



Navigating by Sound in My SmartCity

Romain Barthélémy and Roland Cahen

Wi: Journal of Mobile Media 2015 9: 02

The online version of this article can be found at:

<http://wi.mobilities.ca/romain-barthelemy-and-roland-cahen-navigating-by-sound-in-my-smartcity>

[reference]

[Barthélémy, Romain and Roland and Cahen. “Navigating by Sound in My SmartCity” . *Wi: Journal of Mobile Media*. 09.02 (2014). Web.]

Navigating by Sound in My SmartCity

Romain Barthélémy and Roland Cahen

Abstract

SmartCity+ consists of an immaterial layer to be added to a real city, comprised of a map and a timeline. It is meant to allow people, services and located information to be expressed in real time and, more specifically, through the use of sound. This article is concerned with the phonic/auditory aspects of the user-experience. It seeks to explore an auditory enhancement of the urban experience on mobile devices.

Introduction

In its most basic form, *Smart City+* offers to supplement reality with services providing geographically located auditory notifications. In this version it is necessary for sounds that are heard to be significant for the user in order to make it easier for them to learn how it works. Therefore, we are mainly talking about extending the current phonic lexicon that is currently being used by existing platforms. It also required the development of new functionalities such as: the differentiation between different auditory notification, sound avatars and mediacenters. Certain services of the *Smart City+* platform have also developed ways of contextualising auditory events and notifications (audioguide, localisation of contacts). Finally, in order to facilitate sequence memorisation, it was necessary to develop audio-graphic expressions of both itineraries and the calendar. In its more prospective and extended form, *Smart City+* creates an immersive and auditory augmented reality in situations selected by the user, by making use of spatialised sounds. In order to address these needs, we have chosen the metaphor of the sonar: an activator that is adaptable to usage and to full listening mode, and a mode for continuous

activation for mobility.

Adding to our immersive experience of the city on mobile devices

Our ability to recognise a large number of sounds in our daily environment shows that the quantity of acoustic stimuli is not an obstacle to understanding our auditory environment. We have no problem interpreting the events that we recognize, especially when they possess distinct timbers, when they are spatially distinct or when they are associated to objects, actions and contexts that we are used to. When several sounds come from the same mobile device, they come from the same place and therefore give the illusion of having a single cause. How can we make sure that sound signals emitted by mobile devices will be easier to differentiate and not ignored or perceived as being a nuisance? How can we integrate them to the way we experience our daily space in the same way as external sounds?

Beyond crucial psycho-cognitive questions – notably those of attentional economy and sign-learning – our research intends to address how the acoustic expression of the immaterial aspects of the city can become a consideration in our daily interactions with the physical world.

Our hypothesis is that to be understood, sounds have to be not only distributed and immersive but also intimately associated with actions and usage.

Leave the sound on

In our daily lives, there are times and places where we need silence for concentration and rest; but at other times, like when we commute to and from work, we are drowned in the hubbub of sound. Navigating these different environments is made easier by the fact that physical space is categorised by separate functionalities and closeted off by doors and walls. In the virtual world,

structures are unstable and there are no doors.

Users of mobile devices often mute their machines, especially if they lean towards the noisier side of the spectrum. Sound is often considered – by default – to be a disturbance. We want to show that sound is not only useful, but indispensable. We could even go so far as saying that every sound is a disturbance in one circumstance and useful in another, which by no means facilitates the work of the sound designer. Our project here is to make sound tolerable when it is useful, and discreet or absent when it should be, to dissuade the user from disabling it. Our objective is to change the user's attitude from "shut-up" to "speak if it's important for me to listen to you". If he or she needs to mute the device, they should be incited to reactivate these functionalities afterwards. In order to achieve this we suggest 1) an intelligent use-based design aligning functionalities with their auditory expressions and 2) speedy access to the way in which the platform's sounds express themselves that is easily and quickly adaptable to use (c.f. SmartVolume or a Sound Density Selector)

Sound design of the user experience of a platform of services

The *Smart City+* platform is not meant to realise the services provided by the phone itself or duplicate existing services, but rather to integrate existing and independent services in an environment dedicated to hyperlocability. To this end, it is not about giving sound to service functionalities, but about offering a choice of auditory functionalities, available to partner services. The default, detailed settings set for each service and platform manager can be modified by the user in the platform preferences.

A forward-looking approach to phonic navigation

Smart City+ is first and foremost a map and, as such, it must allow a synthetic representation of the urban space in order to facilitate the search for useful information. With the rise of mobile terminals and geolocalisation, maps have become sorts of parallel cities, superimposed upon the real city. The mass of *in-situ* information habitually given by tangible sources such as street signs, surrounding sounds and human beings are supplemented or even replaced by a flow of contextualised data. In situations of mobility, auditory information is particularly useful given that our eyes are busy finding their way around, recognising obstacles, or controlling the direction of a vehicle. In these situations, our ears can supplement the visual modality to bring us information about non-visible aspects of navigation such as the danger presented by unseen vehicles, external events, or even interactions with people or with the environment.

Modelling interactive functions

For this project we principally used the following modelling tools: animation, video, MaxMSP and Unity 3D. MaxMSP allowed us to rapidly trial models at a very early stage of the project; whereas Unity was used to support the realisation of functional software units ready to be integrated and linked up with partner platforms.

Giving Sound to Mobility: Liberating the Visual

Specifying the phonic lexicon

The phonic lexicon of a digital interface must be concise and yet transmit a precise and intelligible message. Therefore, it is crucial that sounds express their function in a very short period of time and that they should be instantly and easily identifiable, regardless of the listening context. We privileged the creation of sounds whose essential characteristics could be

perceived in mono and in a limited bandwidth (between 200 and 7000 Hz). This facilitates their intelligibility on lower quality devices, such as integrated speakers. However, given that *Smart City+* is a multiplatform tool, sounds must also be adapted to higher-quality devices, such as headphones or Hi-Fi speakers: in stereo and with a larger bandwidth. We developed a phonic lexicon that would be less pleonastic for the listener and more informative when on the move. The number of items in the phonic lexicon has to be limited in order to make sure that the listener can remember them all. We looked for strategies to ensure that the size of the lexicon was developed in accordance with concerns about functionality and memorisation. Wherever possible, we chose to reinforce visual expressions with auditory ones and vice versa, so that users may access the information regardless of their modal sensory competence or the situation of use.

Personalising notifications

Thematic sound-function associations

To facilitate the appropriation of auditory notifications, we decided to let the user select an auditory universe for each known functionality (calendar, email etc.) For example, email could be expressed through the theme of birdcalls, boat-noises or circus-sounds. Notifications upon receiving new messages or for messages having been sent, delivered and read can also be selected from the same universe. The first stage in users' learning curve is to associate the universe (birdcalls) to the function (email). They can then go through subsequent stages such as learning to recognise the different types of notifications: urgent vs. non-urgent, reception of multiple vs. singular messages, etc. As soon as they have received the notification, and without having checked the screen, users can know how many messages they have received and how urgent they are.

Apposing suffixes to notifications: the example of the auditory avatar

The *auditory avatar* is like the auditory “profile picture” or “signature” of the user in *Smart City+*. Thanks to this *avatar*, users can communicate with and notify their contacts. The auditory avatar permits the enhancement of certain notifications (messages, alarms) by incorporating them. In certain situations it renders optional the currently automatic recurrence to the visual interface. For example, a calendar notification, accompanied by an auditory avatar, reminds users both of the imminence of the meeting and the person with whom they are meeting.

Users can personalise their avatars when they complete their profile or pick one by default. This morsel of sound is constructed around a short vocal element. Users are encouraged to record their own voice, but can also “chose their voice” amongst a panel of propositions according to age and gender (see Fig. 1). The selection offered is an ensemble of calls and interjections: they are universal in that they offer non-synthesised voice timbres. By doing this *Smart City+* seeks to offer auditory aesthetics and diversity as we experience them in real humanity.



Fig. 1: The recording interface offered is minimalist: record, select, play.

(Legend: “Chose a voice OR record your own”).

A second stage in the personalisation process allows the user to select a genre (techno, classical,

humour...), made of three short sounds that complement the voice, which insures good audibility in case of poor recording quality or listening conditions.

Indexing auditory POI (Points Of Interest).

Auditory icons¹

Smart City+ is organised around the user's search for local POI. POI can represent local shops, friends nearby or even services provided by local government. POI are classified by category: each POI has its own auditory icon (see Audio 1).

[Audio 1: Examples of auditory icons for "Restaurant", "Sport" and "Nature"](#)

The icon identifies each message notification, alarm, the opening of a *mediacenter* (a popup window containing information about the POI) or the activation of a POI thanks to navigation tools such as the *sonar* (described further). POI categories are exclusive, meaning that a POI cannot be marked as belonging to two different categories. These auditory icons are created from short elements so as to rapidly dispense information. They illustrate the category they represent (child's voice for childcare, musical trait for cultural activities), in order to speed up the learning process and aid memorisation. The POI categories have an auditory grammar, which has been voluntarily simplified so as not to confuse the user.

Learning through use

Even though we have taken care to create iconic sounds, we are aware that there is a part of

¹ Auditory icons are documented in Hermann, Hunt and Neuhoff 2011.

subjectivity implicated in causal listening.² So we are hoping that the context will aid in the learning process; also, there will only be a limited set of sounds to memorise. The auditory icons are audible at the instantiation of each POI, during notifications but also when POI are manually selected on the map. The multiplication of contexts of apparition in the interface allows users to memorise the significance of the auditory icon.

Prospective work on spatial and temporal sequences

A prospective piece of work was carried out better to memorise itineraries and the calendar by using sound synchronised with animation. The aim of this work was to encourage the user to recur to the visual interface less often when they have forgotten the information. The *Smart City+* platform provides a service for itinerary planification in public transport, completed by a pre-visualisation and/or pre-audition functionality. An audio-graphic animation indicates the stages of the itinerary and offers a preview of the transit (see Fig. 3 and Video 1).



Fig. 3: Itinerary + audioguide



Fig. 4: Timeline + alerts

² Causal listening is documented in Schaeffer 1966. It is the natural listening reflex that tries to resolve perceptual issues by recognizing the cause of a sound (as opposed to *semantic* and musical or *reduced* listening). See auditory categorisation of sound sources in Houix 2003.

The visual mode gives spatial, textual and iconic indications: location of starting and finishing points, changes, length of commute, line numbers, etc. whereas the sound illustrates segments of each mode of transport (tram, bus, on foot, train, bicycle, metro), and highlights the stages (departure, change, arrival). The audio-visual bi-modality should permit a faster memorisation for the itinerary. This system has also been modelled for the agenda (see Fig. 4).

[Video 1: A preview of an itinerary through auditory icons](#)

Audioguiding and localised notifications

On the go, the same system operates for in transit navigation, either during commuting time or throughout the day. It then functions like an audioguide: either in the form of contextualised notifications during transit and calendar updates as time passes or by activating POI in our immediate environment during our commutes, which we will refer to as the “auditory navigation” or “full listening mode”.

Approaches to Auditory Augmented Reality

Auditory navigation

This project uses the concepts of *auditory navigation* and of “*topophonie*” developed in the previous projects ICARE, PHASE ^(Ircam 2003-2005), ENIGMES (ENSCI 2006) and TOPOPHONIES (ENSCI 2009-2012). A “*topophonie*” is a space, a geometry or an architecture of sounds, that is to say the structure of an ensemble of distributed sound sources (in a real, virtual or augmented space). *Auditory navigation* is the experience of the visitor in “*topophonies*”: that is to say, the auditory or musical sequences resulting from the movement or actions of the listener. The sound sources are activated – that is to say triggered and controlled by what we call *activation modes*.

For example, the most frequent activation modes used by mouse navigation are *click* and *roll*.

Sonification of POI (Points of Interest) and RTE (Real Time Events)

We can distinguish two categories of auditory elements in spatial or temporal representation that can be expressed during navigation on the timeline or on the map:

- POI are static points of interest (e.g. restaurant) or dynamic ones (e.g. contact) which are heard only when users request that they should be, for example when they search for restaurants close by.
- RTE are punctual elements in real-time which are expressed at the very time they happen, for example when a parking spot becomes free.

Smart City+ offers a collection of activation modalities of the sound sources (POI) adapted to the usage of the platform:

1) Activities linked to selection and search

Individual triggering of POI upon selection or along with search results.

2) *sonar* and *torch* mode (see Video 2)

The *sonar* function triggers POI present in a certain radius by a circular probe activated by a quickly touching the screen or holding it over the user's position. It looks like a ring or bubble around the user.

Torch mode is activated directionally by pointing the device or by manually aligning direction and distance on the screen.

[Video 2: Model of the sonar, torchlight mode \(directional activation\)](#)

3) The *full listening mode*.

Full listening mode

The *full listening mode* is an immersive mode. It corresponds to a sort of auditory “Street View”. It can easily be activated while in motion by zooming in on the map (up to a scale of screen/20m) or by double tapping on the on-screen avatar. This is the visual and functional equivalent of a permanently active sonar around the user’s avatar: by moving, users activate the POI which enter their bubbles (see Fig. 3).

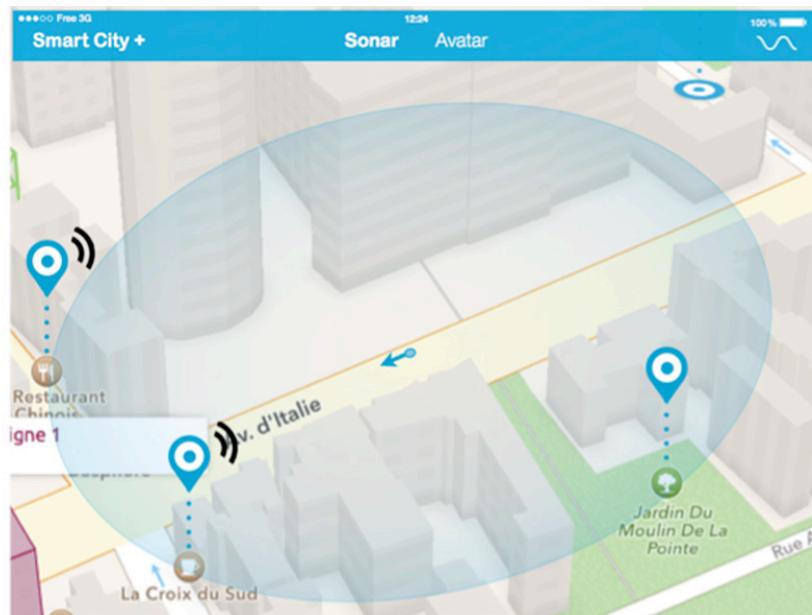


Fig. 3: By walking around in full listening mode, users activate the POI that enters their bubble.

The full listening mode is made of three auditory components. An immersive sound (a city ambiance) is activated with the “Map view” and its volume varies progressively according to the zoom level in order to give users the impression that they are “entering” the map. A drone

(continuous sound) present throughout full listening mode informs users about the activation of their sonar around them. The sounds of POI met along the way are triggered when they enter the sonar perimeter.

Access to User-Preferences

Taking listening contexts into account

The transmedia vocation of *Smart City+*, as well as the mobility it requires, make it necessary for us to take into account the heterogeneity of contexts in which it is used. Auditory needs and the disturbance sound creates vary enormously according to location (at home or on the street), the listening device (headphones or speakers integrated to the device) and the level of urgency of the search for information. Users must be able to easily access a level of preference-management to adapt sounds to their usages.

Classification of acoustic elements at the interface

In order to provide an efficient auditory manager, we had to decide for each of the sounds how useful it is to users. Acoustic events fulfil different functions in a visual interface. Some of them complement visual functionalities (typing sounds), while others dispense specific information (sonar). Therefore they do not all have the same informational value.

We have established a classification of three types of acoustic functionalities:

- 1) The purpose of *interface sounds* is to give an immediate auditory return upon an action performed on the device. They allow a more precise, faster manipulation that reduces visual

attention and gives information about the interpretation of commands (e.g. validation or failure). They possess strong visual equivalences which users are familiar with.

2) *Auditory notifications* are important events. They are exterior to the user, in the sense that they do not directly depend on his or her actions. They aim to rapidly inform users regarding something which requires their attention in the application (e.g. reception of new messages or calendar alerts) at times when they do not necessarily have access to the visual interface.

3) *Advanced auditory functionalities* are auditory services which require a specific initialisation in the application. That is to say that the user has to activate them (e.g. using the sonar to scan surrounding POI or using the audioguiding). These functionalities have been specially designed for *Smart City+* and do not yet exist elsewhere (pre-audition of the day's program, pre-audition of a transit route, in transit audioguiding, sonar, auditory avatar). They need to be easy for users to suspend at times when they only require the visual information.

Designing SmartVolume

From this classification, we have imagined several interface propositions. Our aim in this project was not to design the visual interface but to find a way of accessing an auditory interface. For each proposal, the problem was not to finalise a graphic element, but to introduce different ways of managing the density of auditory demands in *Smart City+*. We need to allow users rapidly to update their preferences. Simplicity is a key factor to ensure that the sound does not get muted by default. However, this requirement goes beyond adding a new volume dial on top of those that already exist in all currently available devices.

Listening modes

A selection of listening modes is already present on numerous mobile devices to manage

ringtones, alerts and vibrating notifications. Most often the activation or deactivation of acoustic events is binary: auditory mode and silent mode. Here, four modes are offered to allow the user to vary the auditory density and the nature of audible sounds in function with predetermined needs (see Fig. 4).



Fig. 4: User audio preferences: selecting an audio mode.

The chosen nomenclature corresponds to contexts of use that are comprehensible to the user (public place, soundlover) rather than being descriptive of the “level of sound”. We even leave open the possibility for the user to personalise one of the modes for it better to correspond to their own use. Additionally, the concentric disposition should diminish the gradation effect, and therefore counteract the assumption of growing importance implied by a linear disposition.

SmartVolume

SmartVolume (Fig. 5) is a selector of auditory density. It is a simple cursor which allows the user to easily adjust the density of auditory events produced by the interface, according to their level of importance. It is determined by default for each event, thanks to attributes (called *priorities*) that can be varied between 0. and 1. *SmartVolume* thus functions like a filter with a variable threshold to mute or play sounds according to their level of priority. Users can define the levels

of priority for each service offered by the platform in order to compose their own SmartVolume. SmartVolume is presented as a visual interface allowing the user to control the sound according to their needs.



Fig. 5: User audio preferences: moving the SmartVolume changes the audio density of the interface.

Conclusion

This project addresses fundamental questions concerning the use of sound in augmented reality: the assumed nuisance, the need for attentional economy, the difficult question of how to make easily memorable sounds and the more general problem of ergonomics. Even though they are far from being resolved, this project offers concrete tools allowing us to project them into real situations.

We have given the *Smart City+* platform a functional auditory interface where sound is more than a simple counterpart to the visual, it completes it. Thus, sound becomes a tool and a

semantic agent in its own right. User-tests of the application still need to be carried out in order to adjust the priorities and to improve the identification of the auditory icons, as well as to address the ergonomic issue.

References

ENSCI. 2006. "ENIGMES project." Accessed March 23, 2014.
<http://projetenigmes.free.fr/wiki/index.php?title=Accueil>

ENSCI. 2009-2012. "Topophonie project." Accessed March 23, 2014.
http://www.ensci.com/fileadmin/_migrated/content_uploads/topophonie_research_project.pdf

Hermann, Thomas, Andy Hunt and John G. Neuhoff. eds. 2011. *The Sonification Handbook* Berlin: Logos Publishing House. Accessed March 23, 2014. <http://sonification.de/handbook/>

Houix, Olivier. 2003. *Catégorisation auditive des sources sonores*. PhD diss., Université du Maine. Accessed March 23, 2014. <http://articles.ircam.fr/textes/Houix03f/index.pdf>

Ircam. 2003-2005. "Phase project." Accessed March 23, 2014.
<http://recherche.ircam.fr/equipes/analyse-synthese/lambert/phase/presentation/08equipe-ircam.html>, <http://roland.cahen.pagesperso-orange.fr/PHASE/PHASE.html>

Schaeffer, Pierre. 1966, *Traité des Objets Musicaux*. Paris: Edition du Seuil.

Biographies

Romain Barthélémy: young composer and industrial sound designer. Master's Degree in industrial sound design in 2013 (barthelemyds@gmail.com).

Roland Cahen: Electroacoustic music composer, sound designer, artistic professor and researcher, head of the sound design studio at l'ENSCI-les Ateliers, responsible for the project.

roland.cahen@ensci.com)

THE SMART CITY + PROJECT: Navidis, Altran, ENSCi-les Ateliers, ESRI France, Grand Paris Seine Ouest, Issymedia, Le Cube, TelecomSudParis.

The *Smart City+* project aims to define and develop a digital platform for the aggregation and distribution of content and of local services for citizens, collectivities and local economic, social or cultural agents. *Smart City+* builds on modes of representation and of urban data valorisation in real time 3D environments, synced to the “Cloud”. Its objective is to inform, interact with and develop local economic activity. It is inscribed in the SOLOMO universe (SOcial / LOcal / MObile) with a multiplatform transmedia approach. www.smartcityplus.com/

ENSCI

Project workshop “Villes trans-apparentes” on mobile interfaces (gloss: Cities you can see through).

From January to June 2013, this project was carried out at L’ENSCI-Les Ateliers with twelve students of industrial design and led by designers Stéphane Villars and Patrick de Glo De Besse. During a whole semester, they spent three and a half days a week working on proposals for services and interfaces for the platform.

From September 2013 to January 2014, the experimental studio Sound Design *Smart City+* assembled a working group led by sound designer Roland Cahen. www.ensci.com

Acknowledgements

This article would not have been possible without generous contributions by Jonathan Tanant, Aurélien Marty, Florian Behejohn and the students of the experimental *Smart City+* Studio at

l'ENSCI (Mayu AGARIE, Arto KUUSISTO, Mathieu EYMELOUD, Camille JEGO, Jean SENEAL,
Tanita KLEIN, Zifan WANG).