Exploring the requirements for multimodal interaction for mobile devices in an end-to-end journey context

Claudia Krehl[a,*] and Sarah Sharples[b]

[a] *Horizon Doctoral Training Centre, University of Nottingham, Human Factors Research Group, Innovative Technology Research Centre, University Park, NG7 2RD, Nottingham, United Kingdom*
[b] *Department of Mechanical, Materials and Manufacturing Engineering ,University of Nottingham, Coates Building, University Park, NG7 2RD, Nottingham, United Kingdom*

Abstract. The paper investigates the requirements for multimodal interaction on mobile devices in an end-to-end journey context. Traditional interfaces are deemed cumbersome and inefficient for exchanging information with the user. Multimodal interaction provides a different user-centred approach allowing for more natural and intuitive interaction between humans and computers. It is especially suitable for mobile interaction as it can overcome additional constraints including small screens, awkward keypads, and continuously changing settings - an inherent property of mobility. This paper is based on end-to-end journeys where users encounter several contexts during their journeys. Interviews and focus groups explore the requirements for multimodal interaction design for mobile devices by examining journey stages and identifying the users' information needs and sources. Findings suggest that multimodal communication is crucial when users multitask. Choosing suitable modalities depend on user context, characteristics and tasks.

Keywords: Requirements Analysis, Multimodal Human-Computer Interaction, Interaction Design, Mobile Devices, End-to-End Journeys

1. Introduction

Human-computer interactions in mobile environments have an integral role in a personal and work setting. Mobile devices provide us with ubiquitous information access, entertainment and is helping users to stay connected to work, friends and family [6]. In comparison with traditional, visual based, interfaces, multimodal interfaces which use a mixture of visual, audio and haptic information, are better able to support rich expressiveness using familiar communication modalities. Additionally, they can be particularly well suited for mobile use as it is crucial to minimise demands on the users' workload while keeping flexibility of interaction available. Therefore, the major themes explored in this paper are the requirements and opportunities for mobile multimodal interfaces in the context of end-to-end journeys.

2. Multimodal human-computer interaction

Human-computer interaction in mobile environments becomes increasingly important as mobility poses further challenges due to a number of factors, such as the need or desire for multitasking, management of interruptions and unpredicted incidents, and additional distractions as well as the fluctuation of settings [1, 16]. These factors enlarge the risk of cognitive overload, which consequently diminishes the potential to carry out the primary task and increases the chance of task failure or even physical accidents. Traditional mobile devices mainly rely on a 12-key system or a QWERTY-keyboard and limited display

---

*Corresponding author. E-mail: psxck@nottingham.ac.uk.

size which complicate interactions in mobile settings. Therefore, Rössler et al [14] state that interfaces in mobile environments should provide the following features to tackle the additional challenges. Firstly, a personal environment should be travelling with the user. Secondly, the user should only gain access to information relevant for task completion in a data format best suited to the user's needs at that time. Finally, the interface should reflect the user's profile, location, situation, accessed network and current context. To achieve these, mobile devices need to gain an understanding of the precise context the user is in.

Turunen et al [16] suggest that location and timing information could be used to automatically determine the mode of transport as well as the environment the user is in. However, care needs to be taken as all modalities used in one system need to be carefully balanced according to the capacity and limits of human sensory channels [2]. If this careful approach is overshadowed by an enthusiasm to try to enable every modality at any time, the user will quickly be annoyed or even cognitively overloaded. Therefore, an understanding of the relationship between user context and presenting the user with suitable communication modalities is critical for multimodal interaction design.

An ideal setting for appreciating multimodal human-computer interaction in a mobile context is the concept of an end-to-end journey. This concept was developed by the Department of Transport [3] when analysing freight journeys such as a deep sea containers, or a food and drink, end-to-end journey. These case studies set out to understand the context of the entire journey with the full complexity of transport chains instead of analysing individual elements of a journey. However, the focus of this paper is on passengers and passenger oriented journeys such as the commute, meeting friends or family, travelling to events, and trips or holidays. In the literature some

efforts have been made towards identifying stages of passenger end-to-end journeys in order to increase efficiency, effectiveness and satisfaction of the journey for individuals. One example of an end-to-end rail journey is shown in Figure 1.

Thus, the aims of the paper are threefold. Firstly, it aims to confirm the stages encountered in end-to-end journeys for other, more universal journey types, with consideration of two different journey types – day trips and commutes. Secondly, it will map user information needs and sources to different journey stages. Finally, it will identify requirements and opportunities for multimodal interaction design in an end-to-end journey context.

3. Methodology

Two methods were used to reach the goals of this paper - focus groups and interviews. The focus groups were deemed the most appropriate method to meet the objectives of this study as little is known about end-to-end journeys and they provide the researcher with a tool for in-depth exploration [15]. The interviews were then used to enrich existing information by reviewing a different type of journey.

3.1. Participants

Two focus groups, discussing day trips, were carried out with three participants respectively, 4 of which were male and 2 were female and their age ranged from 24 to 36. A total of eight interviews were conducted with 4 males and 4 females. The interviews focused on the commute and the participants' age ranged from 22 to 47. All of the participants were in full-time employment and were paid volunteers.
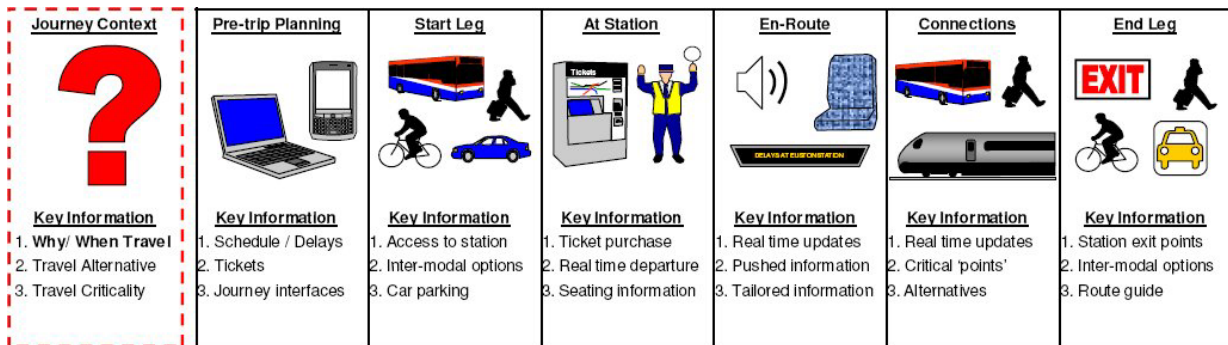

Fig. 1: Stages of the Railway End-to-End Journey (as shown in [4])

3.2. Focus group procedure

The focus groups were organised in two parts. During the first part, participants were asked to talk about any journey they have made or were planning to make, followed by a discussion of a day trip of their choice. This first part served two purposes. It introduced the participants to the topic of interest and the transcription of this part helped to formalise categories needed for content analysis of the data.

The second part of the focus group was scenario based in order to focus the discussion and to foster creative processes. The scenario was a day trip to the Peak District, an area famous for hill-walking in England, which enables the participants to envisage the use of mobile technologies in a specific context. During the first activity, participants discussed the different stages involved in this scenario. This method applies principles of participatory design, where end users are involved as full participants in activities leading to software or hardware computer products and computer based activities [8]. Moreover, they were asked to represent these journey stages and their content using cards with predefined symbols. Blank cards were also provided in case the participants felt that something was missing. The artefact created by the groups via the cards enabled them to communicate with each other and with the researcher using a different medium [10]. Moreover, the second activity focused on the information needs as well as the information sources used by the participants in the context of the scenario for each of the predefined stages.

3.3. Interview procedure

The interviews were carried out as part of a wider study called the Socially Connected Journey Project at the University of Nottingham [7]. The interview themes included questions about commute information and attitudes towards car sharing. During the interview the stages of the commute and the information required during the commute were covered.

Table 1: Categories for Content Analysis

Stages	Planning, Arrival, Activity, Return, Reflection
Information Needs	Choices, Cost, Navigation, People-2-People, Reviews, Tickets, Timing, Traffic, Weather
Sources of Information	Camera, Compass, Environment, Internet, Knowledge, Mobile Phone, Paper Maps, Radio, Satellite Navigation, Signs, Visitor Centre

3.4 Analysis

Content analysis was used to analyse the data gathered in the focus groups and the interviews. Content analysis can be defined as a research technique for making replicable and valid influences from texts [9]. Here the transcription of the focus groups and interviews were used. The method of inductive content analysis was applied and the high-level hierarchies, which were derived from the literature, were defined as stages, information needs and sources of information. The categories itself were then obtained using open coding and a feedback loop [5]. In this case the first two activities in the focus groups were used for open coding. The categories where then grouped to the high-level hierarchies and are applied to the remainder of the texts. Both are shown in Table 1. The unit of analysis used here is theme based rather than word based. Theme-based content analysis uses both qualitative and quantitative aspects of content analysis. Frequencies show general commonalities and to suggest prevalence of specific human factors related problems from the collected data [11].

4. Results

The results aim to firstly confirm the stages encountered in end-to-end journeys regarding two different journey types. The results will also map the user information needs and sources to these stages.

4.1. Stages of end-to-end journeys

The data suggests that end-to-end journeys consist of five stages. Table 2 summarises how often each stage was mentioned in the focus groups covering the day trip and the interviews concerning the commute.

Table 2: Journey Stages

Journey Stages	Focus Group 1	Focus Group 2	Interviews
Planning	10	11	11
Arrival	3	8	9
Activity	7	9	2
Return	1	9	7
Reflection	8	9	none

4.2. Day trips

The analysed data suggests that day trips have five key stages. The first stage is the planning stage, which is crucial to a day trip. The planning stage occurs before leaving the house for the journey to the destination. Both groups mentioned that a significant proportion of time is spent on researching elements of the trip such as choices of transport, timetables, routes, and activities as well as planning these, which becomes more critical when travelling with other people.

DayTrip#3: "Yeah, like when I was talking about that trip to Liverpool we nearly spent as much time researching as we did actually there."

Secondly, during the arrival stage people travel from their door step to the goal location. Here, the plan formulated in the previous stage is executed. Mobile devices are now used to coordinate people. Information, except for route or public transport information, is mainly consulted when a problem occurs which may then trigger adjustments to the original plan.

DayTrip#4: "This is all planning and then you would follow all of this but I think you would leave the computer things at the planning stage and the transfer to a mobile platform because you would not take your laptop with you."

Moreover, in the third stage the actual activity is carried out. The activities can range considerably for day trips, but in this specific example consisted of hill walking or mountain biking. Similarly to the arrival stage, the decision of the nature of the activity is made during the planning stage and executed here. However, it is much less accurate and more open to changes on the day.

DayTrip#6: "I am going to go and I am going to have an experience that I am not quite sure what it is going to be yet."

The return journey, from the place of the activity back home, defines the next stage of the day trip. This stage is almost identical to the arrival stage because the plan is executed unless a problem occurs. Finally, reflection takes place once arrived from the return journey. Parts of the reflection are immediate for instance planning a new trip or forming initial judgements about the trip:

DayTrip#3: "To some extent whenever you get back and you had a good time you feel like you want to plan the next trip."

Other parts of the reflection like organisation of memories, including any digital or physical information, or recollections, take part over time after the trip.

DayTrip#6: "Afterwards is much more, a sort of reflection, introspection and less, it's less information about the outside world and more looking at what has this done to me as a person. Was it worth doing? Have I got out if it what I thought I would get out of it? Have I got out of it completely different things I thought I was going to get out of it?"

These five stages were also represented using cards provided to the participants in the focus groups, as shown in Figure 2 and 3. This enabled the participants to express their ideas using a different media. With the aid of the cards insight on how multimodal devices could be used throughout the course of an end-to-end journey. The representations below illustrate that participants chose different ways to express the stages on paper. One follows a flow, whereas the other is much more modular.

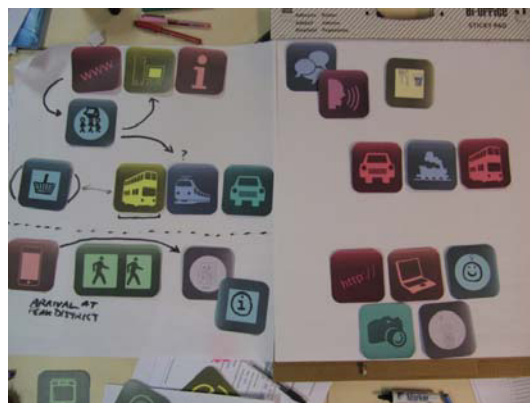Figure 2: Representation of Stages by Focus Group 1

Figure 3: Representation of Stages by Focus Group 2

The representation from one group illustrates the modes of transport used and the information sources consulted in each stage. The different choices, such as train, bus or car for the arrival and return journeys and the resources used for reflection are clearly shown. The second group shows the research sources, like the computer, the internet and other people. It also shows the arrival and return journeys in one area on the right, which were perceived to be equivalent. Entertainment was an additional focus in these stages. The activity, depicted in the middle, reveals that technology is barely used at this stage. Instead people focus on the environment and face-to-face communication. Finally, mobile technology and maps are used throughout the day trip. The reflection stage is not shown as participants felt it was an intrinsic process.

4.3. The commute journey

Contrary to the day trip, the planning stage is much shorter, more intuitive and a quick process just before leaving for work. The importance shifts from research to quick decision making on what transport or route to take depending on factors such as weather, activities during work or after work, and traffic.

Commute#3: *"It is pretty much the same, only if it is really pouring with rain or when it was really icy I got the bus instead of cycling."*

Research is unnecessary as individuals have gained enough experience so that a plan is intrinsically developed and retrievable.

Commute#7: *"When I first moved here I had a loan of my Dad's SatNav and to get to and from here I would use that and it would take me a certain way and then as I got to know the area I kind of experimented taking different way. So I had that sort of information when I was learning my way but now I know it so I don't really need any information at all."*

The arrival takes the individual from home to work. This stage is fairly similar compared to the day trip. Normal routines, rather than carefully planned journeys, are executed and amended if any problems occur. Upon arrival, the working day will define the activity carried out. Similarly to the day trip, the return journey is similar to the arrival. However, no reflection takes place at the end of the journey, individuals know their commute journeys and therefore evaluations as well as reflections are not necessary.

End-to-end journeys overall may not be limited to five stages. For instance, multiple activity stages could occur. Both focus groups which considered a stop at a restaurant or a hotel. To accommodate for

this, the stages can act as components with which an entire end-to-end journey can be represented.

4.4. Information needs and sources

A number of information needs and sources were identified for the day trip and the commute. The most interesting categories in terms of information needs are people-to-people, traffic and weather information.

For the day trips all participants agreed that they would with others and only two considered going on their own. So communication is essential to come to agreements about how to get there, what to do, when to go, and so on. Participants felt that this would become increasingly complicated with rising numbers.

DayTrip#3: *"Discussing it in the office, we were talking about going to the Peak District for a bike ride or something and that's kind of good but there is obviously some logistical problems of getting there and actually finding a decent route that everybody is ok with. Just sorting out the routes on the trip and how to get there, there are a few problems with that as well."*

This contrasts with the information needs of commuters. People mainly travel alone, but even if the commute involves travelling with others, it is very much a routine operation, so that communication is rarely essential. However, commuters consult information about the weather and traffic more frequently than those on a day trip. The information can influence the mode of transport or the route taken. So it is used to plan their commute. But on a day trip weather and certainly traffic information are less important.

Commute#5: *"If I found out on the radio before I leave in the morning that there has been an accident on the M1 that is probably one of the situation, where I would try to change my route and I reroute through Mansfield if I wasn't using the M1. That is a big factor."*

Overall, a considerable amount of information sources are consulted on a day trip including mobile phones, maps, knowledge, the internet and cameras. Three further observations were made. First, computers were mainly used in the planning stage, but people moved to a mobile platform throughout the day to get information on the go.

DayTrip#2: *"So yeah research wise use the internet. I would use the internet for this kind of thing. To find different places to sort of start from essentially."*

DayTrip#4: *"This is all planning and then you would follow all of this but I think you would leave the computer things at the planning stage and then*

transfer to a mobile platform because you would not take your laptop with you."

Secondly, participants used both their own knowledge and knowledge of others in all stages of the day.

DayTrip#4: "I might ask other people if they have heard or have been to the place before and if they have a good route or some good information."

Finally, during the activity itself participants felt that using technology was inappropriate. The goal was to be immersed with the surroundings, enjoying each other's company and the relaxation related with the activity. Thus, all felt that paper maps and compasses would be more suitable to the activity than using devices like smart phones. However, this result may change with the type of activity carried out.

DayTrip#4: "But yeah map, I don't tend to use technology in the Peak District, if I were in the Peak District. And I don't think I would like, unless I was by myself, I think the whole point about being there is,

for something like this, is to immerse yourself in the surroundings and to enjoy the company and the thing."

5. Discussion

This study aimed to identify the requirements and opportunities for multimodal interaction design in an end-to-end journey context comparing day trips and commute journeys. Table 3 shows a summary of the implications for multimodal interaction design.

Initially the research focused on confirming the different stages that occur during journeys. Both journey types show similarities in their stages. Evidence supports that the planning, arrival, activity and return stages exist in both, whereas the reflection stage occurs only in the day trip. Evaluation and reflection

Table 2: Journey Stages

Journey Type	Journey Stage	Information Priorities	Information Sources	Implications for Mobile Interaction Design
Atypical (e.g. Day Trips)	Planning	Coordinating people, choosing transport, considering reviews, cost, and timing, checking the route, buying tickets	Internet, Knowledge, Maps	Mainly performed on a computer at home so little opportunity for improving mobile interaction design
	Arrival	Navigating to the location, coordinating people, entertainment	Environment, Internet, Knowledge, Maps, Mobile Phone, SatNav	Varied and rich information needs Need or desire to multitask
	Activity	Coordinating people, activity related information	Camera, Environment, Maps, Mobile	Use Multimodal interaction to decrease the risk of overload and support for quick decision making
	Return	Considering timing, entertainment	Knowledge, Mobile	Modalities depend on user context, preferences, abilities and the nature of tasks
	Reflection	Create reviews, evaluate memories	Camera, Internet, Knowledge	Mainly user intrinsic so little opportunity for improving mobile interaction design
Routine (e.g. Daily Commute)	Planning	Transport Choices, Timing, Traffic, Weather	Environment, Internet, Knowledge, Radio	Need for real-time information to support quick decision making Need or desire to multitask Use Multimodal interaction to decrease the risk of overload and support for quick decision making
	Arrival	Navigation, Traffic	Knowledge, Radio	Modalities depend on user context, preferences, abilities and the nature of tasks
	Activity	Work related activities	Work related	Mainly performed in an office or static setting so little opportunity for improving mobile interaction design
	Return	None	Based on experience	Mainly based on user experience so little opportunity for improving mobile interaction design

are critical for journeys that are atypical. However, during typical, routine journeys such as the daily commute, these were not required. Secondly, the paper aimed to map the user information needs and sources to different journey stages. Commuters mainly depended on traffic and weather information during the planning and arrival stages. On the contrary, the information needs of day travellers were considerably richer, supporting the claim that the need for information is higher for occasional, atypical journeys such as day trips. Thus, the potential for mobile multimodal devices is much bigger for occasional journeys. However, commuters emphasized the lack of real time traffic information, showing that timely information is key.

During the planning stage of the day trip, using the internet on a computer was the focal starting point for gathering information. Participants felt reluctant to use a mobile interface for this process. However, all participants highlighted the move from computers to a mobile platform during the day. Here, information such as people-2-people communication and navigation were critical and thus shape the requirements for a mobile information device. Multimodal interaction is crucial during arrival, activity and return stages to improve usability of mobile devices especially since users are often forced to multitask. During the day trip, for instance, users may be walking or driving when looking for navigation information or information about their surroundings. Here, insights from cognitive ergonomics theory are key. When tasks become more complex humans need to use distinct resources to ensure task demands are spread in order to avoid cognitive overload [13].

The use of distinct resources can be triggered by communicating information via different modalities. Therefore, the use of audio, visual, tactile or other modalities shape the technical requirements for mobile technologies. The exact application of these still has to be established in future research, where the role of context for mobile interaction has to be established. Moreover, the nature of the tasks that are carried out by the user at the time will also influence which modalities are most appropriate to convey information. For instance, it can be assumed that both walking and driving heavily rely on visual perception and spatial resources. Taking this knowledge as well as the dichotomies from Wickens' Multiple Resource Model into account [17], the use of distinct cognitive resources for a pair of tasks can be triggered. So, when walking or driving, auditory perception and categorical or symbolic processes are available to cope with mobile interaction tasks. Thus, the study

confirms suggestions from the literature, stating that mobile technologies should incorporate new input modalities such as gesture recognition, haptic feedback and speech to allow the user to convey information in the most convenient way [14].

However, the social situation may have an effect on mobile interactions as well, as it is considered rude to use devices as an information source. Future research is required to gage the role of social context in mobile interactions and investigate ways to reduce the barrier when needed..Moreover, the user should not be overloaded by the availability of too many modalities [2]. The nature and number of modalities available depends on the type of activity undergone and varies with the user context and environment. Further research needs to examine during which activities mobile devices are acceptable or how they need to change in order to add to the experience. Turunen et al [16] suggest timing and location information are crucial to determine the user setting. Thus, mobile technologies are required to exploit existing functionality to understand the user context. Finally, understanding the nature of the tasks that the user means to carry out is key to draw on suitable modalities that trigger the use of cognitive resources, which are untapped at the time.

6. Conclusions

Presently, interaction with mobile devices is very cumbersome and often leads to cognitive overload. Moreover, it is not flexible enough to cope with a number of different contexts or the dynamic nature of these. Multimodal human-computer interaction offers much more flexibility in a mobile setting and it provides the user with more natural communication methods. Therefore, this paper set out to explore the requirements and opportunities of multimodal interaction design for mobile devices in the context of end-to-end journeys, by confirming the stages of end-to-end journeys and mapping the users' information needs and sources to these. Focus groups and interviews were used to accomplish these goals. The study found that both day journeys consist of planning, arrival, activity, and return stages with the addition of a reflection stage for the atypical day trip. The information needs varied within the journey types. For the day trip these were much richer, relating to the choices of modes of transport, the activity, people-to-people communication and navigation for commuters these were limited to traffic and weather.

Mobile devices have the greatest potential in the stages of arrival, activity and departure because of the rich and varied information needs particularly for atypical journeys. The remaining stages present fewer opportunities as planning is performed on the computer and reflection is mainly user intrinsic. Furthermore, during the arrival, activity and departure stages, multimodal interaction becomes crucial, because users often multitask. The increased complexity of a mix of tasks calls for the use of distinct cognitive resources, which can be triggered by the use of different communication modalities. The choice of modality depends on the user preferences or abilities, as well as the social and task context. The use of suitable communication modalities would decrease the chance of cognitive overload and increase the efficiency and possibly the enjoyment of interactions.

Future research first of all needs to examine how to deliver information and how individuals would communicate with the mobile device depending on the stage and context they are in. Moreover, considerable research is required in order to determine which elements of mobile contexts are critical for determining when and how information should flow between the user and the device.

Acknowledgements

References

[1] L. Barnard and J.S. Yi, Capturing the Effects of Context on Human Performance in Mobile Computing Systems, Personal Ubiquitous Computing 11 (2007), 81 – 96.

[2] M. Cellario, Human-Centred Intelligent Vehicles: Toward Multimodal Interface Integration, IEEE Intelligent Systems 16, 4 (2001), 78 – 81.

[3] Department of Transport, Towards a Sustainable Transport System - The Logistics Perspective: End-to-end journey case Studies, (2008). http://tinyurl.com/3d783de

[4] Department of Transport , *Integrated Passenger Information:* Delivering the Rail End to End Journey Stage 3 Report (2010). http://tinyurl.com/3lu78yn

[5] S. Elo and H. Kyngäs, The qualitative content analysis process, Journal of Advanced Nursing 62,1 (2007), 107 – 115.

[6] J. Froehlich, M.Y. Chen, S. Consolvo, B. Harrison and J.A. Landay, MyExperience: A System for In situ Tracing and Capturing of User Feedback on Mobile Phones, IN: Proc. MobiSys 2007.

[7] Horizon, Horizon Car Share, Horizon (2010). http://tinyurl.com/43l3tr6

[8] F. Kensing and J. Blomberg, Participatory Design: Issues and Concerns, Computer Supported Cooperative Work 7 (1998), 167 – 185.

[9] K. Krippendorff, Content Analysis: An Introduction to its Methodology. Sage Publications , Beverly Hills (1980).

[10] M.J. Muller, Participatory Design: The Third Space in HCI, IN: The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emergin Applications (2nd Edition), A. Sears and J.A. Jacko, ed., Lawrence Erlbaum Associates, London, (2008), pp. 1061 – 1081.

[11] H. Neale and S. Nichols, Theme-based Content Analysis: A Flexible Method for Virtual Environment Evaluation, International Journal of Human-Computer Studies 55 (2001), 167 – 189.

[12] S. Oviatt, R. Coulston and R. Lunsford, When Do We Interact Multimodally? Cognitive Load and Multimodal Communication Patterns, IN: Proc. 6th International Conference on Multimodal Interfaces (2004), 129 – 136.

[13] S. Oviatt, A. DeAngeli and K. Kuhn, Integration and Synchronisation of Input Modes during Multimodal Human-Computer Interaction, IN: Proc. CHI (1997), 415 – 422.

[14] H. Rössler, J. Sienel, W. Wajda, J. Hoffmann and M. Kostrzewa, Multimodal Interaction for Mobile Environments, IN: Proc. of Workshop IPNMD (2001).

[15] D.W. Stewart, P.N. Shamdasani and D.W. Rook. Focus Groups: Theory and Practice (2nd Edition). Sage Publications, London (2007) .

[16] M. Turunen, T. Hurtig, J. Hakulen, A. Virtanen and S. Koskinen, Mobile Speech-Based and Multimodal Public Transport Information Services, IN: Proc. SiMPE (2006) .

[17] C.D. Wickens, Multiple Resources and Performance Prediction, Theoretical Issues in Ergonomic Science 3, 2 (2002), 159 – 177.