WHEN USERS GENERATE MUSIC PLAYLISTS: WHEN WORDS LEAVE OFF, MUSIC BEGINS?

Simone Stumpf and Sam Muscroft

Centre for HCI Design, School of Informatics
City University London
Northampton Square
London EC1V 0HB
Simone.Stumpf.1@city.ac.uk, Sam.Muscroft.1@city.ac.uk

ABSTRACT

Music systems that generate playlists are gaining increasing popularity, yet ways to select songs to be acceptable to users is still elusive. We present the results of an explorative study that focused on the *language* of musically untrained end users for playlist choices, in a variety of listening contexts. Our results indicate that there are a number of opportunities for playlist recommendation or retrieval systems, particularly by taking context into account.

Index Terms— Music playlists, music retrieval, end users

1. INTRODUCTION

Music playlists that are automatically created for users have emerged as a particular form of music systems. Based on a user's initial *seed song,* a sequence of songs is selected and played, either by drawing on the user's own song library or a wider music repository. Applications that offer these services (e.g. Last.fm, Pandora.com, and iTunes Genius) are immensely popular, and have received some intense research interest, however, little is understood about how end users use these playlists.

Current approaches have attempted to hone the techniques used to suggest items, and even minor improvements can reap great rewards. We propose that viewing the user's interactions with a music recommender or retrieval system from a perspective of end-user programming could lead to dramatic new insights. The end-user programming framework uses professional programming as a perspective to help end users shape a program's behaviour. Previous research has looked at more traditional programs, for example, how end-users can be supported to construct bug-free spreadsheets [2]. A recent move in this area has been to view user interaction with systems that learn from and adapt themselves to end users as a form of end-user programming. Some of these systems construct procedures by watching the steps that a user carries out [11, 12]; others make classifications based on the user's past history of behaviour [22]. We argue that music recommender and retrieval systems are just another instance of a program with which the end-user would like to interact.

A major first step in designing appropriate end-user programming environments, in this case music recommender and retrieval systems, is to understand the existing concepts of the end user about this task. These concepts are then used as "natural" building blocks for new systems. The Natural Programming methodology [17] – which investigates users' existing approaches to complete a task and ways of organising information by observing without influencing them about how the task should be done – has been used to design programming languages and systems, including interfaces that adapt themselves to user preferences [10]. We employed the Natural Programming methodology to understand what matters to end users in music playlists to inform future system design which encourages user involvement.

This paper reports on an explorative study of how end users put together playlists in different music listening contexts in order to inform the design of playlist systems which can respond to users' demands and provide additional interactions. In our study, we focused on the language that end users employ to describe songs and song selections and how these descriptions differ between listening situations. We investigated, firstly, what "features" an individual uses to describe music and, secondly, how context (i.e. listening situations) influence attention on these features. Hence, our results identify users' concepts, their vocabulary and context-dependent constraints. We suggest future work, which will clarify the implications and options for playlist system design using an end-user programming framework.

2. RELATED WORK

Content-based approaches [19] and Collaborative Filtering [7], alongside their hybrids, are used frequently to recommend individual songs that may be of interest to the user. Current content-based features rely on text-based meta-tag descriptions such as Genre, Artist, Year, as well as automatically extracted audio features [23]. More recently, user-generated tags have been explored as meta-tags [5], in addition to expert-based descriptions (for example, Pandora.com's Music Genome Project).

Consideration of the context in recommendations has been confined to group recommender systems, which aim to coordinate and integrate different users' tastes [15, 16, 9]. Our work differs from group recommender systems in that we describe an individual user's notion of context, not how this could be achieved by taking a number of users' preferences into account.

Approaches that deal with recommending *sequences* of songs have received increased attention in recent years. Playlists are usually generated based on similarity to a seed song the user has selected and may use audio features, meta-tags, end-user "path steering" through a meta-tag space, or "graph-walking" of previous playlists [14, 18, 20, 24, 3]. However, this does not investigate the basis on which *end users*' make their choices.

The concepts that user themselves employ when dealing with music have only recently begun to be investigated. Studies conducted with professional DJs have indicated the need for more expressive meta-tag descriptions of songs [3]. Similarly, professional music searchers for films, TV commercials and computer games use different ways to talk about music [8], including existing musical facets such as Artist, Year, Tempo, etc. but also aspects relating to Mood such as "effervescent" and "quirky". A study of musically untrained users found that Artist, Genre, Style, Event/Activity, Mood, and Tempo are prominent themes when putting together a playlist [4, 6]. None of this research has focused in detail on the language – concepts and their associated vocabulary – that naïve end users draw on for songs in a variety of contexts and its implications for playlist recommendations.

3. STUDY SET-UP

The study consisted of lab-based observational sessions, during which participants were asked to "think aloud" about the songs they were choosing for a playlist (Figure 1). We used a standard "think-aloud" set-up: when a participant fell silent, we prompted them to verbalize their thoughts. The sessions and screen activity were audio and video recorded.

To investigate the role that context plays in playlists, we developed three different hypothetical playlist use cases:

- *Large Party:* a friend's large birthday party of around 50 people;
- *Small Gathering:* a small social gathering with close friends;
- *Private Travel:* a journey on public transport for a weekend trip.

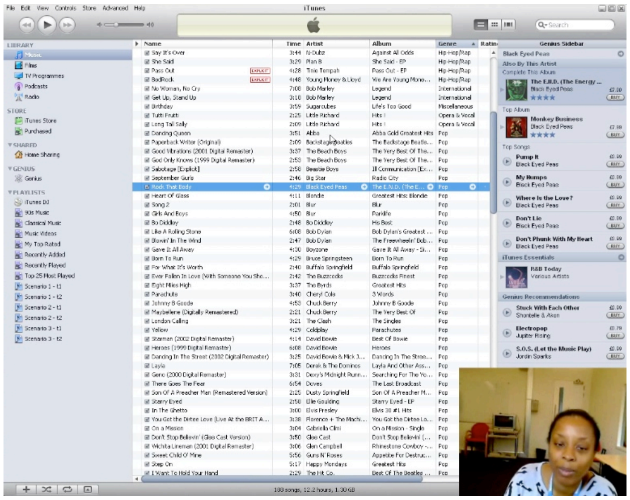
We counterbalanced the order in which use cases were presented to participants.

Each session began with capturing participant demographics, a brief tutorial on the use of the digital media player and a warm-up exercise creating and reviewing a playlist for listening on the way to university or place of work the next day. Then, for each use case, the participant was asked to comment on aspects that would be important in creating a playlist. Their main task was to generate a playlist, by themselves and then via a recommender system. Participants were given 20 minutes to complete the main task. After each task, participants completed a questionnaire asking about their attitudes and perceptions for background analysis.

In our study, seven fluent English-speaking students (three male, four female), ranging in age from 23 years to 48 years (mean 28 years), participated in our study. None had professional interest in music or advanced musical training. Familiarity with desktop digital media players was required; five participants had used recommended playlists previously. Our study followed the Natural Programming methodology to gain an initial understanding of how users generate music playlists and the implications for music recommender and retrieval systems. The early part of this methodology often does not use large user samples but instead explores the range of interactions and behaviours that systems may need to cover with a small number of end users. In the analysis, we use frequencies to provide an indication of importance to our study participants; obviously our sample size does not allow any statistical tests.

As a digital media player, we chose to use iTunes. It has a number of advantages for using in a study such as ours: it is very commonly used, thus does not need extensive additional training for participants, and it also already incorporates a recommender system, iTunes Genius. We used the default set-up and interface features of iTunes (v8.2.1), which does not include any more recent features such as "Ping", etc. Our findings are not dependent on the media player used, as we explored the music choices of our participants.

Creating music libraries for lab-based studies such as ours is difficult: there are financial and copyright considerations, a vast range of music styles that could be of interest, and, most importantly, the ability of participants to



Name	Time	Artist	Album	Genre	Ratin
☑ Say It's Over	3:44	N-Dubz	Against All Odds	Hip-Hop/Rap	
☑ She Said	3:29	Plan B	She Said - EP	Hip-Hop/Rap	
☑ Pass Out	4:28	Tinie Tempah	Pass Out - EP	Hip-Hop/Rap	
☑ BedRock	4:48	Young Money & Lloyd	We Are Young Mone...	Hip-Hop/Rap	
☑ No Woman, No Cry	7:08	Bob Marley	Legend	International	
☑ Get Up, Stand Up	3:18	Bob Marley	Legend	International	
☑ Birthday	3:59	Sugarcubes	Life's Too Good	Miscellaneous	
☑ Tutti Frutti	2:25	Little Richard	Hits !	Opera & Vocal	
☑ Long Tall Sally	2:09	Little Richard	Hits !	Opera & Vocal	
☑ Dancing Queen	3:51	Abba	Abba Gold Greatest Hits	Pop	
☑ Paperback Writer (Original)	2:09	Backstage Beatles	The Backstage Beatle...	Pop	
☑ Good Vibrations (2001 Digital Remaster)	3:37	The Beach Boys	The Very Best Of The...	Pop	
☑ God Only Knows (1999 Digital Remaster)	2:53	The Beach Boys	The Very Best Of The...	Pop	
☑ Sabotage [Explicit]	2:58	Beastie Boys	Ill Communication [Ex...	Pop	
☑ September Gurls	2:46	Big Star	Radio City	Pop	
☑ Rock That Body	4:29	Black Eyed Peas	The E.N.D. (The E...	Pop	
☑ Heart Of Glass	4:11	Blondie	Greatest Hits: Blondie	Pop	
☑ Song 2	2:01	Blur	Blur	Pop	
☑ Girls And Boys	4:50	Blur	Parklife	Pop	
☑ Bo Diddley	2:48	Bo Diddley	His Best	Pop	
☑ Like A Rolling Stone	6:08	Bob Dylan	Bob Dylan's Greatest ...	Pop	
☑ Blowin' In The Wind	2:47	Bob Dylan	The Freewheelin' Bob...	Pop	
☑ Gave It All Away	4:30	Boyzone	Gave It All Away - Si...	Pop	
☑ Born To Run	4:29	Bruce Springsteen	Born To Run	Pop	
☑ For What It's Worth	2:40	Buffalo Springfield	Buffalo Springfield	Pop	
☑ Ever Fallen In Love (With Someone You Sho...	2:42	The Buzzcocks	Buzzcocks Finest	Pop	
☑ Eight Miles High	3:37	The Byrds	Greatest Hits	Pop	
☑ Parachute	3:40	Cheryl Cole	3 Words	Pop	
☑ Johnny B Goode	4:53	Chuck Berry	Johnny B Goode	Pop	
☑ Maybellene (Digitally Remastered)	2:21	Chuck Berry	The Very Best Of	Pop	
☑ London Calling	3:21	The Clash	The Singles	Pop	
☑ Yellow	4:29	Coldplay	Parachutes	Pop	
☑ Starman (2002 Digital Remaster)	4:14	David Bowie	Best Of Bowie	Pop	
☑ Heroes (1999 Digital Remaster)	6:08	David Bowie	Heroes	Pop	
☑ Dancing In The Street (2002 Digital Remaster)	3:25	David Bowie & Mick J...	Dancing In The Stree...	Pop	
☑ Layla	7:05	Derek & The Dominos	Layla And Other Ass...	Pop	
☑ Geno (2000 Digital Remaster)	3:31	Dexy's Midnight Runn...	Searching For The Yo...	Pop	
☑ There Goes The Fear	6:54	Doves	The Last Broadcast	Pop	
☑ Son Of A Preacher Man (Remastered Version)	2:25	Dusty Springfield	Son Of A Preacher M...	Pop	
☑ Starry Eyed	2:58	Ellie Goulding	Starry Eyed - EP	Pop	
☑ In The Ghetto	3:00	Elvis Presley	Elvis 30 #1 Hits	Pop	
☑ You Got the Dirtee Love (Live At the BRIT A...	3:38	Florence + The Machi...	You Got the Dirtee Lo...	Pop	
☑ On a Mission	3:04	Gabriella Cilmi	On a Mission - Single	Pop	
☑ Don't Stop Believin' (Glee Cast Version)	3:50	Glee Cast	Don't Stop Believin' (...	Pop	
☑ Wichita Lineman (2001 Digital Remaster)	3:06	Glen Campbell	Rhinestone Cowboy -...	Pop	
☑ Sweet Child O' Mine	5:56	Guns N' Roses	Appetite For Destruc...	Pop	
☑ Step On	5:17	Happy Mondays	Greatest Hits	Pop	
☑ I Want To Hold Your Hand	2:29	The Hit Co.	Best Of The Beatles ...	Pop	

188 songs, 12.2 hours, 1.30 GB

Fig. 1. The participant (insert bottom right) is constructing a playlist by selecting songs in iTunes. We were using the default layout of iTunes during our studies.

work with an unfamiliar library in a limited amount of time. We therefore constructed a song library, consisting of 200 songs that represented a variety of popular music styles from the past and present, catering to a wide variety of users. All songs were purchased online. Our music library comprised the top 100 best singles of all time as voted for by NME magazine (2002), the top 60 Greatest Songs of all time as voted by Rolling Stone Magazine (2004) and the week's UK top 40 singles (May 1, 2010). We did not include duplicate songs in the library. When it was necessary to replace a duplicate song, we used the next song in the list from the Rolling Stone Magazine top 500 songs of all time.

The analysis was carried out directly on the video itself; video recordings were partially transcribed during analysis. As units of analysis we chose songs and the time spent giving use case descriptions. We employed a Grounded Theory approach [21] to develop a coding scheme. This approach does not use *a priori* codes, instead we identified the vocabulary first, and then derived concepts from the vocabulary as we encountered them in the data through open coding. We used the same approach to analyze participants' open-ended responses from the questionnaires.

Our units of analysis are based on songs and the time that a user took to describe what matters in the use case overall. Any percentages quoted in this paper relate to the amount of concepts applied to these units of analysis. Since the application of concepts to these units may overlap, percentages will not add to 100%.

4. FINDINGS

4.1 Can we leverage users' music descriptions?

The space in which songs could be described is very rich; end users could attend to a myriad of concepts, with language that differs greatly. Participants in our study used a total of 20 concepts (e.g. Mood) overall, with 125 different vocabulary terms (e.g. "sad").

Obviously, if there is no shared vocabulary between users, music systems based on user-generated tags will not succeed. Our results showed there are some idiosyncrasies in users describing music, which will be difficult to overcome: participants did *not* share nearly three quarters of the vocabulary. Even some of the terms which they did share were of a (too) general nature, for example, participants used "a good song" or "nice" to describe what they liked.

Systems could exploit, however, a set of commonly used concepts as organizing categories for songs, especially coupled with a structured vocabulary. We found that there was a large proportion of commonality between participants' concept use, with an average of 12 concepts used per participant, ranging from 45% to 80% of all concepts covered by individual participants.

We now describe the shared concepts and vocabulary examples which participants used in more detail. Since we aim to find commonalities, we do not report rarely used concepts in our results (<1% applied overall).

4.2 What matters in music to users?

If we knew what features matter to users we could attempt to integrate them into the design of music recommender and retrieval systems, especially focusing on the features which matter the most. We will describe the features that mattered to the participants in our study, focusing of two types of features. Table 1 gives an overview of the features found in our study, sorted by frequency which gives an indication of importance.

Table 1. Distribution of concepts over the whole study, sorted in order of frequency of use by participants.

	Percentage of use
Tempo	16%
Mood	14%
Rhythmic Quality	12%
Popularity	11%
Genre	7%
Prominence	6%
Age	6%
Texture	3%
Lyrical Content	3%
Composition	3%
Social catalyst	2%
Audience Type	2%
Volume	1%

4.2.1. Intrinsic song characteristics

Some descriptions are based on what end users know about songs alone; these aspects are *intrinsic* to the songs themselves. There is musical terminology to formally characterize aspects of songs, however, we wanted to investigate what *end users* attended to when listening to songs. We discuss the concepts and vocabulary, in order of frequency of use in our study, which may be an indication for the importance to the general user population.

Audio Content-based Features: There has been much focus on audio feature extraction in current music information retrieval research [e.g. 23]. This could usefully be extended for playlist recommendations as participants focused on audio features to a large extent (36%). The main concepts our participants used were Tempo (16%), Rhythmic Quality (12%), Composition (3%), Texture (3%), and Volume (1%). However, the vocabulary they used to describe these features was not very faceted, even simplistic in nature. Formally, rhythm is the pattern of a musical movement over time; it is often measured by beats per minute and its qualities are expressed through stress and duration (e.g. waltz). Instead, our participants said that they were *"powerful"* or *"not much of a beat"* or that a song was *"danceable"*. In the case of Texture, participants mainly attended to a distinction of voices, such as *"female voice"*. Their differentiation within Composition and Tempo was even more simplistic; it was either *"voice-driven"* or *"instrumental"*, or *"fast"* or *"slow"*.

Mood: Previous research has found that Mood plays an important role in playlists [8, 4, 6] and our results confirm this. Participants frequently mentioned Mood (14%). The range of vocabulary used to describe Mood was very broad, comprising 27 different terms, such as *"sad"* or *"peaceful."* Although this appeared to be an important concept for participants, recommender systems may find it difficult to use these descriptions unless they use a structured vocabulary due to the range of vocabulary.

Age and Genre: Year and Genre are two meta-tags that are commonly used by recommender systems and participants also used concepts related to these meta-tags (13% for Age and Genre combined). When they commented on Age, it was always relative, such as *"new"* or *"classic"*, or they mentioned a specific decade. However, this indicates that recommender and retrieval systems currently only use a fraction of characteristics that users pay attention to and our results suggest that a more flexible approach is required.

Popularity. Popularity (11%) covers a wider concept than simply the number of people playing a song (as in existing recommender systems), in addition it also covers its recognition factor for the potential listeners. Participants were sensitive to this aspect, for example, they considered some songs *"universally recognizable"*.

Lyrical Content-based Features. A previously little-explored concept that could be leveraged by recommender systems is the words set to music and their meaning. Participants paid attention to song lyrics, according to its Content (3%). Participants described a song's content as *"depressing subject"* or *"inappropriate"*. The latter in particular related to either explicit lyrics, or what the participants felt were inappropriate lifestyle choices e.g. taking drugs.

4.2.2. Context-related song characteristics

In addition to intrinsic characteristics, end users may listen to music in a variety of contexts where music is not their primary task or they may not be the only listeners. Hence, some song characteristics are strongly related to context. Three of the concepts we identified fall under this category:

Prominence. Music sometimes takes place in a social setting in which listening to the playlist is not the primary focus. For our participants, Prominence was an important aspect (6%). Participants commented on their song choices' in relation to a primary task, saying that they were *"Distracting", "Background music", "Can talk over it",* or *"Unobtrusive"*.

*Social Catalyst***.** Music can be a social lubricant; participants paid attention to music as a social catalyst in 2% of songs. They constrained playlists by focusing on songs that aimed at *"group reminiscence"*, *"discussion"*, or *"getting people into a party mood"*.

Audience Type. When talking about the listening context, participants paid attention to the particular audience type at which the songs were aimed (2%) and the choices that they made on the audience's behalf. In particular they considered age groups, such as *"diverse age groups"* or *"same age as me"*, and specific audience segments such as *"family and children"* or *"close peers"*.

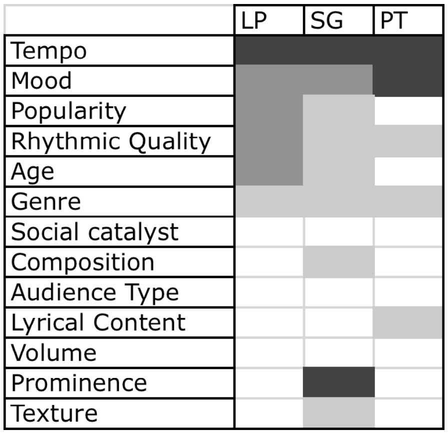
4.3 The effect of context?

If context is addressed in playlist recommender systems, it is usually through modelling preferences that shift over time [13] or group decision-making behaviour [9]. We were interested in the influence that the listening context has on the concepts to which individuals attended; heatmaps can be used to show patterns and differences visually and intuitively. Figure 2 shows a simple heatmap of concepts we previously discussed. The shade of the cell shows the relative frequency of concept use *within* that use case; each shade change indicates a 25% decrease in concept frequency.

It shows that the influence of context is complex, even given just three hypothetical use cases. Tempo and Genre were mentioned by our participants in equal proportion across use cases, indicating that these may be relatively context-independent.

The heatmap also points to concepts which mattered more in certain contexts than in others. For example, in the Private Travel use case, participants used Mood more frequently than in other use cases, whereas in use cases that involved other listeners (Large Party and Small Gathering), Popularity and Age were mentioned more frequently. In addition, in the Large Party use case, Rhythmic Quality was more frequently used. Taken together, this suggests that playlists created for *social* events need to be adapted to take other listeners into account, by focusing on the characteristics of the intended *audience*.

Participants' comments within use cases also



Each shade change indicates a 25% decrease in frequency of use within use case.

Fig. 2. Distribution of concepts within use cases (Large Party, Small Gathering, and Private Travel). Lowest frequency of concept use within a particular use case is shown as lightest shade (highest is shown darkest).

confirmed the task-dependent nature of song choices. Prominence was frequently mentioned in the Small Gathering use case but did not play an important role in others, whereas participants placed more focus on Lyrical Content in the Private Travel use case than in others. This points to the importance of the primary task within contexts, especially given social settings: the main focus may not be on the playlist but when alone you engage more directly with the music.

4.4 Does the playlist need to flow?

Two challenges of playlist recommendations are interaction and ordering of songs [6]. We also paid attention to our participants' comments about ordering and playlist structure. From these comments it appears that they were divided as to whether ordering is important for playlists. In an individual listening context, they often did not attend to ordering, because they can skip easily over songs without having to break off from attending to guests and joining in the party.

Sometimes, listening contexts can matter to playlist ordering. Participants in these instances mentioned that it has to "flow" [6], and common constraints that they used were same Tempo or a playlist Progression, such as "at the beginning of the evening" or "at the end of the playlist". However, most important to them was that they were able to *control* the playlist. Serendipitous inclusion of songs was appreciated by participants but they were less keen on including unfamiliar songs in a social setting. When the participants talked about playlists in these situations, their choices often reflected their personal preference – they knew best.

5. IMPLICATIONS

Our exploratory study has uncovered some potential opportunities and challenges for music playlist recommender and retrieval systems. Our findings have implications for the design of algorithms and interfaces for these systems:

- Interface techniques such as tag suggestions from structured vocabularies may lend themselves to overcome the problem that a large part of the *vocabulary* may not be shared between users. However, a large part of *higher-level concepts* seem to be generalisable and fairly stable across users, and so could already be exploitable by current systems.
- Some of the characteristics we have found are already used extensively in practical applications; for example, Age and Genre are meta-data tags that are already extensively used in current systems. However, users appear to not draw too heavily on these features.
- Extraction of audio features, such as Tempo and Rhythmic Quality, appear to be able to provide substantial pay-off. In addition, ways to extract Mood and Lyrical Content are interesting avenues to pursue.
- Context-dependent aspects have not received much attention so far. For example, context-dependent concepts such as Prominence, Social Catalyst and Audience Type would be an interesting area to explore as features in algorithms and as input parameters for end user interfaces.
- Furthermore, the situation in which end users listen to playlists has an impact on the songs chosen to appear in a playlist. Different song characteristics may come to the fore, particularly in a social setting.
- Control is important to end users, either by being able to skip or re-order quickly, but even more so in the choice of songs to be included.

6. FUTURE WORK

Our study has only begun to examine what matters to end users in generating music playlists. Drawbacks of our study design relate to the small sample size of end users, the limitations of the chosen song library, and the controlled nature of our hypothetical use cases. One way this could be overcome is by drawing on existing music libraries which are based on a large user base, for example Audioscrobbler. However, this cannot replace studies with actual users in a natural setting, taking specific consideration of their interests and contexts into account. Hence, we intend to validate our findings using a larger sample of users in a real setting using their own song libraries in order to explore other contexts in which end users use playlists and the features that they attend to in these situations.

We are proposing a new perspective onto music playlists and interactions with them by end users. End-user programming has been applied to other fields and rests on two main aspects which model professional programmers' approaches: a) *inform the end user about the state of the source code and give feedback about changed run-time behaviour* and b) *provide tools that allow end users to interact with the source code in order to program, test and debug*. Based on the end-user programming framework, future work would focus on explaining how playlists are constructed (i.e. the equivalent of informing end users about the source code and run-time behaviour) and in steering the song selection of playlists, for example through providing new features to attend to in the selection of music items based on context-related characteristics (i.e. provide tools for the end user to interact with the source code) via novel interface functionality. For example, the end user is not usually able to change a recommender's behaviour in any substantial way. By contrast, the user does have more control over the recommendations offered in knowledge-based recommender systems [1], however they still do not allow the user to control the full "source code" of recommendations, including the features that are used. We are interested in exploring novel ways for end users to interact with music recommender and retrieval systems, in order to "program" the playlist generated to be suitable in a particular context.

7. CONCLUSION

Recommended playlists are becoming increasingly popular, yet little is understood about the choices users make when generating a playlist. This paper presented results of an exploratory study that focused on the language of users in a variety of listening contexts. Our results indicate that there are a number of opportunities for extending playlist recommender systems. We found that participants largely shared concepts to describe songs. In particular, audio- and lyrical content-based concepts, alongside mood descriptors and popularity, could supplement existing meta-tags used by recommender systems. However, the interaction has to be carefully managed due to participants' idiosyncrasies. Participants gauged songs as to their suitability to a task and to social situations. When ordering mattered, participants wanted control over the playlist flow. Further research into end users' behaviour in real situations is warranted, to generate better playlists for a range of listening contexts by involving end users.

ACKNOWLEDGMENTS

We thank Charles Inskip and the participants of our study. This work was funded as part of the City University London pump-priming grant scheme.

REFERENCES

[1] Burke, R Knowledge-based Recommender Systems, Encyclopedia of Library and Information Systems, 69(32). (2000)

[2] Burnett, M, Cook, C, Pendse, O, Rothermel, G et al., End-user software engineering with assertions in the spreadsheet paradigm, *Proc. ICSE*, (2003), 93-103.

[3] Crampes, M, Villerd, J, Emery, A, Ranwez, S. Automatic playlist composition in a dynamic music landscape. *Proc. SADPI*, ACM (2007), 15-20.

[4] Cunningham, SJ, Bainbridge, D, Falconer, A. More of an Art than a Science: Supporting the Creation of Playlists and Mixes. *Proc. ISMIR*, (2006).

[5] Gemmis, MD, Lops, P, Semeraro, G, Basile, P. Integrating tags in a semantic content-based recommender. *Proc. Recommender systems*, ACM (2008), 163-170.

[6] Hansen, DL, & Golbeck, J. Mixing it up: recommending collections of items. *Proc. CHI*, ACM (2009), 1217-1226.

[7] Herlocker, JL, Konstan, JA, Borchers, A, Riedl, J. An algorithmic framework for performing collaborative filtering. *Proc. SIGIR*, ACM (1999), pp. 230-237.

[8] Inskip, C, MacFarlane, A, Rafferty, P Upbeat and Quirky, with a bit of a build: Interpretive repertoires in creative music search. *Proc. ISMIR*, (2010).

[9] Jameson, A., Smyth, B. Recommendation to groups. In The adaptive web: methods and strategies of web personalization (pp. 596-627). Springer-Verlag. (2007)

[10] Kulesza. T, Stumpf, S, Burnett, MB, Wong, WK, Riche, Y, Moore, T, Oberst, I, Shinsel, A, McIntosh, K. Explanatory Debugging: Supporting End-User Debugging of Machine-Learned Programs. *Proc. VL/HCC*, IEEE (2010)

[11] Lieberman, H (2001). *Your Wish Is My Command: Programming By Example*. Academic Press.

[12] Little, G, Lau, TA, Cypher, A, Lin J, Haber, EM, Kandogan, E. Koala: capture, share, automate, personalize business processes on the web. *Proc. CHI,* (2007), 943-946.

[13] Liu, NH, Lai, SW, Chen, CY, Hsieh, SJ. Adaptive Music Recommendation Based on User Behavior in Time Slot. *Int. J. Computer Science and Network Security 9*, 2 (2009), 219-227.

[14] Logan, B. Content-based playlist generation: Exploratory experiments. *Proc. ISMIR*, (2002).

[15] O'Connor, M, Cosley, D, Konstan, JA, Riedl, J PolyLens: a recommender system for groups of users, *Proc. European Conference on Computer Supported Cooperative Work,* (2001)

[16] Masthoff, J. Group Modeling: Selecting a Sequence of Television Items to Suit a Group of Viewers. User Modeling and User-Adapted Interaction, 14, 37–85. (2004)

[17] Pane, JF, Myers, BA, Miller, LB. Using HCI Techniques to Design a More Usable Programming System. *Proc. HCC,* IEEE (2002), 198-206.

[18] Pauws, S, Eggen, B. PATS: Realization and User Evaluation of an Automatic Playlist Generator. *Proc. ISMIR*, (2002).

[19] Pazzani, MJ, Billsus, D. Content-Based Recommendation Systems. *LNCS 4321*, (2007), 325-341.

[20] Ragno, R, Burges, CJC, Herley, C. Inferring similarity between music objects with application to playlist generation. *Proc. MIR*, ACM (2005), 73-80.

[21] Strauss, A., & Corbin, J. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. SAGE Publications. (1998)

[22] Stumpf, S, Rajaram, V, Li, L, Wong, W et al. Interacting meaningfully with machine learning systems: Three experiments. *International Journal of Human-Computer Studies 67, 8*, (2009), 639-662.

[23] Tzanetakis, G, Cook, P. Musical genre classification of audio signals. *IEEE Trans. Speech and Audio Processing 10*, 5 (2002), 293-302.

[24] Vignoli, F, Pauws, S, A music retrieval system based on user-driven similarity and its evaluation. *Proc. ISMIR*, (2005)