Interactive Radio: A New Platform for Calm Computing

Citation for published version:
Aylett, MP, Vazquez-Alvarez, Y & Baillie, L 2015, Interactive Radio: A New Platform for Calm Computing. in Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. 2085-2090. DOI: 10.1145/2702613.2732785

Digital Object Identifier (DOI):
10.1145/2702613.2732785

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems

Interactive Radio: A New Platform for Calm Computing

Matthew P. Aylett
University of Edinburgh
School of Informatics
10 Crichton Street
Edinburgh, UK
matthewa@inf.ed.ac.uk

Yolanda Vazquez-Alvarez
University of Glasgow
Lilybank Gardens
Glasgow, UK.
Yolanda.Vazquez-Alvarez@glasgow.ac.uk

Lynne Baillie
Glasgow Caledonian University
Cowcaddens Road
Glasgow, UK
l.baillie@gcu.ac.uk

CHI'15 Extended Abstracts, April 18–23, 2015, Seoul, Republic of Korea.
ACM 978-1-4503-3146-3/15/04.
http://dx.doi.org/10.1145/2702613.2732785

Abstract
Interactive radio is proposed as a platform for Weiser's calm computing vision. An evaluation of CereProc's MyMyRadio is presented as a case study to highlight the potential and challenges of an interactive radio approach: the difficulty of transitioning between passive and active modes of interaction, and the challenge of designing such services. The evaluation showed: 1) A higher workload for MyMyRadio for active tasks compared to default applications (e.g. Facebook app); 2) No significant difference in workload for passive tasks (e.g. listening to audio rendered RSS updates vs Browser app); 3) A higher workload when listening to music within MyMyRadio vs iTunes; and 4) A preference for RSS feed content compared to content from social media. We conclude by discussing the potential of interactive radio as a platform for pervasive eyes-free services.

Author Keywords
Ubiquitous Systems, Interactive Radio, Speech Synthesis

ACM Classification Keywords
H.5.1 [Multimedia Information Systems]: Audio Input/Output, Evaluation/Methodology.

Introduction

In 1991, Weiser and Brown suggested: *"if computers are everywhere they better stay out of the way, and that means designing them so that the people being shared by the computers remain serene and in control."*[9]. Weiser and Brown's vision was not realised, computers are everywhere but they do not stay out of our way. Mobile computing is predominantly stop-and-interact. The web demands our constant engagement with social networking sites continuously emailing their users to get them to re-engage. Yet, we believe there is already a platform available for a pervasive service that could advance Weiser and Brown's vision of calm computing, a platform that until now we have never exploited, the format of radio.

Despite competition from multi-channel TV, and growing Internet content, the popularity of radio has remained surprisingly robust e.g. [7]. The constant pressure from other sources of content on radio media has highlighted its strengths, live content, local content, community relevance[4]. Audio is an example of content which we can selectively attend to, for example, users very often listen to music while they do something else, work, run, drive, make tea etc. Vazquez-Alvarez and Brewster [8] showed that, when designing audio interfaces, there was a significant difference between the user experience of selective attention (where audio was in the background and not requiring the full attention of the user) and divided attention (when two audio streams where competing for the user's attention). This ability for audio to shift between the center of our attention and its periphery fulfills a key element of Weiser and Brown's vision of calm computing[9]. Calm computing argues that systems should remain in the periphery of our attention until we require their services, at which point they would move to the center of our attention for direct interaction.

Radio is so common as a passive medium that it requires a conceptual leap to regard radio as a possible platform for eyes-free interaction. Yet, similar to interactive television, this is what we are proposing (within this paper we use the term *radio* to denote the media format, an audio stream with mixed content, and not the means of transmission, i.e. radio waves *versus* internet).

Interactive television[3] is driven by a commercial model that understands the value of knowing about a user and their historical use of content. In contrast, there has never been a drive to produce a technical solution for interactive radio. For radio there has, instead, been a complete split between live radio streams, downloaded podcasts and personalized music streaming services. Yet, the idea of interacting with radio is not a new one. Local radio stations actively seek to engage users through requests, phone-ins and live competitions. Increasingly, this interaction has been extended from the phone to SMS, email and digital content.

Case Study: MyMyRadio

MyMyRadio was developed by CereProc (www.cereproc.com), a UK speech synthesis company, and was conceived as a *personalized radio service* based on merging: a user's own music, audio dynamically created from news and social media sources, non-speech audio sound effects and background music. It was designed as an antidote to the constant engagement demanded by social media and the web, allowing content to be delivered in the periphery while users listened to music and carried out other activities. If a social, or news headline was of interest, the user would attend to it more closely and could interact directly with the content, moving from a passive or push-down consumption of content to an active or pull-down consumption.

Figure 2: Music player interface for active tasks such as skipping tracks.

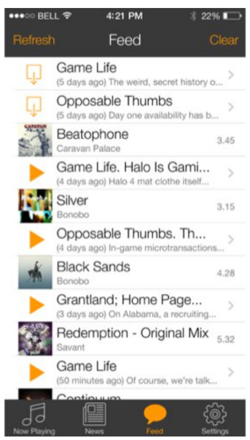

Figure 3: Interface for active tasks such as browsing past and upcoming posts, and music tracks.

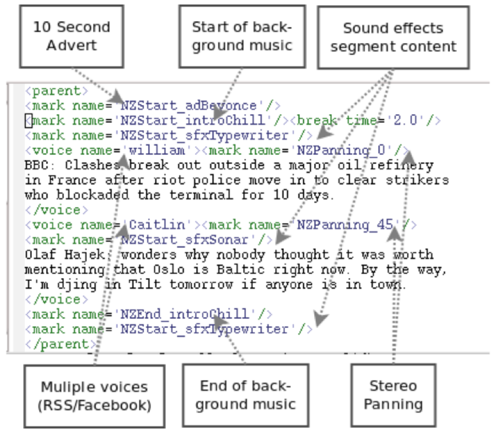

Figure 1: Example of server side XML markup to produce podcast.

This is in contrast with systems which use audio as notification of content, e.g. [5]. Audio notification interrupts current activity. Instead, MyMyRadio inserts content naturally between music tracks to allow continued attention in the periphery. In addition, audio notification does not typically render the actual information. MyMyRadio uses speech synthesis to render the headline[2] so that only content which is of interest to users is brought to their attention. Previous applications have used speech synthesis to present aggregated online social media, such as Twuner, an audio twitter feed for iPhone, and TweetRadio. However, to our knowledge MyMyRadio is unique in merging audio content generated from social media, with non-speech audio and the user's own music to produce a personalised radio service.

A modified music player is used to select music tracks for the user's own music held on the device. A dedicated speech synthesis server renders the user's social media content and RSS news interests as audio. By automatically marking up the content using a dedicated XML markup, a *soundscape* is generated that mixes stereo speech synthesis output with sound effects, background music, and short (10 sec) audio adverts (see Figure 1). Each soundscape is then placed between the users music tracks at a frequency chosen by the user.

Although MyMyRadio is primarily designed for eye-free use, a visual interface is used for active tasks, e.g. Figure 2 and Figure 3. This allows the user to switch between the passive, peripheral use of the system to directly engaging with content, e.g. skipping/reviewing tracks/podcasts and configuring the application, e.g. selecting RSS feeds.

MyMyRadio Design Evaluation

The evaluation was carried out by a research team at Glasgow Caledonian University and extended an earlier pilot evaluation of MyMyRadio[1] (previously called *Noozfeed*). The first phase of the evaluation was comprised of an hour long workshop with 6 usability experts. The aim of the workshop was to identify any usability issues with the initial prototype and to discuss the concepts and use cases behind the design. The concept of MyMyRadio was well received, and the idea of push-down content delivered between music tracks was regarded as desirable. A number of usability issues were highlighted, for example the ease of searching for RSS feeds, confusing button layout, and a number of software crashes connected to skipping tracks. Based on this feedback the application was redesigned for the main evaluation.

Methodology

15 participants were recruited for the main evaluation that was carried out over four days, on the participants' personal iPhone. Participants were asked to carry out a

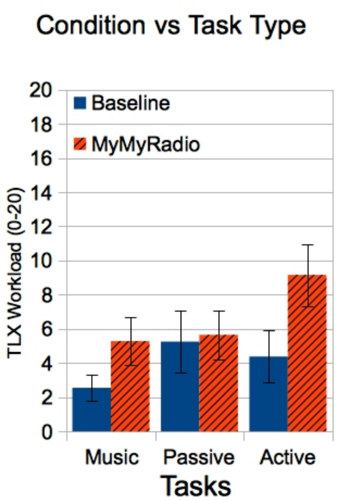

Figure 4: NASA TLX workload means (0-20) by condition and task type averaged over two time points (2 days). Error bars show Standard Error of Mean ± 1.0.

series of tasks with and without the MyMyRadio application and given a set of questionnaires. First, participants were given a set of pre-evaluation questions (age, gender, education level, experience using the iPhone and social media usage) and the experiment instructions. Then, over the four days, participants alternated between two conditions: 1) using MyMyRadio and, 2) Baseline, where participants carried out tasks using their usual methods of accessing social media and playing music. For these two conditions, tasks were categorized as: *music* (listening to music on the device with and without MyMyRadio), *passive* (reading an RSS feed in the baseline condition, listening to RSS updates on MyMyRadio), and *active* tasks (Baseline: Adding new RSS feeds to read; accessing Facebook content and adding Likes; searching through older Facebook updates; adding Likes via Facebook; Join a page liked by a friend. MyMyRadio: Change message frequency; Listen to Facebook updates and add Likes/Comments; Search older message updates; Join a page liked by a friend). Each day participants were asked to rate the difficulty of carrying out the tasks using the NASA TLX workload questionnaire, resulting in two time points (2 days) for each condition. Answers to the final evaluation questionnaire (usability of MyMyRadio, frequency of use, evaluation of speech synthesis voices, preference of social media use within MyMyRadio) were collected at the end of the four-day trial. The hypotheses were:

H1: Active tasks would not be preferred using MyMyRadio.

H2: Passive tasks would be preferred using MyMyRadio.

H3: Listening to music would result in more workload using MyMyRadio.

H1 and H2 offer insight into the relationship between active and passive tasks. H2 supports the MyMyRadio service, whereas H3 would be seen as a negative result for MyMyRadio and call into question whether audio speech content can be perceived in the periphery rather than as an interruption or notification.

Results

10 participants completed the full experiment. A repeated-measures ANOVA was carried out with NASA TLX workload as the dependent variable (averaged across all 6 scales and normalized between 0-20), with time point (1, 2), condition (Baseline, MyMyRadio) and task type (music, passive, active) as within-subjects factors. Time point was significant ($F(1,7)=13.02$, $p<0.01$) showing a learning effect across all conditions and task types (Mean time point 1: 6.513, time point 2: 4.239). Condition was significant ($F(1,7)=11.818$, $p<0.05$) showing an overall higher workload for MyMyRadio (mean= 6.693) than Baseline (mean= 4.059). Task type was significant ($F(2,7)=3.855$, $p<0.05$) with the expected result of workload increasing overall from music (mean= 3.916), passive (mean= 5.448) to active (mean= 6.764) tasks. Condition significantly interacted with task type ($F(2,7)$, $p<0.05$). *Post hoc*, related samples T tests showed a significant difference between Baseline and MyMyRadio music tasks ($t(9)=2.342$, $p<0.05$), and between Baseline active and MyMyRadio active tasks ($t(9)=5.337$, $p<0.001$). See Figure 4 for means.

In the post-evaluation questionnaire, participants rated the ease of using updates for RSS, Twitter and Facebook on a 5-point Likert scale from *very good, good, neither poor nor good, poor, very poor*. A non-parametric Friedman test of differences among repeated measures showed a significant difference between distributions (Chi-square value 9.8, $p<0.01$). *Post hoc* Pair-wise Wilcoxon signed ranks tests showed a significant preference for RSS feeds over Facebook ($Z=2.36$, $p<0.025$), for RSS Feeds over Twitter ($Z=2.12$, $p<0.05$) and for Twitter over Facebook ($Z=2.00$, $p<0.05$). We combined participants' preferences for passive tasks (RSS, Facebook and Twitter Updates) and active tasks (adding likes or joining pages in Facebook). A Wilcoxon signed ranks test showed a significant preference for passive tasks ($Z=2.55$, $p<0.025$). These results support our previous workload results. Active interaction using MyMyRadio was disliked by the participants both in terms of preference and workload and thus we can accept hypothesis H1. There was no evidence that passive tasks were preferred, or required higher workload, in MyMyRadio compared to the Baseline. Thus, we must reject hypothesis H2. There were also significant results showing a higher workload for listening to music in MyMyRadio than in the Baseline, thus we accept hypothesis H3 that adding extra content between tracks increases the perceived workload of listening to music.

Discussion

The result of this evaluation raises a number of issues with the MyMyRadio eyes-free interactive platform. 7/10 participants found the system awkward to interact with and 9/10 participants preferred to use their current methods of checking social media updates. The active tasks were much harder to perform using MyMyRadio compared to the Baseline, whereas there was no significant difference between the workload of passive tasks.

For pervasive eyes-free services to be successful, more work is required to explore the best ways of managing the transition between passive and active use of content. The response to the use of speech synthesis was favorable, with 8/10 agreeing the voices were intelligible, 2/10 neutral, and 10/10 finding the voices pleasant to listen to. One participant reported: *"[I] liked the [CereProc] voices a lot, much more personal and life-like in comparison to Siri"*. When asked specifically in the pre-evaluation whether they would use this eyes-free method to listen to RSS and Facebook updates the result was mixed (3/10 Agree, 3/10 Disagree, 4/10 neutral). This result suggests that despite the difficulties of using the pilot application, it was not the concept as a whole that was received negatively, although it is not a form of interaction that suits all users. In some cases, the advantages of multi-modal stop-and-interact to receive updates is preferred. As one participant remarked: *"I am generally not interested in having an interruption when listening to music. It feels too much like adverts. I usually listen to music and look at news/Facebook"*. The preference in listening to RSS feed updates compared to Facebook could be connected to the challenge in selecting updates that had relevance for the participant. A visual modality is effective for skimming content to find points of interest, whereas audio, with its sequential presentation is less efficient in this process. One participant highlighted the *"need to be able to choose 'favourite' tweeters or 'favourite' friends to be read. Most of the time the tweets/status updates are random people/noise."* Also, another participant commented: *"I like [MyMyRadio] but I prefer how I currently interact with Facebook. I do not like how it links to Facebook but enjoy the ease it provides on the go. Feels like a custom radio."* However, the evaluation process may have skewed these results as all Facebook tasks included active elements which the participants found especially challenging within MyMyRadio. Usability issues were also reported for the MyMyRadio music player. One participant reported: *"I couldn't find a way to simply view the Facebook/BBC News updates without interfering with any music playing."* Given participants are familiar with the music player distributed by iOS, attempting to duplicate and match the functionality of the default system is challenging. Although MyMyRadio used the music player interface available to iOS, it required significant engineering to ensure the system was as robust and effective as the default system. Thus, care must be taken in interpreting H3 as the added workload could also be caused by using a non-standard music player as well as added workload caused by speech content between tracks. Indeed, over time, the workload for the MyMyRadio music task reduced from 6.449 to 4.061, whereas the workload for the Baseline music task remained constant (2.521 - 2.581). One major difference between the pilot evaluated here and the released system was a significant improvement in the functionality and stability of the music player. Issues like background threading, switching processes and interacting with 3rd party APIs in the social networking made MyMyRadio more complex to design and build than the relatively simple underlying idea might suppose. It is not just mobile use that is stop-and-interact, the operating systems on mobile phones are also very strongly oriented to this style of interaction, as well as to a *toolbox* approach to 3rd party applications. Producing applications that fit this format are simpler to build and easier to commercialise.

A more developed interactive radio platform could contain localization information and allow a mixture of localized content, speech synthesis and pre-recorded audio, as well as personalized music streams such as Spotify. An interactive radio platform could help develop a passive-to-active interaction transition model, and offer integration with social media and new digital services.

Conclusion

Future work will examine audio interaction in more depth, and explore further the possibilities of interactive radio. The model of radio has an enduring popularity and there is potential in interactive radio use in mobile environments and could possibly offer a means to realize the calm computing vision that Weiser espoused.

References

[1] Aylett, M., Vazquez-Alvarez, Y., and Baillie, L. Evaluating speech synthesis in a mobile context: Audio presentation of Facebook, Twitter and RSS. In *Proceedings of ITI2013* (2013).

[2] Aylett, M. P., and Pidcock, C. J. The CereVoice characterful speech synthesiser SDK. In *AISB* (2007), 174–8.

[3] Cesar, P., and Chorianopoulos, K. The evolution of TV systems, content, and users toward interactivity. *Foundations and Trends in Human-Computer Interaction 2*, 4 (2009), 373–95.

[4] Denevan, S. L. *The Potentially Bright Future of Radio: An Analysis of Interviews from Radio Professionals Regarding Radio's Past, Present and Future*. PhD thesis, Southern Utah University, 2014.

[5] Dingler, T., Brewster, S. A., and Butz, A. Audiofeeds - a mobile auditory application for monitoring online activities. In *ACM Multimedia* (2010).

[6] Dix, A., Finlay, J., Abowd, G., and Beale, R. *Human-computer interaction*, third ed. Pearson Educational, Harlow, UK, 2004.

[7] Taylor, J. Radio surges in popularity thanks to digital. *The Independent* (1 2009).

[8] Vazquez-Alvarez, Y., and Brewster, S. A. Eyes-free multitasking: the effect of cognitive load on mobile spatial audio interfaces. In *CHI '11*, ACM (2011), 2173–6.

[9] Weiser, M., and Brown, J. S. The coming age of calm technology. In *Beyond calculation*. Springer, 1997, 75–85.