Understanding User Behavior on Online Music Distribution Sites: A Discourse Approach

Hiromi Ishizaki, KDDI R&D Laboratories, Inc.
Susan C. Herring, Indiana University
Gen Hattori, KDDI R&D Laboratories, Inc.
Yasuhiro Takishima, KDDI R&D Laboratories, Inc.

Abstract
We analyze user behavior on two music distribution websites, SoundCloud.com and Last.fm, using a computer-mediated discourse analysis approach. The broad goal is to infer patterns of use that can inform the development of system/site designs to facilitate communication among online users. The specific goal is to analyze the actions users perform through comments posted to the sites. Based on a manually-coded sample of 49,324 comments, we address two research questions: What are the most common speech acts used in comments on the music distribution sites, and how, if at all, do design differences – between the two sites and between the two commenting modes available on SoundCloud – affect speech act use? From these comparisons, we identify site purpose (SoundCloud as a music sharing site, Last.fm as a music streaming site) and commenting mode (regular vs. timed comments) as important factors that contribute to the characteristics of user comments, and suggest implications of these findings for the design of music distribution sites.
Contact: ishizaki@kddilabs.jp, herring@indiana.edu, gen@kddilabs.jp, takisima@kddilabs.jp

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, the increasing popularity of video and music distribution sites has been accompanied by a change in users' activities, from just listening to creating and sharing their own audio recordings. SoundCloud is a popular music distribution site where users can upload their own audio recordings; they can also search and download audio recordings based on their musical preferences. These activities generate many benefits for users: Users who create audio recordings can share them easily and collect immediate listener feedback, and users who want to listen to music specific to their preferences can search for new music that is not available commercially. Furthermore, they can download the recordings they enjoy and leave comments to express their views on recordings and potentially improve them. It follows that facilitation of actions such as user commenting is desirable for users, as well as for developers of multimedia content distribution sites. In order to facilitate the above-mentioned user actions, however, it is important to first understand user behavior on a given site.

Ethnographic approaches enable rich qualitative analysis of user behavior. However, on music distribution sites such as SoundCloud and Last.fm (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 participate anonymously or pseudonymously. Although it might be possible to obtain and analyze logs of user actions from the sites, such as what parts of the site users clicked on, it is challenging to extract meaningful interpretations from such data. In contrast, comments left by users are in themselves an important form of behavior and can reveal much about users' thoughts, impressions, emotions, and intentions. Content analysis is a methodology that permits inferences to be drawn from objective, systematic analysis of specified characteristics of messages (Holsti, 1969). In this study we apply computer-mediated discourse analysis (CMDA), an approach described by Herring (2004) as "[online] language-focused content analysis," to analyze user behavior on two music distribution sites: SoundCloud and Last.fm. Specifically, we apply speech act analysis to analyze the meaning and functions of user comments. This method enables both qualitative and quantitative analysis of comment data in online environments where it is otherwise difficult to observe user behavior directly.

2   Related Work

In the field of computer-mediated communication (CMC), many researchers have analyzed user behavior on communication services such as email (e.g., Cho, 2010), chat (e.g., Werry, 1996), and social network sites (SNS) (e.g., Thelwall, 2009). On these communication services, the main datasets for analysis are text data such as comments and other messages. Because communication among users on these services is primarily via (typed) discourse, computer-mediated discourse analysis (CMDA), an approach grounded in linguistics, was developed to analyze online behavior (Herring, 2004). CMDA has also been used to analyze textual communication in multimodal environments, such as multiplayer online games (Herring et al., 2009) and interactive television (Zelenkauskaite & Herring, 2008). Additionally, social science approaches have been applied to understand user behavior in multimedia systems. For example, some researchers have adopted ethnographically-based approaches to observe user behavior directly and analyze the features of systems in order to find key features and make recommendations for multimedia system design (e.g., Barthet & Dixon, 2011; Stowell & Dixon, 2011; Weigl & Guastavino, 2011). These two approaches analyze user behavior from different, complementary perspectives. Since the data for the present study are exclusively textual, and the researchers do not have access to the users, the CMDA approach was preferred.

3   Research Questions

It is important to analyze user behavior for its own sake, as well as to identify key features on web services and multimedia systems in order to make recommendations for system improvements in terms of user experience. User behavior in different services and systems relates to specific activities on the service platform. For example, listening relates to users' search activities on multimedia distribution sites. In addition, many users post comments and interact with other users by exchanging asynchronous comments. Commenting on multimedia distribution sites is basically realized by short texts. Although often brief, such texts can reveal clues to users' internal states, and are themselves manifest behavior.

In this study, we analyze comments on SoundCloud and Last.fm as a means to identify tendencies in user behavior. In addition, we identify differences in tendencies on the two distribution sites and seek to explain characteristic behaviors that distinguish the sites. Accordingly, the research questions addressed in this study are:

RQ1: What are the most common communicative acts that users engage in through posting comments on the two music distribution sites?

RQ2: What are the main factors that distinguish between the two sites and between the two commenting modes available on SoundCloud in terms of speech act use?

4   Analysis of User Behavior 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, informing, or thanking (Bach & Harnish, 1979; McLaughlin, 1984); speech acts are the building blocks of discourse. In CMDA, act analysis is situated at the level of meaning, which includes pragmatic meanings of language use in context. Analyzing speech acts sheds light on user behavior by revealing the meanings of utterances in context and the messages they purport to convey.

4.1   Data Sample and Coding

Our sample data are user comments posted to SoundCloud and Last.fm. Last.fm is a popular online audio distribution/streaming platform that was created in 2003 as an internet radio-based social network site. The platform distributes music that users want to hear, as well as supporting communication among users by means of blogs, forums, and a 'friend finder.' Users can access songs they like by typing in the name of the song or artist directly or based on recommendations from the system, and they can connect to other users with similar musical tastes, sometimes leading to the formation of friendships and community (Baym & Ledbetter, 2009).

SoundCloud 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. 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 sets it apart from other music streaming sites such as Last.fm is that SoundCloud allows users to insert a comment at a specific point in time of the waveform, 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

point in the song that especially impresses them. We refer to these as "timed comments." Text comments can also be made on the track as a whole; these appear below the waveform, similar to text comments posted below YouTube videos. We refer to these as "regular comments." The two modes of commenting are illustrated in Figure 1.
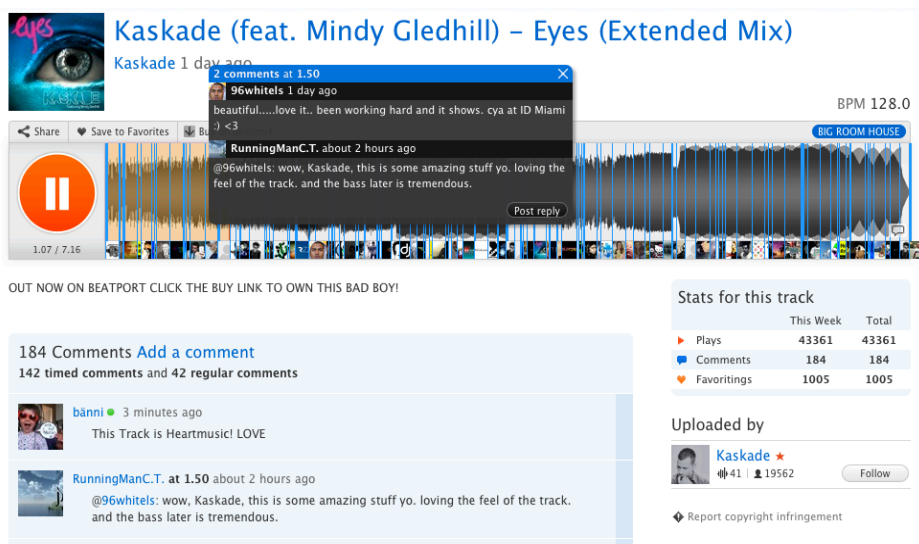

Figure 1. The SoundCloud interface, with two "timed comments" expanded in the waveform. "Regular comments" appear below the waveform on the left.

In order to analyze speech acts on the sites, we first collected 200 music entries from SoundCloud in October 2012. Each entry had more than 100 comments, and some had as many as 1000 comments. From this sample we randomly extracted 58 music entries from the popular genres of "house music" and "pop music." After collecting the SoundCloud entries, we collected all the entries from Last.fm that included the same songs as the SoundCloud sample, for a total of 11 entries. As data for analysis, we extracted all 24,111 comments posted by 14,797 users on SoundCloud from the 58 entries, and all 4,020 comments posted by 2,277 users on Last.fm from the 11 entries. (On average, each SoundCloud entry had 416 comments, and each Last.fm entry had 365 comments.) We divided the comments into utterances based on sentence-final punctuation, resulting in 45,297 utterances in the SoundCloud dataset and 7,971 utterances in the Last.fm dataset. Two coders then manually assigned the most appropriate speech act to each utterance using the CMC act taxonomy developed by Herring, Das, and Penumarthy (2005). 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, along with a description and an example of each act. 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.

CODE	Description	Example
ACCEPT	Concur, Agree, Acquiesce	Yes, I agree.
APOLOGIZE	Humble oneself, Self-deprecatory	Oops my fault :(
CLAIM	(Make a subjective assertion; unverifiable in principle) Assert, Guess, Speculate	I love pizza!
DESIRE	(A cover term including three categories of unrealistic situation) Desire, need (desiderative) , hope, wish dream, speculate (hypothetical, counterfactual) , promise (future action)	I wish I could go with you.
DIRECT	(Attempt to cause action) Require, Prohibit, Permit, Strongly advise	Cool down.
ELABORATE	Comment on, Explain, Paraphrase a previous utterance (usually one's own)	(I can't fake ill…) mum's a teacher
GREET	Greeting, Leave Taking, Inquiries about/wishes for well-being	Hi roley!! / How r u?
INFORM	(Provide "factual" information; verifiable in principle, even if untrue) Inform, State	The capital of India is New Delhi.
INQUIRE	(Seek information) Neutral/Marked Proposal	How long does it take?
INVITE	(Seek participation/acceptance by A) Solicit input, Include,	Let's go outside

3

	Suggest, Offer (Provide goods or opportunity)	
MANAGE	(Manage discourse) Organize, prompt, focus, open or close discussion, preamble, etc.	OK let's get started.
REACT	(Show listenership, engagement - positive, negative, or neutral) Endorse, Approve	Cool!! Eww, ick!
REJECT	Disagree, Dispute, Challenge	No you can't!
REPAIR	Return, Clarify, Correct Misunderstanding	Did you mean "the truth"?
REQUEST	(Seek action politely) Direct or Indirect Request	Can you help me find it?
THANK	Appreciate, Express Gratitude	thxs for showing me / you're welcome

Table 1. Codebook for speech act analysis based on the CMC act taxonomy (Herring et al., 2005)

Following initial training, the coders coded two randomly-selected practice music entries from SoundCloud (524 utterances) that were not part of the sample, after which any issues that arose concerning the codebook and coding process were discussed and resolved. After the two coders independently coded each utterance, we extracted the coded utterances on which both coders agreed. In the end, 49,324 coded utterances (43,190 from SoundCloud and 6,134 from Last.fm) were used for the speech act analysis.

In order to facilitate applying the coding scheme to such a large dataset, we implemented a coding system with three main functions: intuitive coding with a touch screen interface, automated inter-rater agreement computation, and visualization of the coding results. For the touch screen interface, we implemented the coding application on an Android tablet (Figure 2). With this application, coders can apply the speech act coding scheme without a keyboard, just by tapping to select an act from the preset speech acts list. We implemented the prototype application on the HP Slatebook 10 x2 (ANDROID OS 4.2.). After the person administrating the coding prepares the input dataset, the coding application downloads the data from a server via a Wi-Fi network. If there is no Wi-Fi network, the data can be stored in a micro SD card for the tablet.
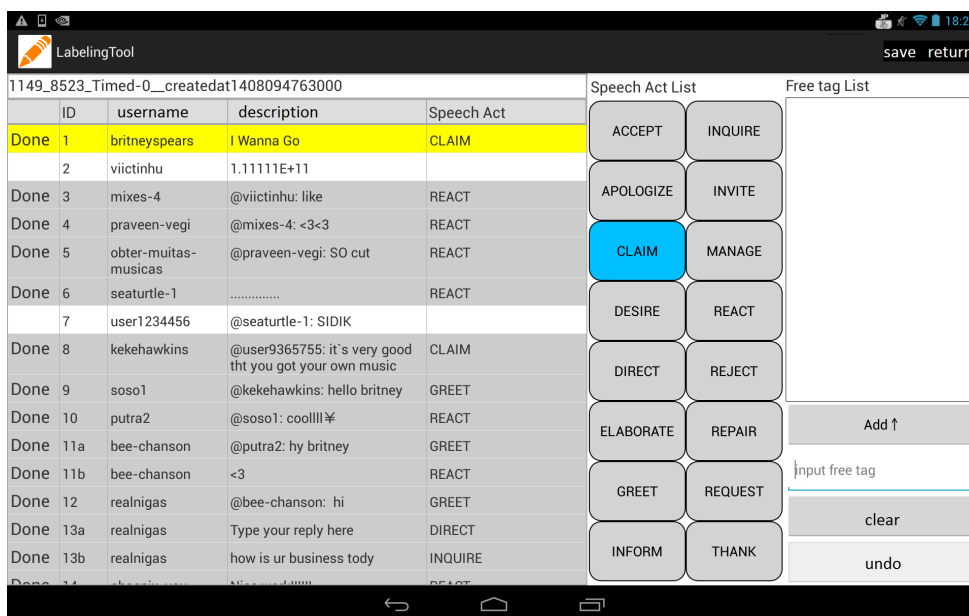

Figure 2. Screenshot of the coding application on Android

4.2   Results: Comparison between SoundCloud and Last.fm

Table 2 shows the frequency of each speech act category in the combined data. As can be seen from the table, CLAIM, REACT, and INFORM are the most common acts in the data. CLAIM is the most frequent, with a combined frequency of 42.1%. The second highest percentage is 33.6% for REACT, followed by 8.3% for INFORM. These results indicate that users tend to comment in order to express their impressions of, and reactions to, a music entry.

Overall, many short CLAIM and REACT comments conveyed concise and simple impressions of the music entry. These comments typically did not occur in conversational exchanges. Indeed, there were

few interactions among the commenters: In the combined data, only 12.9% (3,631) of the comments replied to previous comments. These results suggest that the motivation for commenting on SoundCloud and Last.fm is not so much interaction as expression of simple impressions and reactions to the audio recordings. In further support of this observation, the average utterance length in the sample is only 4.0 words, which is shorter than that of other asynchronous CMC modes such as email (Cho, 2010) and shorter, even, than utterances in much synchronous chat (e.g., Cherny, 1999).

Speech Act	Combined	(%)	SoundCloud	(%)	Last.fm	(%)
ACCEPT	201	(0.4)	175	(0.4)	26	(0.4)
APPOLOGIZE	62	(0.1)	54	(0.1)	8	(0.1)
CLAIM	20776	(42.1)	17060	(39.4)	3716	(60.6)
DESIRE	393	(0.8)	366	(0.8)	27	(0.4)
DIRECT	2060	(4.2)	1981	(4.6)	79	(1.3)
ELABORATE	1252	(2.5)	1222	(2.8)	30	(0.5)
GREET	650	(1.3)	641	(1.5)	9	(0.1)
INFORM	4099	(8.3)	4002	(9.3)	97	(1.6)
INQUIRE	880	(1.8)	761	(1.8)	119	(1.9)
INVITE	28	(0.1)	28	(0.1)	0	(0.0)
MANAGE	26	(0.1)	22	(0.1)	4	(0.1)
REACT	16564	(33.6)	14707	(34.1)	1857	(30.3)
REJECT	30	(0.1)	19	(0.0)	11	(0.2)
REPAIR	10	(0.0)	10	(0.0)	0	(0.0)
REQUEST	944	(1.9)	807	(1.9)	137	(2.2)
THANK	1349	(2.7)	1335	(3.1)	14	(0.2)
TOTAL	49324	(100.0)	43190	(100.0)	6134	(100.0)

Table 2. Coding results of speech act analysis on SoundCloud and Last.fm comments

In order to compare speech act tendencies in SoundCloud and Last.fm in more detail, we calculated the difference between the speech act percentages shown in Table 2 for the two sites. Figure 3 graphically displays the differences that result when the Last.fm percentage is subtracted from the SoundCloud percentage; the bars extending to the left indicate speech acts that occur especially on Last.fm, while the bars extending to the right indicate speech acts that mainly occur on SoundCloud.
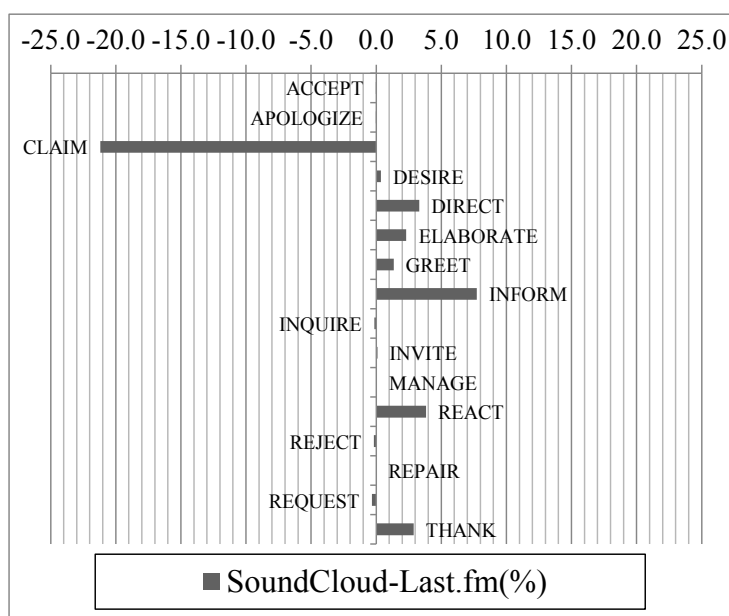

Figure 3. Differences in speech act percentages on two distribution sites: SoundCloud (%) – Last.fm (%)

Figure 3 shows that CLAIM comments are more frequent on Last.fm, whereas INFORM, THANK, DIRECT, and GREET comments are more frequent on SoundCloud. We conducted Chi-square analysis

in order to test the difference in act distribution between these two services and confirmed that the difference between them is statistically significant at the alpha level (alpha=0.01, p<0.01).

CLAIM comments are most frequent overall. A closer examination of CLAIMs on both sites revealed that when users wrote their emotional feelings and impressions of a song, they generally did so in the first person. In addition, the CLAIM comments on SoundCloud often expressed support for the song creator, with compliments such as "Support your song!," "You are a legend," "I like your sound again and again," and "you're a genius." In fact, 29.0% of the CLAIM comments (4,939 utterances) on SoundCloud include a second person pronoun. This observation is in sharp contrast to Last.fm, where only 3.5% of CLAIM comments include a second person pronoun (130 utterances). CLAIM comments on Last.fm tend to mention the artist in the third person – for example, "I feel like she'll be around," "I really like how she sounds," "Thats cuz they're perfection <3," and "her best song since Gimme More." We found this difference in person reference even for songs with the same title.

INFORM is the third most common user act, as can be seen from the combined data in Table 2. However, Figure 3 shows that INFORM comments are more specifically associated with SoundCloud. Some INFORM comments on SoundCloud include URLs and links to other songs; others, such as "Blogged at [URL]," share the origins of the music blogged about. The relatively high percentage of INFORM acts appears to be related to the purpose of SoundCloud, which includes promoting/distributing songs and sharing information about them. In contrast, INFORM comments on Last.fm typically provide a URL that enables the user to download the song. We found only one on Last.fm INFORM comment that might inform the creator/artist about blogs or reviews: "Blogged about it here: [URL]."

Other speech act categories also suggest characteristic behaviors. On SoundCloud, THANK comments usually express appreciation for downloadable MP3 files, such as "Thanks for the freeDL =)" and "too sick, thx for the dl." DIRECT comments tend to direct listeners to the commenters' blog sites or to their audio recordings on SoundCloud; typical examples include "Check out the blog post at [URL]" and "Check out my latest track!" Finally, although INQUIRE acts are not favored by one site more than the other, on SoundCloud users are more likely to ask for information about the system, as well as personal information, such as "how do u download??" and "Who is [ARTIST'S NAME] ex-girlfriend?".

4.3   Results: Comparison between Timed Comments and Regular Comments

For further analysis, we broke the SoundCloud data into two subgroups, regular and timed comments, and compared them. Table 3 shows the frequency of each speech act category for regular and timed comments on SoundCloud. CLAIM, REACT and INFORM are the most common acts in both subgroups, consistent with the overall results in section 4.2. Generally, both regular and timed comments convey concise, simple impressions of the music entry.

Speech Act	SoundCloud			
	Regular	(%)	Timed	(%)
ACCEPT	18	(0.1)	157	(0.6)
APPOLOGIZE	18	(0.1)	36	(0.1)
CLAIM	5907	(39.1)	11153	(39.6)
DESIRE	225	(1.5)	141	(0.5)
DIRECT	997	(6.6)	984	(3.5)
ELABORATE	1117	(7.4)	105	(0.4)
GREET	210	(1.4)	431	(1.5)
INFORM	1879	(12.5)	2123	(7.6)
INQUIRE	296	(2.0)	465	(1.7)
INVITE	10	(0.1)	18	(0.1)
MANAGE	12	(0.1)	10	(0.0)
REACT	3562	(23.6)	11145	(39.6)
REJECT	2	(0.0)	17	(0.1)
REPAIR	2	(0.0)	8	(0.0)
REQUEST	391	(2.6)	416	(1.5)
THANK	437	(2.9)	898	(3.2)
TOTAL	15083	(100.0)	28107	(100.0)

Table 3. Coding results of speech acts: Comparison between SoundCloud regular comments and timed comments

In order to visualize the differences between the two commenting modes, we calculated the difference between the percentages of each speech act category for regular and timed comments as we did in Figure 3. Figure 4 charts the percentage differences; acts on the left side of this figure occur more in timed comments, and acts on the right side occur more in regular comments. Thus ELABORATE, INFORM, and DIRECT tend to be expressed more in regular comments, while REACT tends to be expressed more in timed comments. For regular comments, the percentage of INFORM is much higher than that for timed comments. 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).
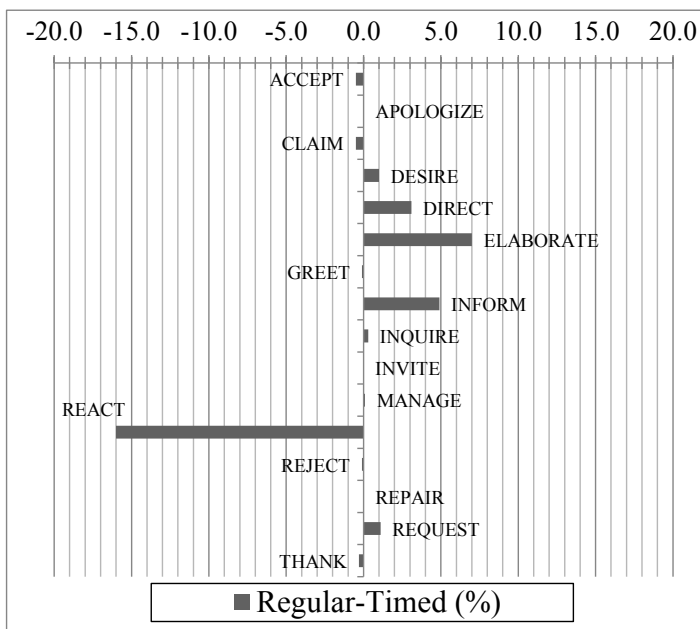

Figure 4. Differences in speech act percentages for two types of comments: SoundCloud regular comments (%) – SoundCloud timed comments (%)

These speech acts can be interpreted as being somewhat characteristic of each comment type. In regular comments in the sample data, we observed many acts that ELABORATE on INFORM acts – for example, to add the title and artist's name after an utterance that provides a URL to a song. It may be that INFORM acts are expressed more often as regular comments because they tend not to be related to particular parts of the track. Conversely, REACT acts mainly express evaluative reactions, as noted in section 4.2, and they tend to be timed comments; these are more related to specific parts of the track. In timed comments, moreover, commenters tend to express the reason after the reaction, such as the CLAIM "Great remix you've got the groove" following the REACT "Incredible!," and the CLAIM "I love this part" following the REACT "Amazing!"

5   Discussion

In response to RQ1, we found that SoundCloud and Last.fm commenters have a tendency to make CLAIMs that briefly express their impressions, such as "I love the song!," and to react simply to the audio recordings and previous comments on the sites, rather than exchanging detailed observations and/or information with other users. This tendency was similar for both sites: The speech act analysis results revealed many CLAIM comments (42.1%) and REACT comments (33.6%) in the combined data. Both CLAIM and REACT are basically subjective evaluations. Furthermore, the comments on both sites are shorter than messages in most other modes of CMC, including email and some chat. It appears that users tend to comment immediately after listening to an audio recording, with short, simple impressions and reactions. These comments do not require much user effort. However, they serve a phatic, as well as an expressive function, in that they connect with other users by expressing shared preferences.

As for RQ2, the speech act analysis revealed systematic differences between SoundCloud and Last.fm. For example, INFORM comments were more common on SoundCloud, and CLAIM comments were relatively more common on Last.fm. The comparisons between the two music distribution sites show

that tendencies of user behavior vary according to several factors, including users' purpose for visiting the site, site content, and site design. As regards purpose and content, we observed community-oriented behavior in the SoundCloud data that seemed particular to the genre of music sharing websites, as distinct from Last.fm, which resembles an Internet radio music streaming site (cf. Baym & Ledbetter, 2009). In their INFORM comments, SoundCloud users commented that they blogged 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. As stated on the SoundCloud web page, "Recording and uploading sounds to SoundCloud lets people easily share them privately with their friends or publicly on blogs, web sites and social networks." Accordingly, SoundCloud music is typically user generated. In contrast, only one of these types of INFORM comments (providing a URL to download the song) was found on Last.fm. It appears that the main motive of Last.fm users is to listen to their favorite songs from famous/popular artists.

Another difference between SoundCloud and Last.fm is the use of personal pronouns in CLAIM comments. SoundCloud users tend to use second person pronouns (e.g., "You are amazing!"), while Last.fm users tend to use third person pronouns (e.g., "She is amazing!") to refer to the music creators, even when the songs extracted from SoundCloud and Last.fm are the same song. This may be because the music creators are more likely to be present on SoundCloud than on Last.fm, which distributes songs by professional recording artists. The differences between SoundCloud and Last.fm in terms of community orientation and use of personal pronouns are thus plausibly caused by the different purposes of the sites: SoundCloud as a music sharing site, and Last.fm as a music streaming site, which in turn has consequences for who can be found on the site. It follows from these observations that vocabularies and specific word frequencies might differ across music distribution sites. Frequencies of words that are not specific to music distribution sites, such as "I" and "you," might be especially good indicators of different sites and would be easy to calculate. In contrast, words that are specific to music distribution sites, such as "music" and "song," are common on both sites and have little value in distinguishing between them.

The comparison between regular and timed comments on SoundCloud also revealed different tendencies regarding use of speech acts. Figure 4 in section 4.3 shows that ELABORATE, INFORM, and DIRECT acts are more common in regular comments. These acts are more likely to refer to the song as a whole than are REACT acts, which occur more commonly in timed comments. These results suggest that commenting format on a site is an important factor that contributes to the characteristics of user behavior. This finding should be explored with a larger dataset including music distribution sites with varied formats.

6   Conclusion

In this study we analyzed over 49,000 manually-coded user comments using speech act analysis, a computer-mediated discourse analysis method, in order to identify the common communicative behaviors of users of two music distribution sites and the main features that distinguish between the two sites. The results of the act analysis showed that a majority of comments fall into the CLAIM and REACT categories, and that the average utterance length is very short. These results indicate a tendency for users to express their impressions of, and reactions to, audio recordings briefly and with minimal effort, rather than to exchange more elaborate comments.

Furthermore, we attempted to identify the factors that condition variation in commenting behavior. The results of a comparison of speech acts between SoundCloud and Last.fm, and between SoundCloud regular and timed comments, showed that the design features of the sites, along with their contents and primary purposes for use, influenced speech act use.

These findings suggest directions for future research. Identifying the common acts on music distribution sites, although just one aspect of user behavior, can facilitate comparison across sites, as well as with user behavior on new, as-yet-unstudied music distribution sites. Other linguistic methods, such as analysis of vocabulary and specific word frequencies in comments, might also be used to characterize the sites.

Moreover, the finding that speech acts tend to differ on the two music distribution sites suggests the possibility of distinguishing among music sites through automated speech act analysis. If an effective automatic speech act estimation technique could be achieved (see, e.g., Twitchell & Nunnemaker, 2004), characteristics could be extracted that indicate which music distribution sites users might be interested in, based on speech act tendencies. System developers might then provide a visualization system based on these characteristics to introduce the sites to new users.

This paper is an initial exploration of a relatively new and unstudied genre of social media. It has several limitations; for example, only two musical genres out of all those that are available on SoundCloud and Last.fm were analyzed. Further research is needed to analyze interactional behavior

around other musical genres, since different genres are associated with different communities of users. Moreover, only one analytical method, speech act analysis, was used; other CMDA methods might reveal other differences (and/or similarities) between how users comment on each site. Finally, it would be useful to analyze comments on other music streaming sites using language-focused techniques in order to conduct a broader comparison of the effects of design features on user communication on this increasingly popular genre of website.

7   References

Bach, K., & Harnish, R. M. (1979). *Linguistic communication and speech acts*. Cambridge, MA: MIT Press.

Barthet, M., & Dixon, S. (2011). Ethnographic observations of musicologists at the British Library: Implications for music information retrieval. In *Proceedings of the International Society for Music Information Retrieval (ISMIR 2011)* (pp. 353-358).

Baym, N. K., & Ledbetter, A. M. (2009). Tunes that bind? Predicting friendship strength in a music-based social network. *Information, Communication & Society, 12*, 408-427.

Cherny, L. (1999). *Conversation and community: Chat in a virtual world.* Stanford, CA: CSLI.

Cho, T. (2010). Linguistic features of electronic mail in the workplace: A comparison with memoranda. *Language@Internet, 7*(3). http://www.languageatinternet.org/articles/2010/2728

Herring, S. C. (2004). 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* (pp. 338-376). New York: Cambridge University Press.

Herring, S. C., Das, A., & Penumarthy, S. (2005). *CMC act taxonomy*. Available at http://info.ils.indiana.edu/~herring/cmc.acts.html

Herring, S. C., Kutz, D. O., Paolillo, J. C., & Zelenkauskaite, A. (2009). Fast talking, fast shooting: Text chat in an online first-person game. In *Proceedings of the 42nd Hawaii International Conference on System Sciences*. Los Alamitos, CA: IEEE Press.

Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley.

McLaughlin, M. L. (1984). *Conversation. How talk is organized*. Beverly Hills, CA: Sage.

Stowell, D., & Dixon, S. (2011). MIR in school? Lessons from ethnographic observation of secondary school music classes. In *ISMIR 2011* (pp. 347-352).

Thelwall, M. (2009). MySpace comments. *Online Information Review, 33*(1), 58-76.

Twitchell, D. P., & Nunamaker, J. F., Jr. (2004). Speech act profiling: A probabilistic method for analyzing persistent conversations and their participants. In *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*. Los Alamitos, CA: IEEE Press.

Werry, C. C. (1996). Linguistic and interactional features of Internet Relay Chat. In S. C. Herring (Ed.), *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives* (pp. 47-63). Amsterdam/Philadelphia: John Benjamins.

Weigl, D., & Guastavino, C. (2011). User studies in the music information retrieval literature. In *ISMIR 2011* (pp. 335-340).

Zelenkauskaite, A., & Herring, S. C. (2008). Television-mediated conversation: Coherence in Italian iTV SMS chat. In *Proceedings of the Forty-First Hawai'i International Conference on System Sciences*. Los Alamitos, CA: IEEE Press.

8   Table of Figures

9   Table of Tables