e.g.*, choosing "1" for all questions), and 77 of them finished answering the questions in a very short time[†]. The remaining 206 participants were well balanced in gender and age range: 95 males (10s: 2; 20s: 20; 30s: 22; 40s: 26; 50s: 25) and 111 females (10s: 4; 20s: 21; 30s: 27; 40s: 28; 50s: 31).

3.2 Results and Discussion

3.2.1 Reasons

To understand why people view lyrics on a smartphone while listening to music, we listed the following eight candidate reasons.

(1) *Confirmation*: The user wants to confirm what the artist sings.
(2) *Understanding*: The user wants to more deeply understand the lyrics.
(3) *Singing*: The user wants to sing to herself (not in public).
(4) *Structure*: The user wants to figure out the structure of the lyrics, such as verse and chorus.
(5) *Karaoke*: The user wants to practice for singing in public, as in karaoke.
(6) *Boredom*: The user wants to get rid of her boredom by viewing lyrics.
(7) *Language*: The user wants to learn a language with the lyrics.
(8) *Writing*: The user wants to study for writing lyrics.

The participants were asked to rate the frequency of viewing lyrics for each reason on a scale of 1 to 5 (1: never; 5: very often). The reasons were displayed in a random order to each participant[††].

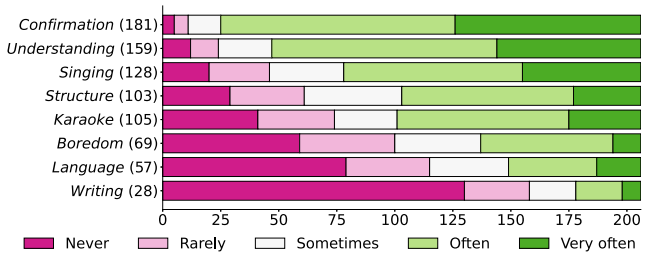

Fig. 1 Frequency of reasons why people view lyrics on a smartphone while listening to music (1: never; 5: very often). The number in parentheses represents the number of participants rating 4 or 5.

For each reason, Fig. 1 shows the frequency distribution and the number of users whose score was 4 or 5 (*i.e.*, the number who often viewed lyrics for that reason). We can see that the ratings for *Confirmation* and *Understanding* are high: in fact, the paired Wilcoxon signed-rank tests with Bonferroni correction reveal that the medians of *Confirmation* and *Understanding* are statistically higher than the remaining six reasons at $p < 0.01$. It would be beneficial to provide additional functions according to users' reasons for viewing lyrics. For example, for a user whose reason is *Understanding*, displaying diverse interpretations of lyrics could help her understand them more deeply. An interesting future work would be to automatically mine web pages that describe interpretations of given song's lyrics and display the collected interpretations along with the lyrics.

Among the remaining six reasons, more than half of the participants gave a rating of 4 or 5 for *Singing*, *Structure*, and *Karaoke*. For users who view lyrics to sing (*Singing* and *Karaoke*), some smartphone applications already provide a function that automatically scrolls lyrics by synchronizing them with the playback time. To improve their singing performance, we suggest more advanced functions that display tips for singing each part of the lyrics and automatically judge their singing skill [59]. As for the *Structure* reason, one possible application is coloring blocks of lyrics according to the estimated structure [60], [61]; this would enable the user to quickly figure out the structure.

Although *Boredom*, *Language*, and *Writing* are relatively minor reasons, it is still worth considering functions for them, not only because it is important to build systems to support niche uses but also because more users may begin to view lyrics to use such functions. This may give users chances to listen to music more frequently and eventually provide benefits for music streaming platforms. For a user who views lyrics because of *Boredom*, displaying information related to the played song, such as similar songs by dif-

[†]We applied a tight rule for this filtering to reduce the risk of noisy answers as much as possible. Nonetheless, the remaining 206 participants are sufficient to discuss the general tendency of people's behavior [16].

[††]We also provided an open-ended answer format for asking the participants to freely describe other reasons. However, only three participants used it. We therefore think that the eight candidate reasons covered the possible reasons well. Using a fully open-ended answer format to compare results could be an interesting future work.

Table 1 Number of participants who gave a rating of 4 or 5 to at least k reasons.

k	1	2	3	4	5	6	7	8
#participants	196	185	168	125	87	39	22	2
Percentage	95.1	89.8	81.6	60.7	42.2	18.9	10.7	3.88

ferent artists, may help her discover unfamiliar songs. When a user views lyrics for learning (*Language* and *Writing*), she may want to use functions that improve the efficiency of the learning process. Examples for *Language* include enabling the user to see the meaning of a word in lyrics just by tapping the word and recommending a song by the same artist with more intelligible lyrics [58]. Examples for *Writing* include explaining poetic and rhetorical techniques used in writing lyrics and recommending songs with the same techniques.

Finally, Table 1 lists the number of participants who gave a rating of 4 or 5 to at least k reasons. Because 89.8% of the participants gave high scores for more than one reason and over 60% of them often view lyrics for more than three reasons, we can say that the reason for viewing lyrics is not exclusive; rather, it is common to have multiple reasons. An interesting future work would be to predict and recommend lyrics-related functions (like those described above) to use next according to those already used.

3.2.2 Behavior

We now investigate users' detailed behavior in viewing lyrics for different reasons in terms of three aspects. Note that, for each reason, we asked follow-up questions to participants who gave a rating of 4 or 5 so that we could interpret the characteristics of the reasons more accurately (see Fig. 1 for the number of such participants for each reason).

Aspect 1: timing. First, to each participant, we showed a reason for which she gave a rating of 4 or 5 and asked, "When you view lyrics for this reason, do you decide to do so (a) before playing a song or (b) after playing a song?" The possible answers were (1) mostly (a), (2) moderately (a), (3) about the same, (4) moderately (b), and (5) mostly (b). Answers (1) and (2) ((4) and (5)) were then merged into a "Before" ("After") group. The "Timing" column of Table 2 lists the frequency of responses in each group for each reason. For *Structure* that has a statistically high frequency in the "After" group, it would be effective to enable users to more quickly execute the corresponding function proposed in Sect. 3.2.1 while listening to a song, as compared to the functions for other reasons. On the other hand, *Karaoke* has a statistically high frequency in the "Before" group. Therefore, if a smartphone application provided an option to play a song in the setting of the *Karaoke* function explained in Sect. 3.2.1, users would be expected to use the application more frequently to practice for karaoke. In Table 2, although both *Singing* and *Karaoke* are related to singing a song, it is interesting that *Singing* has almost the same frequencies in the "Before" and "After" groups.

Aspect 2: repetition. Our next question was "When you view lyrics for this reason, how many times do you continuously view them while repeatedly playing a song?" The answers consisted of (1) mostly once (*i.e.*, no repetition), (2) mostly two or three times, and (3) mostly more than three times. Because no significant difference was observed between answers (2) and (3), we report the results with answers (2) and (3) merged as a "Many" group, while answer (1) is labeled as "Once." The "Repetition" column of Table 2 lists the results. It can be observed that, for all reasons, the "Many" group has higher frequency. It is thus common behavior to continuously view lyrics while repeating a song. Therefore, it would be helpful for users to change the displayed information according to the number of repetitions (*e.g.*, when a user listens to a song for the *Understanding* reason, different interpretations of the lyrics can be shown every time she plays it).

Table 2 Behavior frequency in terms of three aspects: timing, repetition, and percentage. $*$ ($**$) denotes the statistical difference at $p < 0.05$ ($p < 0.01$) based on a two-tailed z-test.

	Timing		Repetition		Percentage	
Reason	Before	After	Once	Many	Partial	Most
Confirmation	49	**95**$**$	70	**111**$**$	53	**84**$**$
Understanding	60	**70**	38	**121**$**$	20	**116**$**$
Singing	50	**51**	36	**92**$**$	16	**85**$**$
Structure	29	**46**$*$	33	**70**$**$	18	**57**$**$
Karaoke	**55**$*$	33	14	**91**$**$	13	**78**$**$
Boredom	12	**39**$**$	29	**40**	19	**31**
Language	**27**	18	12	**45**$**$	2	**39**$**$
Writing	**11**	10	3	**25**$**$	2	**17**$**$

Aspect 3: percentage. In our last question, we asked, "When you view the lyrics for this reason, what percentage of the lyrics do you view?" The answers were (1) ≤20%, (2) 21%–40%, (3) 41%–60%, (4) 61%–80%, and (5) ≥81%. We merged answers (1) and (2) ((4) and (5)) into a "Partial" ("Most") group. The "Percentage" column of Table 2 lists the frequency of responses in each group. Because "Most" was more popular for all reasons, people tend to view most of the lyrics in any situation. However, a significant difference between "Partial" and "Most" was not observed for *Boredom* only. This result indicates that when a user stops viewing lyrics within a short time, she is likely bored. Therefore, music streaming platforms have a big opportunity to give such users valuable information, as illustrated in Sect. 3.2.1.

4. How People View Lyrics

In this section, we report how people view lyrics based on over 23 million lyrics request logs sent from smartphone applications for playing music.

4.1 Dataset

4.1.1 Lyrics Viewing Log

For lyrics viewing, we used log data given by a lyrics distribution company (SyncPower Corporation) in Japan. Although this company provides lyrics text to various music-

listening smartphone applications, we focused on the iOS application of a Japanese online music service and used logs collected from it. In the application, a user can view the lyrics of a played song while listening to the song. The application gets the lyrics by using an API provided by the lyrics distribution company. The company stores request logs that include the timestamp, user ID, and song ID. Note that the application does not automatically get lyrics when a song is played; rather, it only gets them when a user explicitly requests them. Therefore, the logs are suitable for analyzing how users view lyrics.

We first collected logs whose timestamp was between 1/1/2018 and 12/31/2018. We then removed logs whose duration was less than 30 seconds, because such short-term logs may have resulted from users' wrong operations. Finally, our dataset (hereafter, *LyLog*) consisted of 611,895 users, 214,434 unique songs, and 23,034,417 logs.

4.1.2 Music Listening Log

To investigate the difference between lyrics viewing behavior and music listening behavior, we used the Last.fm dataset released by Schedl [24]. This dataset consists of users' play logs, each of which includes the timestamp, user ID, song ID, and artist ID. To align the users' nationality with the *LyLog* dataset, we first extracted logs of Japanese users (the dataset also includes each user's nationality). We then collected logs whose timestamp was between 1/1/2013 and 12/31/2013 and removed logs whose duration was less than 30 seconds. This gave us a music listening dataset (hereafter, *Lastfm*) consisting of 660 users[†], 718,466 unique songs, and 2,932,430 logs[††].

We do acknowledge some limitations of using *Lastfm* for comparison. For example, the years in *Lastfm* are different from those in *LyLog*, and *Lastfm* includes play histories from not only smartphones but also PCs. Therefore, it should be noted that the purpose of the comparison in this paper is not to provide generalizable insights about the differences between lyrics viewing and music listening. Nonetheless, we think it is still worth comparing the differences as a first step toward understanding the characteristics of lyrics viewing behavior. We leave it as a future work to compare lyrics viewing logs and music listening logs from the same platform.

4.2 Basic Statistics

We first investigated several basic characteristics of lyrics viewing. Figure 2 shows the distribution of the number of consumption logs per song[†††]. Although the curves of both

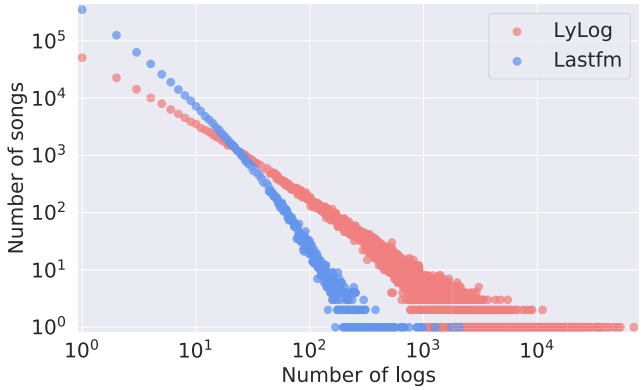

Fig. 2 Distribution of the number of logs per song. There are y songs that have x logs.

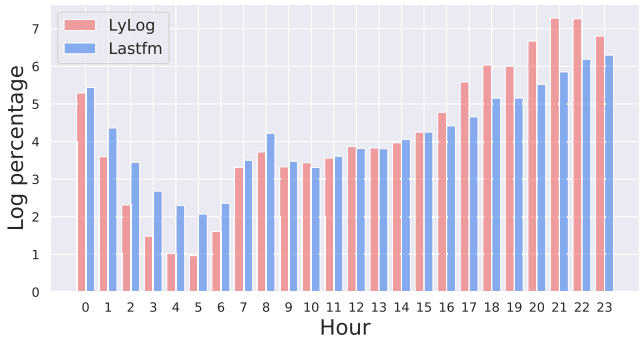

Fig. 3 Distribution of logs over the hours of the day.

LyLog and *Lastfm* show the heavy tail of their consumption patterns, lyrics viewing behavior is more biased to popular songs: in *Lastfm*, 80% of the whole logs are dominated by the top 34.8% of the songs in terms of the number of logs, while in *LyLog*, those are dominated by only the top 6.64% of the songs.

In Fig. 3, we show the distribution of logs over the hours of the day. According to a survey on time use by the Statistics Bureau of Japan [62], the average Japanese person gets up at 6:32 am, commutes to school or work between 7:30 am and 8:30 am, commutes from school or work between 6:00 pm and 7:00 pm, and goes to sleep at 11:15 pm. Referring to this time schedule, we can see some common characteristics in both datasets: the number of logs increases during the morning commute and after returning home in the evening; then, the number gradually decreases as people go to sleep. Between 5:00 pm and 11:59 pm, however, *LyLog* has a higher percentage than *Lastfm* does. Viewing lyrics on a smartphone requires users to interact with the application more actively, as in tapping the screen to request and look at lyrics; in contrast, users can listen to songs even with a smartphone in a pocket. Therefore, we can guess that users often view lyrics in a relaxed state after coming back home. When a smartphone application recommends some of the functions described in Sect. 3.2.1 to a user, night would be a more suitable time, because the user would engage more actively in viewing lyrics than during the daytime: it would be

[†]There is no correspondence between the users in *LyLog* and those in *Lastfm*.

[††]A similar Last.fm dataset was released more recently [25], but the included logs are older than those in Schedl's dataset [24]. Therefore, we decided to use the latter dataset.

[†††]Throughout our investigation, the word "consumption" refers to viewing lyrics in *LyLog* or listening to music in *Lastfm*.

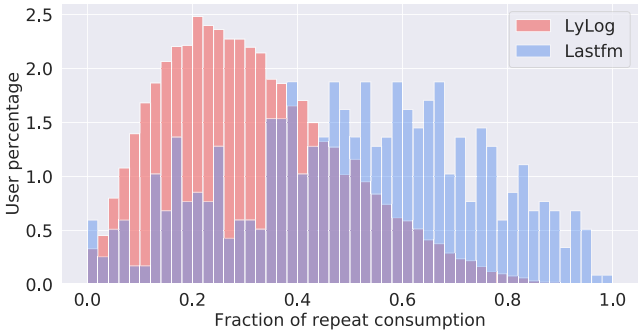

Fig. 4 Fraction of repeat consumption for each user.

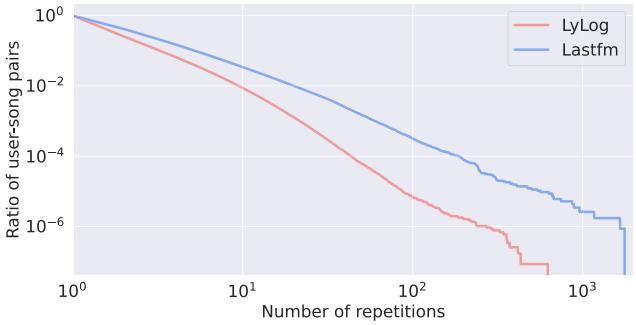

Fig. 5 Distribution of the ratio of user-song pairs that are repeatedly consumed x or more times.

an interesting future work to verify the usefulness of changing the recommendation frequency of each function according to time. Regarding the distribution of logs over the days of week, people view lyrics and listen to music 6.64% and 6.53% more often on weekends than on weekdays, respectively; and no significant difference was observed between the datasets.

4.3 Repeat Consumption

We next investigated repeat consumption behavior in which a user consumes the same song repeatedly over time. We first computed the fraction of repeat consumption for each user. For example, if a user's fraction is 0.4, then 40% of viewed lyrics have been already viewed by her. Figure 4 shows this fraction's distribution. It can be observed that the fraction for *LyLog* tends to be lower than that for *Lastfm*; in fact, the average fractions for *LyLog* and *Lastfm* are 0.378 and 0.604, respectively. However, we can say that the fraction of repeat consumption for lyrics viewing is still high compared to that of other domains such as watching videos (fraction: 0.26) and clicking on English Wikipedia pages (fraction: 0.15) [30]. The above analysis did not consider how many times each song was repeatedly consumed. Thus, we also computed the ratio of user-song pairs, in which each song was repeatedly consumed x or more times by each user, to all user-song pairs, as shown in Fig. 5. We can see that the numbers of repetitions for both *LyLog* and *Lastfm* have a heavy tail. However, because the *LyLog* curve is located below and to the left of the *Lastfm* curve, people do not repeatedly view the same lyrics as many times as they listen to the same song.

Benson et al. [30] reported that, in repeat consumption, each item has its own lifetime for a user, as described in Sect. 2.1. Following their processes, we investigated the lifetime characteristics of lyrics viewing as follows. Given a user, we first sorted all songs for which she requested lyrics in ascending order of the timestamp. We then extracted songs whose first and last consumption events were in the middle 60% of the list, so that we could consider songs that certainly began and ended their lifetimes during the period of data collection. Suppose that a user's extracted consumption list consists of N songs and is represented by

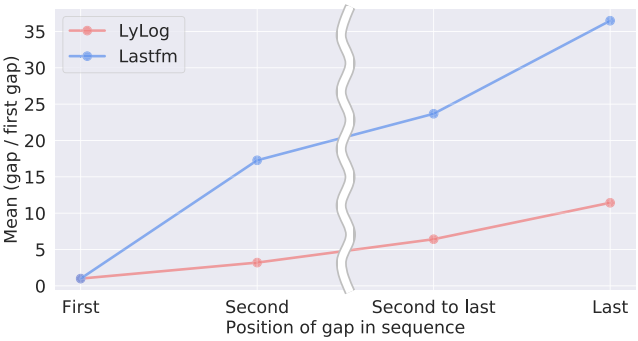

Fig. 6 Normalized mean index gaps.

$L = \{i_1, \ldots, i_N\}$. When a particular song s is consumed k times at indices $\{i^s_1, \ldots, i^s_k\} \in L$, the index gap between the jth and j+1th consumption events is defined by $g_j = i^s_{j+1} - i^s_j$. Figure 6 shows the transition of the mean gap, with all gaps normalized by the first gap g_1 (the average values of g_1 for *LyLog* and *Lastfm* were 19.0 and 248, respectively). As in the report by Benson et al. [30], in lyrics viewing behavior, too, the gap tends to grow over time. This means that when a user repeatedly views the lyrics of a song, she views it again within a short span at the beginning; the span gradually increases as she gets bored with it, and eventually she stops viewing the lyrics. As can be seen in Fig. 6, the gap increase rate for *LyLog* was smaller than that for *Lastfm*.

Because the gap grows over time, there is a possibility that we can detect a user who begins to get bored with particular lyrics by using the method proposed by Benson et al. [30]. When such a user is detected, suggesting functions (from those described in Sect. 3.2.1) that she has not used for the lyrics is one possible way to hold her attention on the lyrics for a longer time. In contrast, recommending novel lyrics related in terms of, say, the topic [10]–[13] would be a good trigger for the user to listen to new songs and expand her interest to other artists; this would also be beneficial for music streaming platforms.

5. Conclusion

In this paper, we investigated why and how people view lyrics while listening to music on a smartphone. Regard-

ing the "why" part, we conducted an online user survey involving 206 participants; regarding the "how" part, we analyzed over 23 million lyrics request logs. From the results, we discussed reusable insights that are beneficial for researchers and music streaming platforms, such as the extent of the demand for the eight major reasons to view lyrics and the generality of repeatedly viewing the same lyrics. We also suggested several functions according to users' reasons for viewing lyrics. We believe that realizing the functions would diversify and enrich users' music listening experiences. Some of the reported findings might be obvious (*e.g.*, people view lyrics more often at night). However, in this kind of study that investigates research questions on an unexplored topic, it is valuable to report not only unexpected results but also such obvious results based on the data; obvious but verified results can then be used as evidence for claiming the appropriateness of proposed methods or systems in later studies.

We acknowledge a limitation of this paper in that we investigated lyrics viewing behavior by only Japanese people in both the "why" and "how" parts. Nonetheless, we believe that our study is a worthwhile contribution because this is the first attempt to reveal lyrics viewing behavior and verifies the fundamental characteristics of the behavior. At the same time, this limitation indicates the possibilities of this research topic and guides future work such as investigating the differences in lyrics viewing behavior among countries. It would also be an important future work to investigate lyrics viewing behavior on other devices (*e.g.*, PCs and tablets) and at various locations (*e.g.*, homes, restaurants, and public transportation).

Acknowledgments

The authors would like to extend their appreciation to SyncPower Corporation for providing the lyrics request logs. This work was supported in part by JST CREST Grant Number JPMJCR20D4 and JSPS KAKENHI Grant Number JP20K19934.

References

[1] K. Tsukuda, M. Hamasaki, and M. Goto, "Toward an understanding of lyrics-viewing behavior while listening to music on a smartphone," Proc. 22nd International Society for Music Information Retrieval Conference, ISMIR 2021, pp.705–713, 2021.

[2] J.H. Lee and J.S. Downie, "Survey of music information needs, uses, and seeking behaviours: Preliminary findings," Proc. 5th International Conference on Music Information Retrieval, ISMIR 2004, pp.989–992, 2004.

[3] J.H. Lee, J.S. Downie, and S.J. Cunningham, "Challenges in cross-cultural/multilingual music information seeking," Proc. 6th International Conference on Music Information Retrieval, ISMIR 2005, pp.1–7, 2005.

[4] D. Bainbridge, S.J. Cunningham, and J.S. Downie, "How people describe their music information needs: A grounded theory analysis of music queries," Proc. 4th International Conference on Music Information Retrieval, ISMIR 2003, pp.221–222, 2003.

[5] A. Demetriou, A. Jansson, A. Kumar, and R.M. Bittner, "Vocals in music matter: The relevance of vocals in the minds of listeners," Proc. 19th International Society for Music Information Retrieval Conference, ISMIR 2018, pp.514–520, 2018.

[6] E. Brochu and N. de Freitas, "'Name that song!": A probabilistic approach to querying on music and text," Proc. 15th International Conference on Neural Information Processing Systems, NIPS 2002, pp.1505–1512, 2002.

[7] M. Müller, F. Kurth, D. Damm, C. Fremerey, and M. Clausen, "Lyrics-based audio retrieval and multimodal navigation in music collections," Proc. 11th European Conference on Digital Libraries, ECDL 2007, pp.112–123, 2007.

[8] T. Hosoya, M. Suzuki, A. Ito, and S. Makino, "Lyrics recognition from a singing voice based on finite state automaton for music information retrieval," Proc. 6th International Conference on Music Information Retrieval, ISMIR 2005, pp.532–535, 2005.

[9] C. Wang, J.R. Jang, and W. Wang, "An improved query by singing/humming system using melody and lyrics information," Proc. 11th International Society for Music Information Retrieval Conference, ISMIR 2010, pp.45–50, 2010.

[10] S. Sasaki, K. Yoshii, T. Nakano, M. Goto, and S. Morishima, "LyricsRadar: A lyrics retrieval system based on latent topics of lyrics," Proc. 15th International Society for Music Information Retrieval Conference, ISMIR 2014, pp.585–590, 2014.

[11] T. Nakano and M. Goto, "LyricListPlayer: A consecutive-query-by-playback interface for retrieving similar word sequences from different song lyrics," Proc. Sound and Music Computing Conference 2016, SMC 2016, pp.344–349, 2016.

[12] K. Tsukuda, K. Ishida, and M. Goto, "Lyric Jumper: A lyrics-based music exploratory web service by modeling lyrics generative process," Proc. 18th International Society for Music Information Retrieval Conference, ISMIR 2017, pp.544–551, 2017.

[13] K. Watanabe and M. Goto, "Query-by-Blending: A music exploration system blending latent vector representations of lyric word, song audio, and artist," Proc. 20th International Society for Music Information Retrieval Conference, ISMIR 2019, pp.144–151, 2019.

[14] D. Baur, J. Büttgen, and A. Butz, "Listening factors: A large-scale principal components analysis of long-term music listening histories," Proc. SIGCHI Conference on Human Factors in Computing Systems, CHI 2012, pp.1273–1276, May 2012.

[15] J.H. Lee and R. Price, "Understanding users of commercial music services through personas: Design implications," Proc. 16th International Society for Music Information Retrieval Conference, ISMIR 2015, pp.476–482, 2015.

[16] J.H. Lee and N.M. Waterman, "Understanding user requirements for music information services," Proc. 13th International Society for Music Information Retrieval Conference, ISMIR 2012, pp.253–258, 2012.

[17] A.J. Lonsdale and A.C. North, "Why do we listen to music? A uses and gratifications analysis," British Journal of Psychology, vol.102, no.1, pp.108–134, 2011.

[18] W.M. Randall and N.S. Rickard, "Reasons for personal music listening: A mobile experience sampling study of emotional outcomes," Psychology of Music, vol.45, no.4, pp.479–495, 2017.

[19] A.B. Haake, "Individual music listening in workplace settings: An exploratory survey of offices in the UK," Musicae Scientiae, vol.15, no.1, pp.107–129, 2011.

[20] J.H. Lee, L. Pritchard, and C. Hubbles, "Can we listen to it together?: Factors influencing reception of music recommendations and post-recommendation behavior," Proc. 20th International Society for Music Information Retrieval Conference, ISMIR 2019, pp.663–669, 2019.

[21] A.C. North, D.J. Hargreaves, and J.J. Hargreaves, "Uses of music in everyday life," Music Perception: An Interdisciplinary Journal, vol.22, no.1, pp.41–77, 2004.

[22] L.A. Liikkanen and A. Salovaara, "Music on YouTube: User engagement with traditional, user-appropriated and derivative videos," Computers in Human Behavior, vol.50, pp.108–124, Sept. 2015.

[23] J.H. Lee, R. Wishkoski, L. Aase, P. Meas, and C. Hubbles, "Un-

derstanding users of cloud music services: Selection factors, management and access behavior, and perceptions," Journal of the Association for Information Science and Technology, vol.68, no.5, pp.1186–1200, May 2017.
[24] M. Schedl, "The LFM-1b dataset for music retrieval and recommendation," Proc. 2016 ACM International Conference on Multimedia Retrieval, ICMR 2016, pp.103–110, June 2016.
[25] G. Vigliensoni and I. Fujinaga, "The music listening histories dataset," Proc. 18th International Society for Music Information Retrieval Conference, ISMIR 2017, pp.96–102, 2017.
[26] M. Schedl, "Leveraging microblogs for spatiotemporal music information retrieval," Proc. 35th European Conference on Advances in Information Retrieval, ECIR 2013, pp.796–799, 2013.
[27] D. Hauger, M. Schedl, A. Kosir, and M. Tkalcic, "The million musical tweet dataset - What we can learn from microblogs," Proc. 14th International Society for Music Information Retrieval Conference, ISMIR 2013, pp.189–194, 2013.
[28] E. Zangerle, M. Pichl, W. Gassler, and G. Specht, "#nowplaying music dataset: Extracting listening behavior from twitter," Proc. First International Workshop on Internet-Scale Multimedia Management, WISMM 2014, pp.21–26, Nov. 2014.
[29] A. Anderson, R. Kumar, A. Tomkins, and S. Vassilvitskii, "The dynamics of repeat consumption," Proc. 23rd International Conference on World Wide Web, WWW 2014, pp.419–430, April 2014.
[30] A.R. Benson, R. Kumar, and A. Tomkins, "Modeling user consumption sequences," Proc. 25th International Conference on World Wide Web, WWW 2016, pp.519–529, April 2016.
[31] G. Dzhambazov, A. Srinivasamurthy, S. Sentürk, and X. Serra, "On the use of note onsets for improved lyrics-to-audio alignment in Turkish makam music," Proc. 17th International Society for Music Information Retrieval Conference, ISMIR 2016, pp.716–722, 2016.
[32] K. Lee and M. Cremer, "Segmentation-based lyrics-audio alignment using dynamic programming," Proc. 9th International Conference on Music Information Retrieval, ISMIR 2008, pp.395–400, 2008.
[33] V. Thomas, C. Fremerey, D. Damm, and M. Clausen, "Slave: A score-lyrics-audio-video-explorer," Proc. 10th International Society for Music Information Retrieval Conference, ISMIR 2009, pp.717–722, 2009.
[34] H. Fujihara, M. Goto, J. Ogata, K. Komatani, T. Ogata, and H.G. Okuno, "Automatic synchronization between lyrics and music CD recordings based on viterbi alignment of segregated vocal signals," Proc. 8th IEEE International Symposium on Multimedia, ISM 2006, pp.257–264, 2006.
[35] M.Y. Kan, Y. Wang, D. Iskandar, T.L. Nwe, and A. Shenoy, "LyricAlly: Automatic synchronization of textual lyrics to acoustic music signals," IEEE Transactions on Audio, Speech, and Language Processing, vol.16, no.2, pp.338–349, Feb. 2008.
[36] H. Fujihara, M. Goto, J. Ogata, and H.G. Okuno, "LyricSynchronizer: Automatic synchronization system between musical audio signals and lyrics," IEEE Journal of Selected Topics in Signal Processing, vol.5, no.6, pp.1252–1261, Oct. 2011.
[37] M. Mauch, H. Fujihara, and M. Goto, "Integrating additional chord information into HMM-based lyrics-to-audio alignment," IEEE Transactions on Audio, Speech, and Language Processing, vol.20, no.1, pp.200–210, Jan. 2012.
[38] R.J. Ellis, Z. Xing, J. Fang, and Y. Wang, "Quantifying lexical novelty in song lyrics," Proc. 16th International Society for Music Information Retrieval Conference, ISMIR 2015, pp.694–700, 2015.
[39] H. Hirjee and D.G. Brown, "Automatic detection of internal and imperfect rhymes in rap lyrics," Proc. 10th International Society for Music Information Retrieval Conference, ISMIR 2009, pp.711–716, 2009.
[40] X. Hu and B. Yu, "Exploring the relationship between mood and creativity in rock lyrics," Proc. 12th International Society for Music Information Retrieval Conference, ISMIR 2011, pp.789–794, 2011.
[41] E. Nichols, D. Morris, S. Basu, and C. Raphael, "Relationships between lyrics and melody in popular music," Proc. 10th International Society for Music Information Retrieval Conference, ISMIR 2009, pp.471–476, 2009.
[42] A. Singhi and D.G. Brown, "Are poetry and lyrics all that different?," Proc. 15th International Society for Music Information Retrieval Conference, ISMIR 2014, pp.471–476, 2014.
[43] C. Johnson-Roberson and M. Johnson-Roberson, "Temporal and regional variation in rap lyrics," NIPS Workshop on Topic Models: Computation, Application and Evaluation, NIPSW 2013, pp.1–5, 2013.
[44] P. Knees, M. Schedl, and G. Widmer, "Multiple lyrics alignment: Automatic retrieval of song lyrics," Proc. 6th International Conference on Music Information Retrieval, ISMIR 2005, pp.564–569, 2005.
[45] G. Geleijnse and J.H.M. Korst, "Efficient lyrics extraction from the web," Proc. 7th International Conference on Music Information Retrieval, ISMIR 2006, pp.371–372, 2006.
[46] R. Macrae and S. Dixon, "Ranking lyrics for online search," Proc. 13th International Society for Music Information Retrieval Conference, ISMIR 2012, pp.361–366, 2012.
[47] X. Hu and J.S. Downie, "When lyrics outperform audio for music mood classification: A feature analysis," Proc. 11th International Society for Music Information Retrieval Conference, ISMIR 2010, pp.619–624, 2010.
[48] R. Mayer, R. Neumayer, and A. Rauber, "Rhyme and style features for musical genre classification by song lyrics," Proc. 9th International Conference on Music Information Retrieval, ISMIR 2008, pp.337–342, 2008.
[49] R. Mayer and A. Rauber, "Music genre classification by ensembles of audio and lyrics features," Proc. 12th International Society for Music Information Retrieval Conference, ISMIR 2011, pp.675–680, 2011.
[50] X. Wang, X. Chen, D. Yang, and Y. Wu, "Music emotion classification of Chinese songs based on lyrics using TF*IDF and rhyme," Proc. 12th International Society for Music Information Retrieval Conference, ISMIR 2011, pp.765–770, 2011.
[51] B. Wei, C. Zhang, and M. Ogihara, "Keyword generation for lyrics," Proc. 8th International Conference on Music Information Retrieval, ISMIR 2007, pp.121–122, 2007.
[52] M. van Zaanen and P. Kanters, "Automatic mood classification using TF*IDF based on lyrics," Proc. 11th International Society for Music Information Retrieval Conference, ISMIR 2010, pp.75–80, 2010.
[53] H. Fujihara, M. Goto, and J. Ogata, "Hyperlinking lyrics: A method for creating hyperlinks between phrases in song lyrics," Proc. 9th International Conference on Music Information Retrieval, ISMIR 2008, pp.281–286, 2008.
[54] D. Baur, B. Steinmayr, and A. Butz, "SongWords: Exploring music collections through lyrics," Proc. 11th International Society for Music Information Retrieval Conference, ISMIR 2010, pp.531–536, 2010.
[55] W. Machida and T. Itoh, "Lyricon: A visual music selection interface featuring multiple icons," Proc. 15th International Conference on Information Visualisation, IV 2011, pp.145–150, 2011.
[56] S. Funasawa, H. Ishizaki, K. Hoashi, Y. Takishima, and J. Katto, "Automated music slideshow generation using web images based on lyrics," Proc. 11th International Society for Music Information Retrieval Conference, ISMIR 2010, pp.63–68, 2010.
[57] T. O'Hara, N. Schüler, Y. Lu, and D. Tamir, "Inferring chord sequence meanings via lyrics: Process and evaluation," Proc. 13th International Society for Music Information Retrieval Conference, ISMIR 2012, pp.463–468, 2012.
[58] K.M. Ibrahim, D. Grunberg, K. Agres, C. Gupta, and Y. Wang, "Intelligibility of sung lyrics: A pilot study," Proc. 18th International Society for Music Information Retrieval Conference, ISMIR 2017, pp.686–693, 2017.
[59] T. Nakano, M. Goto, and Y. Hiraga, "An automatic singing skill evaluation method for unknown melodies using pitch interval accuracy and vibrato features," Proc. 9th International Conference on Spoken

Language Processing, INTERSPEECH 2006, pp.1706–1709, 2006.
[60] 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, Sept. 2006.
[61] K. Watanabe and M. Goto, "A chorus-section detection method for lyrics text," Proc. 21st International Society for Music Information Retrieval Conference, ISMIR 2020, pp.351–359, 2020.
[62] Statistics Bureau of Japan, Survey on time use and leisure activities, Japan Statistical Association, Tokyo, 2016.

Kosetsu Tsukuda received the Ph.D. degree in Informatics from Kyoto University, Japan in 2014. He is currently a Senior Researcher at the National Institute of Advanced Industrial Science and Technology (AIST), Japan. His research interests lie in the areas of recommender systems, user generated content, and user behabior analysis. He has received 16 awards including IPSJ Computer Science Research Award for Young Scientists and IPSJ Yamashita SIG Research Award.

Masahiro Hamasaki received the Ph.D. degree in informatics from the Graduate University for Advanced Studies (SOKENDAI), Japan in 2005. He is currently the Leader of the Media Interaction Group at the National Institute of Advance Industrial Science and Technology (AIST), Japan. His research interests include Web mining, semantic Web, and social media analysis. He is a member of the JSAI, the IPSJ, and ACM.

Masataka Goto received the Doctor of Engineering degree from Waseda University in 1998. He is currently a Prime Senior Researcher at the National Institute of Advanced Industrial Science and Technology (AIST), Japan. Over the past 30 years he has published more than 300 papers in refereed journals and international conferences and has received 58 awards, including several best paper awards, best presentation awards, the Tenth Japan Academy Medal, and the Tenth JSPS PRIZE. He has served as a committee member of over 120 scientific societies and conferences, including the General Chair of ISMIR 2009 and 2014. As the research director, he began OngaACCEL project in 2016 and RecMus project in 2021, which are five-year JST-funded research projects (ACCEL and CREST) related to music technologies.