

A Computer-Mediated Discourse Analysis of User Commenting Behavior on an Online Music Distribution Site

Hiromi Ishizaki† Susan C. Herring‡ Gen Hattori† Chihiro Ono† Yasuhiro Takishima†

Abstract

We analyze user behavior on SoundCloud.com, a web-based music distribution site, using the methodology of computer-mediated discourse analysis. The broad goal is to infer patterns of use that can inform the development of methods to facilitate communication among online users. In order to achieve this goal, we analyze SoundCloud comments and identify site design features that facilitate users' actions, using speech act analysis and dynamic topic analysis. We address two research questions: What are the most common communicative acts on the music distribution site, and which commenting function of the site contributes most to facilitating interactive commenting? Results of speech act analysis show that 56% of comments are *REACT* which means site users mostly react to the song in short, simple comments. Additionally the users tend to be more interactive when using the timed comment function, which appear below the song waveform, provided by the site. Dynamic topic analysis reveals that the structure of regular comments, made on the song as a whole, shows prompt focused such as blog comment, and structure of timed comments shows sequential threads are more interactive as is the case in chatting. These findings indicate that comments which focus on particular part of the song have possibility to facilitate user's interactive comments.

1. Introduction

The rapid increase in digital devices and the development of web-based music streaming sites such as Last.fm and Pandora have made it possible for many users to listen to large numbers of songs via the Internet. At the same time, increases of the popularity of video and music distribution sites have been accompanied by a change in user activities, from just listening to creating and sharing their own audio recordings. SoundCloud (SoundCloud.com <http://soundcloud.com/>) is a popular music distribution site where users can upload their own audio recordings. Users can also search and download audio recordings based on their musical preferences. These activities have many advantages: Users who create audio recordings can share them easily and collect listener feedback immediately, and they can also search for new music that is not available commercially. Furthermore, they can download the recordings they like and leave comments to express their views of, and potentially improve, the recordings. It follows that facilitation of actions such as user commenting is desirable for users and developers of multimedia content distribution sites alike.

In order to facilitate user actions, it is important to first understand user behavior on a given site. In addition to its value for scholarship, understanding behavioral tendencies on music distribution sites facilitates the development of appropriate system designs. In the field of computer-mediated communication (CMC), there have been many research efforts to analyze user behavior on communication services such as email, chat, and social network sites (SNS) [1][2][3]. On these communication services, the main datasets for analysis are log

data such as user records from mobile devices and computers, and text data such as comments and messages. Because communication among users on these services is primarily via (typed) discourse, computer-mediated discourse analysis (CMDA) [4], an approach grounded in linguistics, was developed to analyze online behavior in such communication services. Additionally, approaches from social science have been applied to understand user behavior in multimedia systems. For example, ethnography has been applied to observe user behavior and analyze the features of systems, in order to find key features and make recommendations for multimedia system design [5][6][7]. These approaches attempt to analyze user behavior from different perspectives.

Ethnographic approaches enable rich qualitative analysis of user behavior. However, on music distribution sites such as SoundCloud and Pandora (and on websites more generally), it is difficult to observe the behaviors of users directly, since the users are widely dispersed geographically and in many cases anonymous or pseudonymous. Although it might be possible to obtain and analyze logs of user actions, such as what parts of the site they clicked on, from the sites, it is difficult to extract meaningful interpretations from such data. In contrast, the comments that users leave on such sites are in themselves an important form of behavior and can reveal much about users' thoughts, impressions, emotions, and intentions. Therefore, in this study we apply computer-mediated discourse analysis to analyze user behavior on SoundCloud. In particular, we apply speech act analysis and dynamic topic analysis to analyze SoundCloud comment data from different discourse perspectives, one semantic and the other interactional. These methods enable both qualitative and quantitative analysis of comment data in online environments where it is otherwise difficult to observe user behavior directly.

2. Research Questions

Besides listening, commenting is the basic activity that users engage in on music distribution sites. Therefore, we analyze comments on SoundCloud in order to extract tendencies in user behavior, controlling for the placement of comments in the interface. In addition, we attempt to identify which placement option most facilitates user commenting, based on the results of the first analysis. Accordingly, the research questions addressed in this study are:

RQ1: What are the most common communicative acts that users engage in on the music distribution site?

RQ2: Which feature of the site design as regards comment placement contributes most to facilitating interactive user commenting?

†KDDI R&D Laboratories, Inc

‡Indiana University

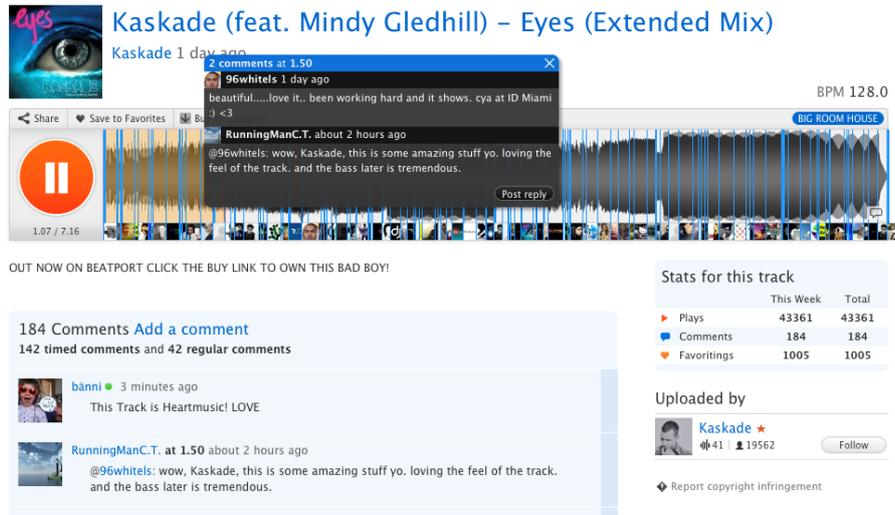


Figure 1. The SoundCloud interface, with two “timed comments” expanded in the waveform. “Regular comments” appear on the left below the waveform.

Table 1. Acts in the CMC act taxonomy [11]

Acts
ACCEPT, APOLOGIZE, CLAIM, DESIRE, DIRECT, ELABORATE, GREET, INFORM, INQUIRE, INVITE, MANAGE, REACT, REJECT, REPAIR, REQUEST, THANK

As regards RQ1, we could not find any studies in the literature of commenting behavior on music distribution sites. Hence, basic research is needed in this area. RQ2 is also important in order to understand and facilitate user interaction. In what follows, we describe the methods we employed, the data, and our findings, organized into two separate sections for the speech act and dynamic topic analyses.

3. Analysis of User Behavior Based on Speech Act Analysis

In this section, we analyze user comments on SoundCloud based on speech act analysis. A speech act is defined in the field of linguistics as the use of language to perform some act, such as claiming or thanking [8][9], and it is the unit at the lowest rank of the discourse level of language patterning. In computer-mediated discourse analysis, it is situated at level two, meaning. Analyzing speech acts sheds light on user behavior by revealing the meaning of each utterance in context.

3.1 Sample Data and Coding

Our sample data are comments from SoundCloud, a popular online audio distribution platform that was created in 2007 by a Swedish sound designer and a Swedish musician, initially out of frustration with the limited affordances of music sharing sites such as MySpace [10]. On the site, audio recordings uploaded by users are graphically represented as waveforms, one per page at the top of the page, and users can listen to and comment on the audio recordings. In particular, a feature of SoundCloud that differs from other music streaming sites is that SoundCloud

Table 2. Coding results of speech act analysis using the CMC act taxonomy

Speech Act	Combined (%)	Regular (%)	Timed (%)
ACCEPT	4 (0.2)	0 (0.0)	4 (0.2)
APOLOGIZE	1 (0.0)	0 (0.0)	1 (0.1)
CLAIM	317 (15.0)	81 (16.1)	236 (14.7)
DESIRE	18 (0.9)	3 (0.6)	15 (0.9)
DIRECT	143 (6.8)	48 (9.6)	95 (5.9)
ELABORATE	29 (1.4)	16 (3.2)	13 (0.8)
GREET	59 (2.8)	10 (2.0)	49 (3.0)
INFORM	237 (11.2)	111 (22.1)	126 (7.8)
INQUIRE	23 (1.1)	6 (1.2)	17 (1.1)
INVITE	6 (0.3)	1 (0.2)	5 (0.3)
REACT	1178 (55.8)	198 (39.4)	980 (60.9)
REQUEST	45 (2.1)	14 (2.8)	31 (1.9)
THANK	51 (2.4)	14 (2.8)	37 (2.3)
TOTAL	2111 (100.0)	502 (100.0)	1609 (100.0)

allows users to insert a comment at a specific point in time of the track, such as at 5.0 seconds. Using this feature, users can comment wherever they want to, such as on the introduction or the chorus, or at a time when they were especially impressed. Text comments can also be made on the track as a whole; these appear below the waveform, similar to text comments on YouTube videos. Hereafter, we refer to users' general comments that appear below the waveform as "regular comments", and comments inserted at a specific point along the waveform as "timed comments" (Figure 1).

In order to analyze speech acts in the sample data, we first collected 43 music entries from SoundCloud in October 2012. Each entry had over 200 comments and was selected from the genre of "house music". As data for analysis, we then selected five music entries randomly. From these sample entries, we extracted 1295 comments posted by 922 participants. We divided the comments into utterances based on punctuation, resulting in 2173 utterances. Two coders then manually assigned the most

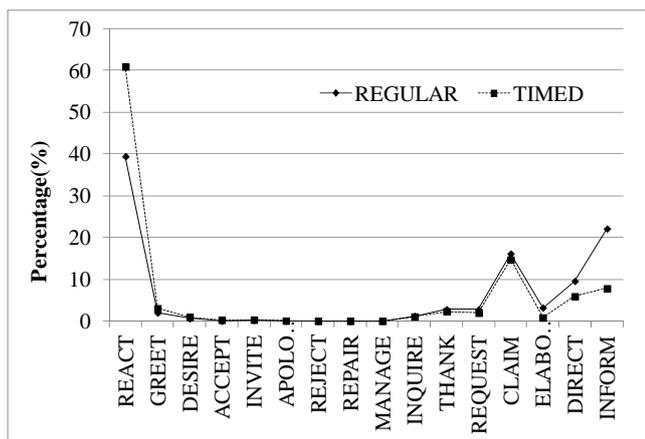


Figure 2. Percentages of speech acts in regular and timed comments

appropriate speech act to each utterance using the CMC act taxonomy [11]. This taxonomy is a classification scheme for coding speech acts designed for computer-mediated discourse that includes 16 acts; these are listed in Table 1. Each coder participated in about 15 hours of training to become familiar with all of the definitions in the codebook and to practice coding actual comments.

Following initial training, the coders coded two music practice entries (including 524 utterances) that were not part of the sample, after which issues and questions concerning the codebook or coding process were discussed. After the two coders independently coded each utterance, we extracted the coded utterances on which both coders agreed. In the end, we used 2111 coded utterances for the speech act analysis. Additionally, we broke down the utterances into two subgroups, regular and timed comments, in order to compare them.

3.2 Results

Table 2 shows the frequency of each speech act category in the code sample. We did not observe any instances of REPAIR, REJECT, APOLOGIZE, or MANAGE; therefore, we excluded these acts from the table. According to Table 2, REACT, CLAIM, and INFORM are the most common acts in the sample data. REACT is the most frequent, with a combined frequency of 55.8%. The second highest percentage is 15.2% for CLAIM, followed by 11.2% for INFORM. These results show that users tend to comment in order to express simple reactions to, and impressions of, the music entry. In further support of this claim, the average utterance length in the sample is only 3.4 words, which is shorter than that of email [1] and shorter, even, than much synchronous chat [12].

Overall, we observed many short comments that described concise and simple impressions of the music entry, such as REACT and CLAIM. These comments did not involve conversational exchanges [13]; indeed, there were few interactions among commenters. Rather, these results suggest that the motivation of commenters on SoundCloud is not so much communication as expression of simple reactions to the audio recordings.

For further analysis, we broke the data into two subgroups, regular and timed comments, and compared them. In order to

Table 3. Coding results of topical units (T: On-topic, P: Parallel Shift, B: Break) and average semantic distance of Parallel Shifts

Topical Units	Number of topical units (%)		
	Combined (%)	Regular (%)	Timed (%)
T	77 (24.4)	9 (14.1)	68 (27.0)
P1	191 (60.4)	46 (71.9)	145 (57.5)
P2	24 (7.6)	0 (.0)	24 (9.5)
P3	9 (2.8)	0 (.0)	9 (3.6)
B	15 (4.7)	9 (14.1)	6 (2.4)
Total	316 (100.0)	64 (100.0)	252 (100.0)

clarify the differences, we reordered the speech act results based on their proportions. Figure 2 charts the differences between regular and timed comments. Acts on the left side in this figure were observed more in timed comments, and acts on the right side were observed more in regular comments. These speech acts can be interpreted as being somewhat characteristic of each comment type. Thus INFORM, DIRECT, and ELABORATE tended to be expressed more in regular comments, while REACT tended to be expressed more in timed comments. For regular comments, percentages of INFORM and DIRECT are much higher than those of timed comments. INFORM and DIRECT may be expressed more as regular comments because they tend not to be related to particular parts of the track. We conducted Chi-square analysis in order to test the difference in act distribution between these two subgroups and confirmed that the difference between them is statistically significant at the alpha level ($\alpha=0.01$, $p<0.01$).

Each speech act category was associated with characteristic behaviors. Many INFORM comments included URLs and links to other songs. Other INFORM comments, such as "Blogged at [URL]", are required by "netiquette" in order to share the origins of the music blogged about. The relatively high percentage of INFORM use appears to be related to the purpose of SoundCloud, which includes promotion/distribution of songs and sharing information about them.

For INQUIRE, we observed that users asked for information about the system, as well as personal information, such as "how do u download??" and "Who is [artist name's] ex-girlfriend?". THANK comments usually expressed appreciation for downloadable MP3 files, such as "Thanks for the freeDL =)" and "too sick, thx for the dl." DIRECT comments tended to direct listeners to the commenters' blog sites or their audio recordings on SoundCloud. Typical expressions of DIRECT include "Check out the blog post at [URL]" and "Check out my latest track!". A CLAIM was a longer utterance than REACT, and mainly expressed subjective impressions of the uploaded audio recordings, such as "I love this tune!" and "I FUCKIN LOVE ITXD". In contrast, REACT was typically a single word comment that mainly expressed evaluative reactions, such as "AMAZING!" and "AWESOME", although it could sometimes be longer. A REACT comment did not necessarily include the reason for the commenter's evaluation.

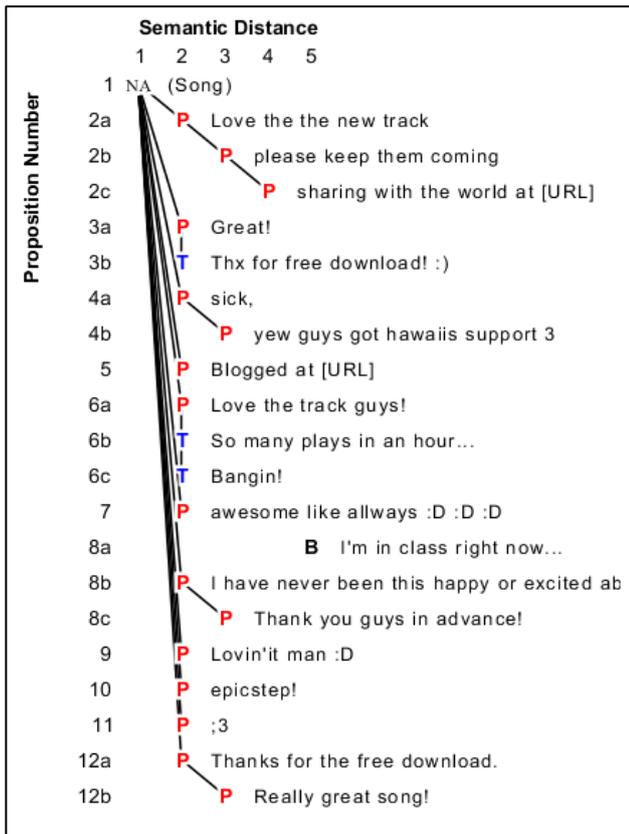


Figure 3. Visualization of dynamic topic transitions (first 20 regular comments)

4. Analysis of Topic Development Based on Dynamic Topic Analysis

In this section, we analyze the development of topics in user comment threads using the dynamic topic analysis (DTA) method [14]. Dynamic topic analysis is a computer-mediated discourse analysis technique for analyzing how discussion/conversation evolves over time by focusing on transitions between topical units. We apply this methodology to analyze topical transitions and what they reveal about how users interact on SoundCloud.

4.1 Coding

To apply dynamic topic analysis, we coded each topical unit for topic relation type (On-topic, Parallel Shift, or Break) and semantic distance (On-topic = 0; Parallel Shift = 1-3; Break = 4). In the sample data, one topical unit corresponded roughly to one proposition. "On-topic" was coded when a proposition expressed reaction to, elaboration on, or continuation of the same topic or provides an expected response to a question. A "Parallel Shift" expressed movement of the conversation onto new ground that was related to what came before. A "Break" expressed a non-sequitur or abrupt topic change, unrelated to what came before. As for the semantic distance of parallel shifts, "1" means that the relation between propositions was immediately obvious, "2" means that the relation was understandable after a moment's thought, and "3" means that the relation was ultimately understandable, but only after puzzling it out [14]. As sample data, we use a music entry randomly selected from sample data sets

mentioned in the speech act analysis section, which include 213 comments, and split propositions to code topical units. After the two coders independently coded each utterance, we extracted the coded utterances on which both coders agreed. In total, we used 320 topical units for the dynamic topic analysis.

4.2 Results

Table 3 shows the percentages of topical units and the average semantic distance of parallel shifts. For this analysis we again divided the sample data into regular comments and timed comments. For the combined results, the percentages of topical move types are 24.4% On-topic, 70.9% Parallel Shifts, and 4.7% Breaks. The percentage of Parallel Shifts is much higher than the others. This result is due in part to the fact that the percentages of REACT and CLAIM are much higher than the other speech acts, as described in the speech act analysis section. Since REACT is basically an evaluative reaction to a music entry, it was coded as a Parallel Shift with a semantic distance of 1. Depending on a coder's interpretation of the data, these REACTS could also plausibly have been coded as "On-Topic" with a semantic distance of 0. However, this would have compromised the legibility of the VisualDTA diagrams, because most of the propositions would have appeared along the same vertical axis as the prompt to which they responded, and it would have been impossible to differentiate visually among them (cf. Figures 3 and 4). These percentages are consistent with the tendency for users to comment in order to react to a music creator and/or to express impressions of the song.

For further analysis we again divided the sample data into regular comments and timed comments. Some differences are visible from Table 3; notably, the timed comments have more On-topic propositions, more Parallel Shifts with a semantic distance of 2 or 3, and fewer Breaks compared to regular comments. To help explain these results, we visualized the patterns of topic transition for each subgroup using the VisualDTA tool [15]. As input data for the visualizations, we extracted the longest timed comment sequences (9 comments, at 0.0 sec) and the first 20 regular comments. Figures 3 and 4 display the output. (Note: In VisualDTA diagrams, the y-axis represents the chronological sequence in which the propositions were posted, and the x-axis represents cumulative semantic distance from the initiating proposition.)

Figure 3 shows that propositions in the first 20 regular comments reply exclusively to the song prompt rather than to propositions by other commenters. While there appear to be some chained sequences of propositions (branching rightward), all of them are within a single comment, meaning that there is no topic development across comments. In contrast, propositions in the timed comments tend to respond to previous propositions, as shown in Figure 4. Propositions are connected to next propositions in timed comments via reactions and (sometimes rather tenuously-connected) topic shifts, several of which are followed by sequences of brief, ratifying (on-topic) responses.

As regards coherence, the sample data overall are generally coherent in that most propositions respond directly or indirectly to a song. However, regular comments show little topic development; they are narrowly prompt focused. In contrast,

timed comment sequences tend to develop ideas interactively.

messages, including email and some chats, and they take part in

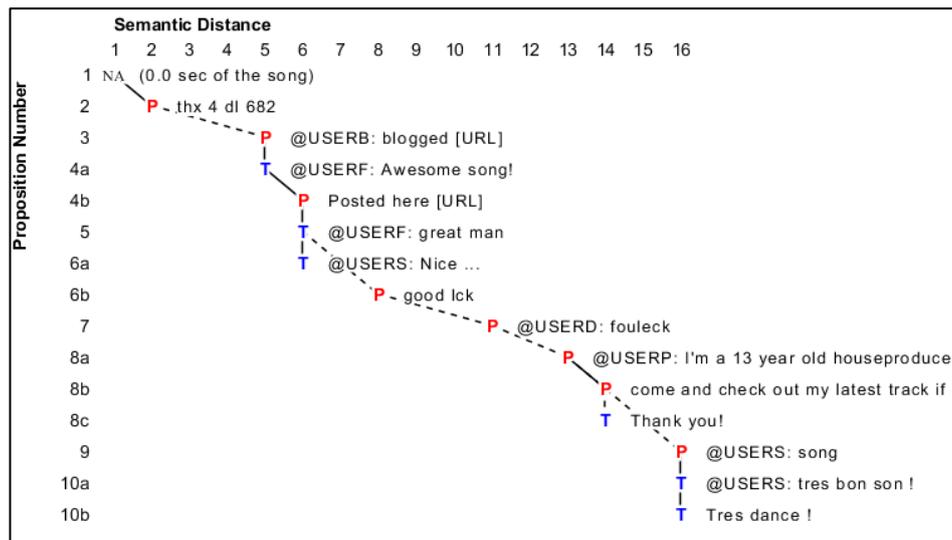


Figure 4. Visualization of dynamic topic transitions (timed comments posted at 0.0 sec of the song)

Figure 4 shows that comment interaction progresses from top left to bottom right in a more or less continuous line. These differences suggests that users choose to make timed or regular comments based on their communicative intention, and that they perceive timed comments to be more interactive. In support of this interpretation, eight out of the nine timed comments at 0.0 sec include the user ID of a direct addressee, such as "@[USER A]". On the use of the @ sign as a marker of direct address in another social media site, Twitter [16].

Additionally, we calculated the consecutiveness of propositions. We defined consecutiveness as the percentage of propositions that respond to previous propositions. Based on the overall dynamic topic analysis results, consecutiveness of timed comments is 24.1% (45/252 topical units) and that of regular comments is only 5.4% (2/64 topical units). This further supports the interpretation that timed comments are more interactive. The regular comment feature is selected for general reactions, whereas selection of the timed comment features is more associated with an intention to engage with other users. At the same time, the interactive nature of timed comments on SoundCloud is weaker than that observed in previous CMC such as discussion forums and chat rooms [14], in that the sequential threads tend to be rather short and digressive.

5. Discussions

In response to RQ1, we found that SoundCloud commenters have a tendency to react simply to the audio recordings and previous comments on the site, rather than users exchanging detailed observations and/or information. This tendency was similar in both commenting locations. The speech act analysis results revealed many REACT and CLAIM acts, both of which are basically subjective evaluations. Furthermore, the SoundCloud comments are shorter than other kinds of CMC

shorter thread sequences. Thus users tend to comment immediately after listening to an audio recording, with short, simple reactions and impressions. Moreover, although politeness was not part of our research questions, we also observed polite behavior in the data that seemed particular to the genre of music sharing website. Users commented that they wrote blogs about the audio recordings, and they freely shared URLs with one another. It is possible that the motivation underlying this behavior was to "give back" to the website community by promoting good music that impressed the commenters.

To address RQ2, we compared the two subgroups as revealed through the results of the speech act analysis and the dynamic topic analysis, especially the latter. The speech act analysis showed that there were more timed comments than regular comments. Dynamic topic analysis showed that roughly one-quarter of timed comments were consecutive, and that consecutiveness was much greater in timed comments than in regular comments, which overwhelmingly responded to the initial song prompt rather than to other user comments. In the timed comments in our sample data, the users were more likely to respond to each other and even generated mini-threads. These are typically not very long, and therefore it is easy to look through an entire mini-thread. Taken together, these results suggest that timed comments that focus on particular parts of an audio recording have a greater potential to facilitate user commenting and user interaction on SoundCloud than do regular comments.

At the same time, our analysis shows that timed comments are lacking in certain respects. As described above, sequential coherence is sometimes weak due to the loose semantic relatedness of responding propositions. Moreover, there is little detailed information in the timed comments. The SoundCloud site states that it provides timed comments on audio waveforms primarily for users to provide detailed feedback on the audio recordings.¹ We counted the number of timed comments that included detailed feedback, such as reactions pointing out a

1. "Timed comments let your friends and fans give you valuable feedback at specific moments throughout the waveform" (<http://soundcloud.com/tour/>, retrieved November 1, 2012)

particular part where a listener was impressed and indicating the reason why he or she likes/dislikes the track, and found only 21 detailed timed comments out of 1111 timed comments, or 1.2%. This difference between the SoundCloud system design and users' commenting behavior is one possible area in which to improve the site's usability. It could provide an improved timed comment function; for example, the system could enable creators to identify particular parts of their recordings on which they wish to solicit feedback. A player that enables listeners to see comments during listening might also facilitate detailed user feedback on particular parts. Conversely, if a primary goal of commenting on the site is communication among users without reference to particular parts of a musical entry, the system should implement other functions to support interactive communication on music entry pages. User testing would be required to explore these possibilities further.

Conclusion

In this paper, we conducted two analyses using speech act and dynamic topic analysis, which are computer-mediated discourse analysis methods, in order to begin to understand the common communicative and interactional behaviors of users of the SoundCloud music distribution site. Results of speech act analysis showed that 56% of comments are *REACT* and the average utterance length is only 3.4 words. The results indicated a tendency for users to simply react to audio recordings and/or previous comments, rather than exchanging detailed informative comments. In addition, we attempted to identify which type of commenting function best facilitates user commenting activity on the site: "timed" or "regular" commenting. The results of the two analyses showed that the timed comment function tends to encourage commenting more than the regular comment function, and that comments are more consecutively related in timed comments than in regular comments. We conclude from these findings that a commenting function that allows users to insert their comments in particular parts of an audio recording has good potential to facilitate user interaction and engagement with a music streaming website, taking into account the issues noted in the previous section.

This paper is an initial exploration of a relatively new and unstudied type of social media. Further research is needed to analyze other music entries on SoundCloud in order to validate the findings presented here and investigate them more deeply. In addition, it would be useful to analyze comments on other music streaming sites using computer-mediated discourse analysis techniques in order to compare the effects on communication of different site designs.

References

- [1] [1] Cho, T., Linguistic features of electronic mail in the workplace: A comparison with memoranda. *Language@Internet*, 7, article 3, 2010.
- [2] Werry, C. C., Linguistic and interactional features of internet relay chat. In S. C. Herring (Ed.), *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives*, 47-63. Amsterdam/Philadelphia: John Benjamins, 1996.
- [3] Thelwall, M., MySpace comments. In *Online Information Review*, 33(1), pp. 58 - 76, 2009.
- [4] Herring, S. C., Computer-mediated discourse analysis: An approach to researching online behavior. In: S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for virtual communities in the service of learning*, 338-376. New York: Cambridge University Press, 2004.
- [5] Weigl, D. M., & Guastavino, C., User studies in the music information retrieval literature. In *Proceedings of International Society for Music Information Retrieval*, pp. 335-340, 2011.
- [6] Stowell, D., & Dixon, S., MIR in school? Lessons from ethnographic observation of secondary school music classes. In *Proceedings of International Society for Music Information Retrieval*, pp. 347-352, 2011.
- [7] Barthet, M., & Dixon, S., Ethnographic observations of musicologists at the British Library: Implications for music information retrieval. In *Proceedings of International Society for Music Information Retrieval*, pp. 353-358, 2011.
- [8] Bach, K., & Harnish, R. M., *Linguistic communication and speech acts*. Cambridge, MA: MIT Press, 1979.
- [9] McLaughlin, M., Ch. 4. *Conversation: How talk is organized*. Thousand Oaks, CA: Sage, 1984.
- [10] Wikipedia. *SoundCloud*. Retrieved November 4, 2012 at 6:30 pm, EST from <http://en.wikipedia.org/wiki/SoundCloud>
- [11] Herring, S. C., Das, A., & S. Penumarthy. CMC act taxonomy, 2005. Available at <http://www.slis.indiana.edu/faculty/herring/cmc.acts.html>
- [12] Herring, S. C., Kutz, D. O., Paolillo, J. C., & Zelenkauskaitė, A., Fast talking, fast shooting: Text chat in an online first-person game. In *Proceedings of HICSS-42*. Los Alamitos, CA: IEEE Press, 2009.
- [13] Francis, G., & Hunston, S., Analysing everyday conversation. In M. Coulthard (Ed.), *Advances in Spoken Discourse Analysis*, 1-34. London: Routledge, 1992.
- [14] Herring, S. C., Dynamic topic analysis of synchronous chat. In *New Research for New Media: Innovative Research Methodologies Symposium Working Papers and Readings*. MN: University of Minnesota School of Journalism and Mass Communication, 2003.
- [15] Herring, S. C., & Kurtz, A. J., Visualizing dynamic topic analysis. In *Proceedings of CHI'06*, 1-6. New York: ACM.
- [16] Honeycutt, C., & Herring, S. C., Beyond microblogging: Conversation and collaboration via Twitter. *Proceedings of the Forty-Second Hawai'i International Conference on System Sciences (HICSS-42)*. Los Alamitos, CA: IEEE Press, 2009.