The Cognitive Perception of a Multi-room Music System with Spatial Interaction

Henrik Sørensen, Jesper Kjeldskov[(✉)], Mikael B. Skov,
and Mathies G. Kristensen

Research Centre for Socio + Interactive Design/Department of Computer
Science, Aalborg University, 9220 Aalborg, Denmark
{hesor, jesper, dubois}@cs.aau.dk

Abstract. In recent years we have seen a growing interest in exploring spatial interaction as a means of interacting with computer systems through what has been labelled "proxemic interaction". In order to explore the potentials and challenges of spatial interaction spanning across separate physical locations, we have developed a multi-room music system and performed a field evaluation of use. The system extends Apple AirPlay to allow spatial interaction with one's music player, for example, adapting an App interface to the current location of the user, and allowing music to follow the user around the house. The prototype was deployed in two households over a three-week period, where data was collected through logging, user-written diaries and interviews. The field evaluation revealed a number of findings related to the cognitive perception of the spaces it was used in, such as importance of a simple interaction, the importance of providing local interaction, the challenge of foreground and background interactions, and challenges in designing interaction with music in discrete zones.

Keywords: Proxemic interaction · Spatial interaction · Multi-room · Music

1 Introduction

Our homes are becoming populated by an increasing number of devices technically capable of interacting with each other. New challenges, therefore, emerge within the field of HCI and work is needed to explore how to take advantage of the possibilities that emerge in the home as a ubiquitous computing (ubicomp) environment.

One way of approaching the interaction design of systems within ubicomp environments is through the perspectives of proxemics and spatial interaction. Hall originally coined the term proxemics as a way of describing interpersonal spatial interaction based on physical measures [11]. In his work he defined discrete zones surrounding us, which are meaningful to the way we interact with each other. The notion of proxemic and spatial interaction used in this chapter is based on the elaboration of Hall's work by Greenberg et al. [10], where they apply proxemics to interaction within ubicomp environments. The purpose is to take advantage of the way proxemics influence how we cognitively perceive, and naturally interact, with each other, and apply this knowledge to interaction within ubicomp environments.

Recent work has explored the possibilities of proxemic and spatial interactions for both work related contexts [10, 13] and leisure contexts [3, 7, 10, 29]. Most studies provide valuable results on what can be called the micro level of proxemic interaction. An example is where content on a large display changes appropriately according to the distance and orientation of the user, as shown in the Proxemic Media Player [10]. On a larger scale, the macro level, spatial interaction may span multiple locations. In many cases this introduces a real cognitive difference and although the macro level sometimes consists of several micro level systems, it is important to investigate this type of interaction designs, and see how they are used in real-life contexts.

Music consumption is an area where we see a natural application of proxemic and spatial interactions. Music plays an important role in many people's lives across age, gender and culture. The digitisation of music and advances in mobile and network technologies have opened up for new opportunities in interaction design. Holmquist describes this phenomenon as ubiquitous music [12] and several new advances have recently found its way into the consumer market. An example is the emergence of online music services that contain millions of songs, available through a subscription, making it accessible from several different devices. The growing integration of wireless networks, in our homes, has additionally changed the way we listen to and control music at home. Several music systems, such as Sonos, Bose link, Bang & Olufsen BeoLink and Apple AirPlay, allow us to listen to the same music collection in our entire home, and use mobile devices to control playback. However, on top of the opportunities offered by this new infrastructure there is also an extra layer of complexity for interaction design and for research within the cognitive effects of spatial interaction. How do people cognitively perceive interactive systems based on proxemics and spatial interaction, especially those spanning across separate physical locations? How does the user choose where to play music? How does the user direct control towards a specific location? The same music could of course play in the entire house simultaneously, but what if different persons want to listen to different songs in different rooms? What if the music needs to be louder in a large room and quieter in a small room?

In order to explore some of these questions, we have developed a multi-room music system, called AirPlayer, and performed a field evaluation of use. AirPlayer was designed to hide some of the technical complexity of multi-room music systems through an integration of proxemic and spatial interactions. In our work we focus on how proxemic interactions are used during everyday situations. Due to the importance of the spatial context in which the interaction occurs, findings are based on field evaluations conducted in actual households. In the following we will present related work to proxemic and spatial interaction, spatial cognition, and ubiquitous music. We then describe the AirPlayer system in terms of interaction design and system implementation. We then present the field evaluation, and discuss our findings.

2   Related Work

This section will put our work in relation to existing work to clarify the motivation for exploring proxemic interactions in a music consumption context.

2.1 Proxemics and Spatial Interaction

In the work of Greenberg et al. [10] proxemics and spatial interactions are operationalized in a way suitable to the interaction within ubicomp environments between people, digital devices, and non-digital things. They break proxemics into five specific measurable dimensions: Distance, orientation, movement, identity and location. The dimensions provide discrete and/or continuous measures, which can influence the interaction that takes place. Work originating from this operationalization of proxemics and spatial interaction includes a proximity toolkit for fast prototyping [18], and the use of the sociological constructs, F-formations and micro-mobility, in the design of cross-device interaction [19]. The five dimensions provide a great framework for exploration of proxemics in various ubicomp contexts, and in their work they have furthermore identified six immediate challenges for spatial interaction design [20]. These are, for example, the challenges of directing actions or providing feedback to the user, when the interaction is based on proxemics.

The application of proxemics and spatial interaction in HCI has primarily moved towards different aspects in relation to a central device of focus. Vogel and Balakrishnan [28] have previously specified a set of design principles for public ambient displays. In their work they define a framework, which they refer to as interaction phases. The idea is that the area in front of the device is divided into four discrete phases similar to Hall's proxemic zones [11] surrounding a person. Each phase is determined by the distance to nearby users, and the spatial interaction is described as transitioning from implicit and public to explicit and personal interaction. Recent work has explored similar applications of proxemics, such as public displays [29], whiteboards [13], or tabletops [1, 2]. These studies provide important insight into proxemics in relatively small spaces, and help us understand how proxemics and spatiality can facilitate a different interaction form in such ubicomp environments. There has, however, not been the same focus on proxemics and spatiality on a larger scale.

The idea of exploring proxemic and spatial interactions in larger spaces than a single room, or the immediate area surrounding a display, is of course not entirely new and unexplored. An early, well-known, example is the Active Badge [30]. The Active Badge is aimed at a work context, where employees can be tracked via a wearable badge. The badge contains an infrared beacon from which a sensory network picks up the signal and updates the user position every 15 s Information can for instance be used by receptionists to direct calls to the correct location. The system has successfully been deployed on a large scale at several locations. The UbER-Badge [14] is a different approach to proxemic aware wearable badges. In this case, the badges act as sensor nodes used to facilitate social interaction at large meetings. Each badge contains sensors that can detect other badges or stationary tags. Possible applications include locating other badge-wearers, exchanging contact information wirelessly, and as an interaction device at appropriate locations. The EasyLiving project [4] is an example of proxemic and spatial interaction in the home, with a focus on an architecture that aggregates devices into a coherent user experience. This is accomplished through technologies that track people spatially, and devices and applications binding it together.

Proxemic and spatial interactions on a larger scale, where interaction spans separate rooms of a house, are closely related to the work on indoor positioning, which is a huge research area within ubicomp. Based on measured physical quantity and hardware technology, indoor positioning technologies can be categorised into radio frequency, photonic, sonic waves, mechanical and others [26]. An accurate indoor positioning system is out of the scope of this paper and focus is on the interaction, designed on top of the position system.

2.2   Spatial Cognition

Spatial cognition is a research field dealing with human knowledge and beliefs about the spatial environment around us [21], and working to understand spatial cognition in humans. The field builds on contributions from several disciplines such as Psychology, Geographic Information Science, Human-Computer Interaction, and Cartography. People's spatial perception and cognition is fundamental for our ability to move and navigate through physical space, and for our ability to identify, locate, and track objects and entities in motion [9]. According to recent research, many of the mechanisms that humans make use of in spatial perception and cognition are mainly instinctive, providing us with an ability to store cognitive spatial representations, or cognitive maps [25], for locating themselves, others, things, and directions. In addition to this, because humans can use language for representing space, people are able to create very rich and creative extensions of representations for three-dimensional physical space. This means that physical space plays an important role for humans as a "memorial structure" used to organise memory by, for example, attaching it to specific locations [23].

It is well known that a variety of technologies influence people's spatial cognition, such as the use of the Global Positioning Service (GPS) for navigation, Geographical Information Systems (GIS), and other types of location-based services. Such systems have today become commonplace, and are integrated in hundreds of millions smartphones throughout the world. Nevertheless, it is still a relevant research question to investigate the effect of these systems on people's spatial cognition, and especially how interacting with such systems affect and change people's spatial experience and behaviours, and to investigate how the interaction with such systems is best designed accordingly [21]. How does, for example, the use of location-based services, or proxemics/spatial interaction design, influence people's perception and thinking about their physical surroundings? These questions make it particularly relevant for researchers within the areas of computing and human-computer interaction to apply a spatial cognition perspective on their work, as is exemplified in this chapter.

2.3   Ubiquitous Music

We find the case of music consumption in ubicomp environments particularly interesting as a case for proxemic interaction, as the music itself provides immediate non-visual feedback. This is particularly suitable for a case where location and movement are separated from orientation and in some way distance. Holmquist's description of ubiquitous music [12] embraces music in interaction design for several research

directions. The point is that technological advances have changed the way we listen to music. Liikkanen et al. [17] encourages a renewed interest within HCI in music interaction, due to the cultural, social and commercial significance of music consumption, but also because music interaction as a research topic has become less visible.

A radical change in music consumption is the way advances in mobile technologies have enabled us to listen to music on-the-move. In research this has been manifested in a relatively large body of work on mobile music interaction. Recent interesting examples include MusicalHeart, which recommends music based on heart rate [22], and the +++ wearable player [27] that "infects" passers-by with music while jogging. Ubiquitous music is however not restricted to mobile devices. Integration of wireless network infrastructures in our homes open up for similar novel interaction possibilities. The fact that music has become more ubiquitous does not mean that we have stopped listening to music in fixed locations.

Rose [24] reflected on music consumption in the home, back in 2000, and suggested an architecture and interaction design with a central music library and a touch interface. A similar use of music as a case for interaction designs in ubicomp environments is seen in the work of Chang and Kim [6], where a context-aware music playing service is presented. They use a Bluetooth signal to facilitate location recognition, where moving nodes carried by users are detected by a set of fixed nodes. Fixed nodes are then capable of playing music according to the preference of nearby users. What we see is a potential for smarter music systems capable of utilising various devices and a wireless network infrastructure. It is however important also to explore these advances from an HCI perspective.

3 AirPlayer

AirPlayer has been developed to explore proxemics and spatial interaction for a multi-room music system. As a multi-room music system installed throughout the home, it allows users to centralise storage and management of their music collection. Remote control of the music can furthermore be integrated into existing devices. For interaction, AirPlayer makes use of proxemic spatial interaction principles. By making the system aware of the user's location and movement around the house, AirPlayer is able to infer what part of the music system the user wants to interact with, and where music should be played. The goal of this interaction design is to explore the use of proxemic interaction as a form of remote controlling a multi-room music system, and to explore people's cognitive perception of such system in use.

3.1 Multi-room Proxemic Interactions

The proxemic interactions of AirPlayer are implemented as separate spatial features of the system that can be activated independently. The two spatial features, location and movement, are each based on the corresponding dimension from the operationalization of proxemics by Greenberg et al. [10]. In this section, each of these features will be
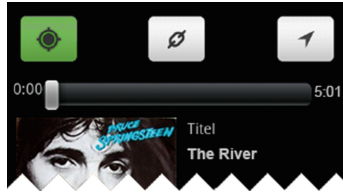
Fig. 1. The location feature can be activated on the main screen of the music player with the touch of a button.

presented along with a concrete scenario and followed by our interpretation of the spatial dimension in AirPlayer.

Location. The location feature of AirPlayer is activated from the top-left corner of the music player interface as shown in Fig. 1.

The location feature allows the graphical user interface of the smartphone application to adapt to the current location of the user. The following scenario, illustrated in Fig. 2 describes a typical use situation.

> *Alice and Bob enjoy a Sunday evening at home. They are in the living room where Alice is reading a book and Bob is browsing the news on his laptop. Bob has previously used the music browser on the AirPlayer smartphone application to queue a number of songs from their common iTunes library. Because the system knew that he was in the living room, it immediately started playing there as he pressed play. Bob had been up early that morning and goes to the bedroom to get a nap. Bob likes to listen to music as he falls asleep and as a big fan of Bruce Springsteen, he takes out his smartphone to put on his playlist of favourite Bruce Springsteen songs. The smartphone application shows that he is currently in the bedroom and as he selects the playlist, the speakers in the bedroom start playing. The smartphone application shows that it is "The River" playing. Alice is still in the living room listening to the same music as before.*
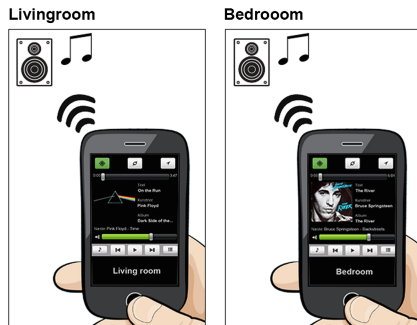

Fig. 2. When the *location* feature is active, the GUI adapts to the current location of the user.

Greenberg et al. [10] defines the location dimension as a description of the physical context in which the entities reside. Entities in this case can be people, digital devices and non-digital things. They use a particular room as an example of a distinct location,

which is also the basis of how it is interpreted in AirPlayer. In AirPlayer the spatial locations are however referred to as zones since the system allows the user to combine separate locations into larger zones. As described in the scenario, the location feature introduces location-awareness of the controlling digital device. When Bob is in the living room with his smartphone, AirPlayer knows this and automatically makes his interactions control the music in the living room. When Bob moves to the bedroom, control is automatically shifted to the bedroom as well, independent of what is playing in the living room. The smartphone application furthermore provides real-time visual feedback to the corresponding zone, showing which zone he is currently in, as well as information about the music playing in that particular zone. Greenberg et al. [10] puts emphasis on the importance of the location dimension, as the meaning applied to the other dimensions can be dependent on the physical context. In AirPlayer this is the case for the movement dimension.

Movement. The movement feature is activated from the top-right corner of the music player interface as shown in Fig. 3.
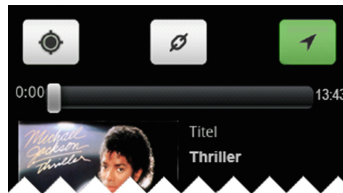

Fig. 3. The movement feature is activated in the same way as the location feature on the main screen of the music player.

The movement feature tracks the user's movement around between different locations. This information is used to make music follow the user around continuously. The following scenario, illustrated in Fig. 4, explains the use situation.

> *Charlie is not much of a morning person but music normally helps him get a fresh start of the day. He uses his music system as an alarm clock and at 7:00 AM his stereo starts playing in the bedroom. He grabs his smartphone and goes to the kitchen to get a cup of coffee. The music stops in the bedroom and "follows" him to the kitchen. In the kitchen he suddenly remembers a great song he has not heard for a while. He picks up his smartphone, finds the song in the AirPlayer application, and puts it in the queue. After he finishes his coffee he goes to the bathroom to take a shower and gets ready for work. The music "follows" him to a small speaker in the bathroom and during his shower the song he queued earlier starts playing.*

In AirPlayer, movement is interpreted as a discrete measure of changes in the user's spatial location. As illustrated by the scenario, it allows music to follow a person around, by tracking the location of a smartphone. As the person moves in between zones, the system is able to anticipate where the user is going, thereby preparing the music in the zone(s) ahead. Similarly the system is aware of the zone a person is moving away from and stops music playback in that place. As illustrated by the scenario, not only the song playing is transferred, but also queues are transferred between zones as the user moves around.
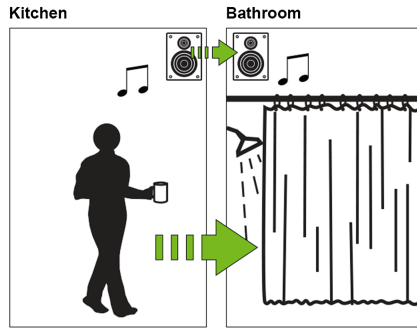
Fig. 4. When the *movement* feature is active, music follows the user around the house automatically.

3.2   Graphical User Interface

The proxemic spatial interaction in AirPlayer is not a replacement for a graphical user interface (GUI), but a supplement that adds to specific parts of the interaction that otherwise usually complicates the control of a multi-room music system. In addition to the location and movement features, AirPlayer therefore also has a more traditional

GUI enabling the user to control other aspects of music playback, such as choosing artists, albums, songs etc. This GUI is available through a smartphone application, which serves as a remote control for the system, and as a display providing visual feedback to the user on their interactions. Using the smartphone application, the user can control music playback in each of the zones through a touch interface, controlling, for example, queuing music on a zone, playing and pausing the music, and adjusting the volume. In order to focus on the proxemics and spatial interactions facilitated by AirPlayer, the GUI functionality is, however, deliberately kept simple, and even when the proxemics and spatial features are turned off, only very basic functionality is available. The GUI consists of three underlying screens: a music player, a queue and a music browser, which we will briefly describe in the following.

Music Player. The music player is the main screen of the application and is also the first screen the user is presented to when starting the application (Fig. 5). The location and movement features can be activated/deactivated by the click of a button placed in the top bar. Once activated, the button will highlight, showing that the feature is active. When movement is active, music follows the user around between zones and naturally so does the control, i.e. the location feature is implicitly active whenever the movement feature is active. The top bar furthermore contains a link button. While the link feature is not directly related to the proxemic interaction, it is used to control the zones, which the proxemic features utilise.

The music playback component is the main part of the music player interface and contains information and controls commonly seen in music players. The interface features a seek bar, controls to start or stop the music, skip to next or previous track, adjust volume, as well as buttons to open the browse and query screens. The music playback component additionally contains a large display of information regarding the
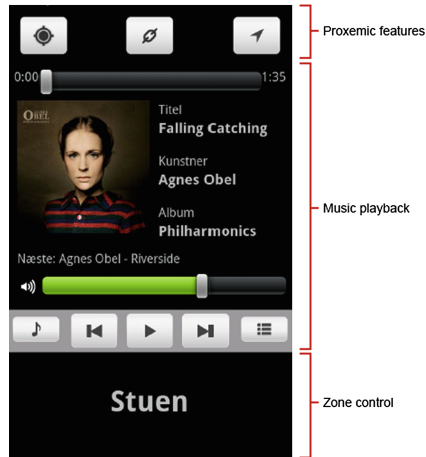
Fig. 5. The music player is the point of entry and primary interaction screen of the application.

music currently playing. The bottom of the music player is reserved for the zone indicator, which displays the name of the current zone being controlled. A swipe gesture to the right or left will manually cycle through the configured zones. When the location feature is active, it will automatically change the zone according to the location of the user.

Queue. The queue screen, shown in Fig. 6, presents the current queue of songs for the active zone, as well as controls to manage the queue. The primary part of the screen is the queue of songs presented in a scrollable list. The clear button in the top right corner removes all songs queued in the active zone, and stops playback. Controls in the bottom can be used to play a song, rearrange the queue, or remove songs.
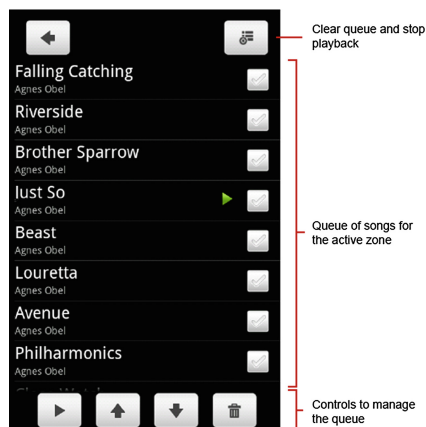

Fig. 6. The queue for the active zone can be accessed and managed through the user interface.

Music Browser. The music browser, shown in Fig. 7, has a top bar with a back button, an add button, and an add-all button. If the user presses the add button, all selected songs are added to the queue of the active zone. The primary part of the music browser, is a list of items from the music library, i.e. playlists, albums, artists and songs. The list is sorted alphabetically and the user can scroll using swipe gestures. The bottom bar of the interface contains a tabbed interface for the user to browse through the music library. For example, when the user selects the Albums tab, a list of available albums from the iTunes library is presented. From here, the user can select an album, and all songs on the album will be displayed on the list. From the song level, the user can also select a number of songs to add to the zone queue, or simply add all songs from the album to the queue. In this way, the user can browse through the iTunes library from his smartphone, and add the desired songs to the queue of currently playing songs.
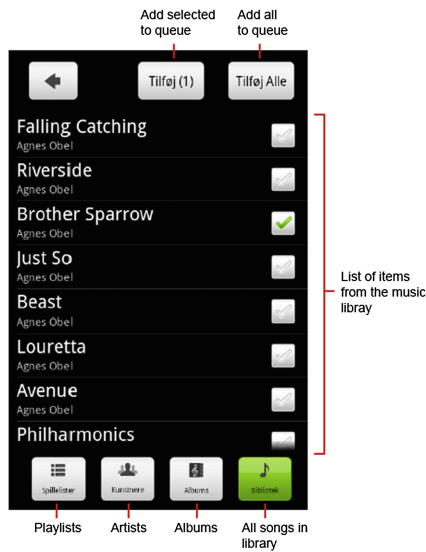

Fig. 7. On the song level of the music browser, songs can be added to the queue.

4   System Implementation

In this section we provide details about the implementation of AirPlayer, including system architecture and details about the location estimation and server application.

4.1   System Architecture

The AirPlayer prototype system platform is build on top of Apple's AirPlay, which allows streaming of media content between Apple products, and Apple compatible products, using a regular wireless network. The AirPlayer infrastructure relies on three types of devices: (1) a smartphone which hosts the AirPlayer client application,

(2) a Mac Mini server that hosts the AirPlayer server application, the music collection, and iTunes, and (3) a number of AirPort Express wireless network base stations, connected to either active speakers or a hi-fi stereo system.

The AirPlayer system architecture is depicted in Fig. 8, which shows relations between components. The server application is installed on a Mac Mini server and is basically a remote controlled music player capable of streaming to AirPort Express stations. It gets the music data directly from the music collection residing on the server. Apple's music player iTunes is also installed on the server, but only serves the purpose of accessing meta-data from the music library, and providing an interface where the user can manage the music collection and playlists. In addition to receiving the music stream, the AirPort Express wireless base stations also provide location measurements handled by the smartphone application. The smartphone application communicates directly with the server application to issue commands like "skip to next song" or "move music stream to another AirPort Express". It furthermore receives data from the server, such as meta-data about the currently playing song, which can be presented to the user. The user can control the system using a smartphone from anywhere in the house directly through touch input or indirectly by activating the proxemic interaction features.
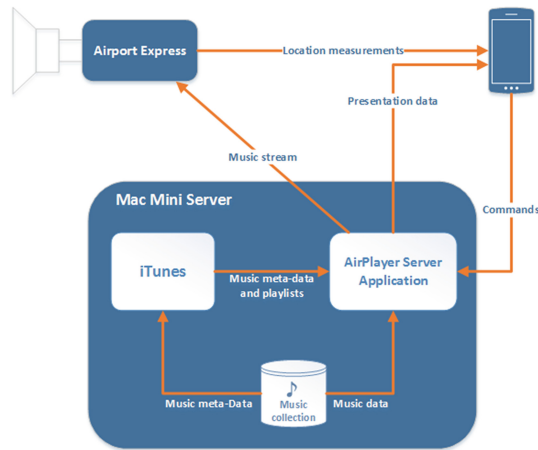

Fig. 8. The system architecture of AirPlayer consists of three types of inter-connected devices.

4.2 Location Estimation

The use of location awareness is a central part of AirPlayer. Both the location and movement features are reliant on knowledge about the user's current spatial position within the environment. In AirPlayer this is achieved through a simple comparison of received signal strength indicator (RSSI) measurements, handled by the smartphone application. For the prototype application, only coarse-grained information regarding the user's position is required, i.e. in which zone the user is currently located and whether his location has changed over time. By using results from Wi-Fi RSSI measurements from the AirPort Express devices, the relative location of the user can be estimated, thus creating the basis for the implementation of the proxemic interaction.

AirPlayer is a self-positioning system [8], meaning that the positioning receiver is responsible for doing appropriate measurements and uses these to position itself. In AirPlayer the mobile application is responsible for collecting the RSSI measurements, from each available AirPort Express, and use them to determine which zone it is currently in. Having the smartphone application as a self-positioning receiver makes it less dependent on changes in the infrastructure and therefore does not need to be aware of details about the setup such as the number and location of available AirPort Express stations.

4.3   Server Application

The server application is the backend of the system and is basically a remote controlled music player capable of streaming wirelessly to speakers connected to an AirPort Express. It contains information about the user's iTunes music collection including playlists. Actual management of the music, like editing playlists or reorganising the music, is handled directly in iTunes. This provides a familiar interface for the user and results in a simpler implementation on the server side of the application.

The server application is furthermore responsible for registering and managing the configured zones. It makes use of Apple's implementation of zero configuration networking, called Bonjour, which provides a DNS based Service Discovery API that enables automatic discovery of network services. This means that every time an AirPort Express is connected to the wireless network, it will be discovered automatically by AirPlayer. The default behaviour is for each newly connected AirPort Express to create its own zone, but as mentioned, individual zones can be linked to form larger ones. This is managed in the server application by letting one zone handle a number of AirPort Express connections, synchronizing the control and stream of music.

5   Field Evaluation

We conducted a field evaluation of AirPlayer with the goal of investigating people's use of proxemics and spatial interactions, and exploring their cognitive perception of such system and interaction in use.

5.1   Method

The field evaluation spanned three weeks where AirPlayer was installed and integrated with participants' existing multi-room music system at home. The participants were introduced to the system in the beginning of the evaluation and then used it in their everyday lives for the duration of the study. Participants were asked to note their thoughts about the system in a diary. After the three weeks a semi-structured interview was conducted, following the guidelines of Lazar et al. [16]. Entries from the diaries were included as a basis for a conversation regarding the system. Interviews were conducted in the homes of the participants, to provide a comfortable environment and to let the participants talk about their experiences in the place where interaction took place.

Two interviewers were present at each session. One would specifically be responsible for taking notes including information about esoteric remarks, visual references etc. Interviews were furthermore recorded.

5.2   Participants

Two households, which will be referred to as A, and B, participated in the field evaluation. Both had a multi-room music system installed prior to the evaluation and were therefore familiar with the basic concept and use of such systems. The participants received a small token of appreciation for participating in the evaluation.

Household A. The first participating household had two residents. A woman aged 47 (A1) living with her 16 year old son (A2) on a small farm. The music player currently installed in their home was a Sonos system. The woman had four years of experience with the system and her son one year. Three AirPort Express stations were installed in the locations where the participants usually listened to music using their current system as illustrated in Fig. 9.
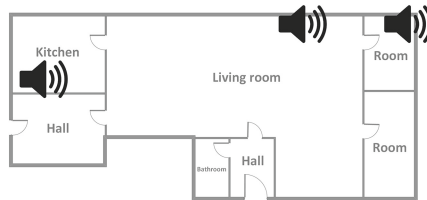

Fig. 9. Floor plan of household A illustrating placement of AirPort Express stations. The floor plan only includes the part of the farm where AirPlayer was installed.

Household B. The second participant was a 28 year old man living alone in an apartment (B1). He already had a setup consisting of Apple products and used iTunes to play music. AirPlayer could therefore easily connect to the existing system. An additional AirPort Express was added to get more than two locations. The floor plan is shown in Fig. 10.
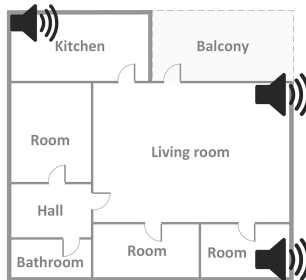

Fig. 10. Floor plan of household B illustrating placement of AirPort Express stations.

6   Findings

Basic usage was logged by the system itself, which will be presented to give an overview of the foundation for the qualitative findings.

Over the three-week period household A used the system for 70 h in total, averaging 3.3 h per day. The system would however be used more in the first half of the evaluation period. During the 70 h of use, the location and movement features were active for 57 % and 23 % respectively. The usage in household B had a more even distribution with a total of 43 h, averaging 2.0 h per day. The location feature was active 47 % of the time and the movement feature 39 %. Reasons for the relatively low percentages of time the features were active can partly be found in the findings presented later. There are however also a few low-practical explanations. One is the location feature being implicitly active when the movement feature is active, but not counted in the statistics if it is not explicitly activated in the system. Another is the fact that participants would turn features on one at a time to explore this particular feature.

From this usage, participants were able to continuously report experiences in their diary and conclusively report on their use of the system during the interview. The following sections describe our findings, structured by the two different proxemic interaction features, Location and Movement.

6.1   Location

The location feature allows the smartphone application's user interface to automatically adapt to the zone the user is currently in. Interaction is thereby directed towards the current spatially location of the user, and visual feedback on the smartphone reflects what is playing in that particular location. From the interviews two main findings in relation to this type of spatial interaction design were identified, specifically regarding simple interaction and local interaction.

Simple Interaction. The purpose of the proxemics features of AirPlayer is to hide the complexity of interaction with the music system by introducing a layer of spatial interaction. Both households described the location feature a very simple way of interacting "spatially" with the system. They perceived this form of interaction with the music player as intuitive, but also as very transparent. In fact, A1 expressed that she was not even always consciously aware that she was using it:

> *To me it had to be fantastic, because I used it and I didn't even think about whether or not the feature was enabled. I actually thought it was connected with the other [feature] where music followed me from room to room. It worked when I adjusted the volume and other things. But I had not thought about the fact that I used it, but I really did. (A1)*

The fact that she was not always consciously aware that she had enabled the location feature, despite of not having this function available in her existing Sonos system, indicates that the feature was very subtle, and that automatically adapting the interface of a smartphone application to the user's immediate spatial surroundings matches the users' spatial cognition in relation to that place and that type of interaction. This is especially evident since the location feature was in fact active 57 % of the time in

household A, and also implicitly active when the movement feature was enabled as well. Looking back at her interaction s with the system, the user of household A understood immediately how the location feature worked, and that interacting spatially in her home, by moving from one room to another had a direct effect on the user interface. She explicitly thought of it as a "simple" form of interaction to be able to use her smartphone to control the music in the room she was currently in – just as she was doing when the movement feature was active. In household B, the participant similarly experienced the location feature to be simple. Compared to household A, however, he was consciously aware that he was using this particular feature:

> As I said earlier, I think that a relatively simple feature as this one is extremely good. The fact that I do not have to find the room that I am about to play music in, makes it easy to utilise the mobile phone to control the music according to your current location. (B1)

Another point where the need for simple interaction was visible was not in the added proxemic features but instead in features missing. Participants from both households commented that they were missing the radio or Internet radio functionality of the systems they were used to. B1 also mentioned that he would have liked a larger queue size. Both of these comments indicate a need for simple interaction where it is not up to the user to constantly control the music in several locations.

Local Interaction. What we experienced from the field evaluation was that the location feature, which limits interaction to the immediate spatial surroundings of the user, was greatly appreciated by the participants for its spatially localised interaction. By enabling the location feature, they found it easy to use the application as a localised remote control for the music player, as the application was aware of its current location and could therefore easily manage the music in the given zone. Both households used the location feature during the majority of the test period. As the participant from household B expresses:

> I would estimate that I just used it. I just used it most of the time. Actually, I have had no real need to control another room other than the one I was present in. (B1)

This is also supported by the previous comments from A1 who used the location functionality without being consciously aware of it, because she experienced that it just "naturally responded nearby" whenever she interacted with it.

6.2   Movement

The second proxemic dimension, movement, was also implemented as a part of the interaction design of AirPlayer. When active, the movement feature allows music to follow the user around the house additionally moving the queue and control of music as well. From the interviews, two main findings in relation to this type of spatial interaction design were identified, specifically regarding foreground and background interaction and music in discrete locations.

Foreground and Background Interactions. None of the participants initially expressed any uncertainties regarding the concept or functionality of AirPlayer. When asked to describe the system's procemic and spatial interaction features, using their

own words, they were all able to give a brief and correct overall description of how it actually worked, suggesting that the implemented interaction design matched the users' spatial cognition in relation to that specific physical place very well. However, B1 also elaborated:

> *Of course, I had to get used to how it worked the first time I used it, but I do not think that we had any doubts of how it was supposed to work. (B1)*

Both households had a good overall perception of how the different features worked spatially. Despite that, however, we did find examples where the details of the Air-Player system's behaviour was not clear:

> *I would have thought that the music stopped when I went upstairs to bed. It didn't. It stayed in the last room I left – That puzzled me. (A1)*

Because the music "followed" A1 into rooms where a speaker was present, as in started playing when entering a room, she also expected the music to stop playing when she left one of those rooms - or at least when she left the floor of the house where that particular room was located. This was not the case in the current implementation of AirPlayer, leading to some surprise and uncertainty. From a spatial cognition perspective, the issue at play here is simply that the system does not consistently behave naturally, but has different behaviour depending of whether the user enters or leaves a room. If the spatial experience of music playback "following" the user was to be realised fully, music should, of course, also stop playing when leaving a particular room. Alternatively, the spatial experience created would be one of "leaving a trail" of music rather than the desired one of "taking it with you". In relation to this issue, B1 stated that he actually thought that the music had stopped playing in the rooms he had been to and then left. But because he had left the room, he was unaware that the music was in fact still playing.

Music in Discrete Zones. Interesting findings also emerged regarding the relationship between music and the spatial subdivision of the house into discrete zones. In particular, both families found that the movement feature worked well. As one of the participants from household B stated:

> *We have used it a lot, partly because we wanted to test the system, but also because we found it to be clever and fun to use. The thing about having the music following you is nice when you walk around at home in your own thoughts. (B1)*

Household A shared this opinion towards using the feature; however they also experienced a slightly unintended behaviour. When asking how the feature behaved, they explained that the music should overlap for longer periods than it currently did:

> *It could have waited 15 s before it stopped the room you were leaving, and then start the music immediately as you walk into the next room. I think more overlap would have worked wonders (A1)*

As participant A1 stated, she would have liked AirPlayer to have a timer-based threshold instead of having one purely based on distance between two zones and the smartphone. The result of the behaviour was that she would turn off the feature while cleaning the house, as she would move frequently between different rooms.

At times, the participants also experienced that the music would not be entirely synchronised, which were obvious when standing in the middle of two different zones with the movement feature active. The result was a slightly asynchronous playback of the same song:

> *As I said earlier, there were a few incidents where the music was not entirely synchronised, but we learned that we could fix it by changing the track being played and then switching back (B1)*

The asynchronous behaviour was caused by technical limitations discovered in the preliminary study when the framework was developed.

7   Discussion

Several systems previously presented within research on proxemics and spatial interactions are based on a central visual focal point, like a large public display, where interaction takes place in the immediate surroundings. Complementing this research, the AirPlayer prototype has been developed to explore proxemic interactions in environments that span separate locations, and where the primary medium is not visual but auditory. Our focus has furthermore been on exploring specific spatial dimensions of proxemic interaction, namely location and movement, and to understand better the effects of such spatial interaction in real world use. In this section we will discuss further the cognitive perception of the proxemics and spatial interaction with AirPlayer in people's homes.

7.1   Simple Interaction

One of the things that stood out from our field evaluation was that proxemics and spatial interaction provided a means to facilitate simple interaction with the multi-room music system. The point of simplicity in this case is not only that the interaction is simple to understand and use, but also that the proxemic and spatial features reduce the complexity of what is in fact a relatively complicated setup of distributed speakers and wireless media players. The fact that one of the participants (A1) was unaware that she had used the location feature, but still found it useful when reflecting on it in hindsight, points out a particular positive potential of proxemics and spatial interactions when they are designed and implemented well. Interaction can be very transparent to the user as long as it feels natural, and letting features become invisible to the user can even be a success criterion of the design. Our findings about the use of the location feature did not indicate importance of being informed every time control shifted from one spatial zone to another, or that people needed advanced features to feel more in control of the system at their spatial location. One reason can be that the metaphor of a single music system in each room simply becomes stronger with the location feature active, and that it then just feels natural to use such a system as a remote control for the speakers in people's immediate spatial surroundings.

7.2   Local Interaction

The main differences between multi-room music systems and traditional music players is that the music source and the control can be centralised when playing music in different spatial locations of the home. What the location feature of AirPlayer does is really to automatically limit the control to the room the user is currently in. During the field evaluation of AirPlayer, this feature was well received, and it was found to be very intuitive to control music "locally" in spatial vicinity of the user. While integration of wireless networks in our homes provide great opportunities for remotely controlling everything from everywhere, our findings point toward positive aspects of doing the opposite, and use spatial positioning to localise control to only those systems present in the users immediate vicinity. By doing so, we found that people very quickly perceive control systems, for example an application on a smartphone, as extensions of the interaction possibilities in their spatial location, and very quickly understand that the system changes when they move between physical locations. This observation confirms previous work on context-awareness, which has, for example, described spatially dependent mobile systems as dynamic indexical signs, which people are generally highly capable of interpreting [14]. This finding also contributes to the argument that discrete zones with localised interaction can be powerful for the user in order to understand where interactions are directed [20].

7.3   Foreground and Background Interactions

Buxton talks about foreground and background interactions within HCI [5]. Where the graphical user interface of the smartphone application facilitates foreground interaction, proxemic and spatial interactions facilitate background interactions. Making the user understand the intended design of background interactions can be rather challenging exactly because they are supposed to be more discrete. What the use of the movement feature in AirPlayer showed us was that although the overall concept was easy to understand, the functional details about the interaction were difficult to communicate to the user. One participant knew that the movement feature allowed music to follow her around and therefore also expected it to stop when she went to a room where music could not be played. As a consequence of making an effort to simplify the interaction, in our case through proxemic and spatial interactions, challenges in making details about this interaction design apparent to the user were unintentionally introduced. In our specific case this was a particularly problematic challenge because the activity of listening to music is already typically something that takes place in the background.

7.4   Music in Discrete Zones

Certain issues and unexpected behaviour were experienced by the participants in relation to the movement feature. In response to the way music moved between zones, A1 preferred not to use the movement feature while cleaning the house, as she would be moving frequently between rooms. The system behaviour she experienced might be due to the specific layout of the house, but it does raise an interesting issue in relation to

the combination of spatial interaction and auditory output. One of the reasons for the annoyance could be because the speakers, often placed in opposite sides of adjacent rooms, created the experience of music jumping from one side of the house to the other, instead of creating a seamless experience. Another reason is possibly that while the location system is divided into discrete zones, music is not restricted to be audible within a specific zone but easily travels across zones, even through walls. Because of this, when standing between two zones, users also sometimes experienced that playback from different sets of speakers was slightly out of sync. While this may be avoided by implementing the system slightly differently in order to achieve better synchronisation, the fact that sound travels differently to visual media presents some fundamental challenges to proxemics and spatial interaction with such systems.

7.5   Cognitive Effects of Spatial Interaction with AirPlayer

Our study has showed that deploying an auditory media system with proxemic and spatial interaction into people's homes has an effect on how people cognitively experience the relationship between such system and the physical space where it is being used. It is clear that the layout of the physical space influences people's experience of the system, and their expectations about how it should behave. But it is also clear that the use of the system conversely influences people's experience of the physical space it is being used in.

Starting with the effect of physical space on expected system behaviour, we saw very clearly that people readily experienced the localised user interface, and the music being played in a room, as just another natural property of a particular space, like its physical shape, colour, lighting, etc. Hence, it was also readily expected that the user interface would change when moving to another spatial location, and that when in another spatial location one would be able to control the music playing there. This leads us to conclude that there is a significant cognitive effect of people's placed-ness in the physical world on their perception of interactive digital systems that respond to proxemic and spatial relationships and changes in their physical surroundings.

In terms of the effect of systems with proxemic and spatial interaction on the experience of the physical space it is being used in, we saw a change in people's spatial experience of their homes in the sense that some physically separated locations were suddenly less separated because sound triggered by the interaction with the system would now flow between rooms. On the positive side this meant that previously disconnected physical spaces could now be experienced as connected ones, contributing to an experience of wholeness when moving through the house. On the negative side, however, it also meant that those spaces that people wanted to keep separated could be more difficult to experience as such. This was specifically the case when separate music sources sometimes interfered with each other across rooms, but it was also the case when people experienced that they were unable to go into a silent room, for example to answer a phone call, without the music "chasing" them. This exemplifies an unwanted cognitive effect of perceiving the house, rather than the inhabitants, being in charge of how the physical space is used, and how it can be negotiated. In relation to this, we find it very important that systems with proxemic and spatial interaction do not take control away from the user.

7.6   Technical Limitations of the Spatial Tracking

We are aware of the limitation imposed by the direct use of RSSI values as indication of physical distance, due to uncontrollable external influences. A more accurate indoor positioning system is however out of the scope of this paper and the focus has been on the concept of proxemic interactions in real-life contexts. What is interesting to note is that even though participants were asked about the usefulness of the proxemic interaction features, no comments were directly addressing problems with accuracy of the location estimation, despite the simple implementation. Other issues have surfaced during the interviews, but an overall perception of proxemic awareness and satisfaction with the features have been experienced by the participants.

8   Conclusion

We have explored people's experiences of spatial interaction with a multi-room music system by developing a functional prototype and deploying it in actual households. This has enabled us to investigate the effects of two specific dimensions of proxemics and spatial interaction, namely location and movement. The study yielded a number of specific findings related to simple interaction, local interaction, foreground and background interaction and music in discrete zones. The findings suggest that proxemics and spatial interactions have a great potential for "hiding" parts of the user interaction with a complex system and thereby making it simpler for people to use. However, we also found that it is challenging to simplify background interactions in a way where details about the functionality and possibilities for interaction are still apparent to the user. We also identified challenges in relation to combining the specific medium of music with spatial interactions, specifically related to the fact that audio may travel across the physical boundaries delimitating our spatial surroundings.

References

1. Ackad, C.J., Clayphan, A., Maldonado, R.M., Kay, J.: Seamless and continuous user identification for interactive tabletops using personal device handshaking and body tracking. In: Proceedings of CHI EA 2012, pp. 1775–1780. ACM Press (2012)
2. Annett, M., Grossman, T., Wigdor, D., Fitzmaurice, G.: Medusa: a proximity-aware multi-touch tabletop. In: Proceedings of UIST 2011, pp. 337–346. ACM Press (2011)
3. Ballendat, T., Marquardt, N., Greenberg, S.: Proxemic interaction: designing for a proximity and orientation-aware environment. In: Proceedings of ITS 2010, pp. 121–130. ACM Press (2010)
4. Brumitt, B., Meyers, B., Krumm, J., Kern, A., Shafer, S.: Easyliving: technologies for intelligent environments. In: Thomas, P., Gellersen, H.W. (eds.) HUC 2000. LNCS, vol. 1927, pp. 12–29. Springer, Heidelberg (2000)

5. Buxton, B.: Integrating the periphery and context: a new model of telematics. In: Proceedings of GI 1995, pp. 239–246. Canadian Information Processing Society (1995)

6. Chang, J.-W., Kim, Y.-K.: Context-aware application system for music playing services. In: Gabrys, B., Howlett, R.J., Jain, L.C. (eds.) KES 2006. LNCS (LNAI), vol. 4253, pp. 76–83. Springer, Heidelberg (2006)

7. Clark, A., Dünser, A., Billinghurst, M., Piumsomboon, T., Altimira, D.: Seamless interaction in space. In: Proceedings of OZCHI 2011, pp. 88–97. ACM Press (2011)

8. Drane, C., Macnaughtan, M., Scott, C.: Positioning GSM telephones. IEEE Commun. Mag. **36**(4), 46–54, 59 (1998)

9. Evans, V., Chilton, P.: Language, Cognition and Space: The State of the Art and New Directions. Equinox Publishing, London (2010)

10. Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., Wang, M.: Proxemic interactions: the new ubicomp? Interactions **18**(1), 42–50 (2011). ACM Press

11. Hall, E.T.: The Hidden Dimension. Doubleday, Garden City (1966)

12. Holmquist, L.E.: Ubiquitous music. Interactions **12**(4), 71 (2005)

13. Ju, W., Lee, B.A., Klemmer, S.R.: Range: exploring implicit interaction through electronic whiteboard design. In: Proceedings of CSCW 2008, pp. 17–26. ACM Press (2008)

14. Kjeldskov J., Paay J.: Indexicality: understanding mobile human-computer interaction in context. ACM TOCHI 17(4), 14 pp. (2010)

15. Laibowitz, M., Gips, J., Aylward, R., Pentland, A., Paradiso, J.A.: A sensor network for social dynamics. In: Proceedings of ISPN 2006, pp. 483–491. ACM Press (2006)

16. Lazar, J., Heidi, J., Hochheiser, H.: Research Methods in Human-Computer Interaction. Wiley, Chichester (2010)

17. Liikkanen, L., Amos, C., Cunningham, S.J., Downie, J.S., McDonald, D.: Music interaction research in HCI: let's get the band back together. In: Proceedings of CHI EA 2012, pp. 1119–1122. ACM Press (2012)

18. Marquardt, N., Diaz-Marino, R., Boring, S., Greenberg, S.: The proximity toolkit: prototyping proxemic interactions in ubiquitous computing ecologies. In: Proceedings of UIST 2011, pp. 315–326. ACM Press (2011)

19. Marquardt, N., Greenberg, S.: Informing the design of proxemic interactions. IEEE Pervasive Computing **11**(2), 14–23 (2012)

20. Marquardt, N., Hinckley, K., Greenberg, S.: Cross-device interaction via micro-mobility and f-formations. In: Proceedings of UIST 2012, pp. 13–22. ACM Press (2012)

21. Montello, D.R.: Spatial cognition. In: Smelser, N.J., Baltes, P.B. (eds.) International Encyclopedia of the Social & Behavioral Sciences, pp. 14771–14775. Pergamon Press, Oxford (2001)

22. Nirjon, S., Dickerson, R.F., Li, Q., Asare, P., Stankovic, J.A., Hong, D., Zhang, B., Jiang, X., Shen, G., Zhao, F.: MusicalHeart: a hearty way of listening to music. In: Proceedings of SenSys 2012, pp. 43–56. ACM press (2012)

23. Nova, N.: Socio-cognitive functions of space in collaborative settings: a literature review about Space, Cognition and Collaboration. CRAFT Research Report No. 1 (2003)

24. Rose, M.: Music in the home: interfaces for music appliances. Pers. Technol. **4**(1), 45–53 (2000). Springer

25. Tolman, E.C.: Cognitive maps in rats and men. Psychol. Rev. **55**, 189–208 (1948)

26. Torres-Solis, J., Falk, T.H., Chau, T.: A review of indoor localization technologies: towards navigational assistance for topographical disorientation. In: Jesus, F., Molina, V., (eds.) Ambient Intelligence (2010)

27. Trotto, A., Tittarelli, M.: Musical viruses for graceful seduction. In: Proceedings of NordiCHI 2012, pp. 731–735. ACM Press (2012)

28. Vogel, D., Balakrishnan, R.: Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In: Proceedings of UIST 2004, pp. 137–146. ACM Press (2004)
29. Wang, M., Boring, S., Greenberg, S.: Proxemic peddler: a public advertising display that captures and preserves the attention of a passerby. In: Proceedings of PerDis 2012, ACM Press, Article no. 3 (2012)
30. Want, R., Hopper, A., Falcão, V., Gibbons, J.: The active badge location system. ACM Trans. Inf. Syst. **10**(1), 91–102 (1992)