

Generating Situated Assisting Utterances to Facilitate Tactile-Map Understanding: A Prototype System

Kris Lohmann, Ole Eichhorn, and Timo Baumann

Department of Informatics, University of Hamburg

Vogt-Kölln-Straße 30

22527 Hamburg, Germany

{lohmann, oeichhor, baumann}@informatik.uni-hamburg.de

Abstract

Tactile maps are important substitutes for visual maps for blind and visually impaired people and the efficiency of tactile-map reading can largely be improved by giving assisting utterances that make use of spatial language. In this paper, we elaborate earlier ideas for a system that generates such utterances and present a prototype implementation based on a semantic conceptualization of the movements that the map user performs. A worked example shows the plausibility of the solution and the output that the prototype generates given input derived from experimental data.

1 Introduction

Humans use maps in everyday scenarios. Especially for blind and visually impaired people, tactile maps are helpful accessible substitutes for visual maps (Espinosa, Ungar, Ochaita, Blades, & Spencer, 1998; Ungar, 2000). However, tactile maps are less efficient than visual maps, as they have to be read sequentially. A further problem of physical tactile maps is restricted availability. While physical tactile maps are rarely available and costly to produce, modern haptic human-computer interfaces can be used to present virtual variants of tactile maps (*virtual tactile maps*) providing a similar functionality. For example, the Sensable Phantom Omni device used in our research enables a user to feel virtual three-dimensional objects (see Figure 1). It can be thought of as a reverse robotic arm that makes virtual haptic perception possible by generating force feedback. In the context of the research discussed, these objects are virtual tactile maps. These consist of a virtual plane on which

streets and potential landmarks (such as buildings) are presented as cavities.

In recent work, Habel, Kerzel, and Lohmann (2010) have suggested a multi-modal map called *Verbally Assisting Virtual-Environment Tactile Map* (VAVETaM) with the goal to enable more efficient acquisition of spatial survey (overview) knowledge for blind and visually impaired people.

VAVETaM extends the approaches towards multi-modal maps (see Section 2) by generating situated spatial language. The prototype described reacts to the user's exploration movements more like a human verbally assisting a tactile map reader would do, e.g., by describing spatial relations between objects on the map. The users may explore the map freely, i.e., they choose which map objects are of interest and in which order they explore them. This demands for situated natural language generation (Roy & Reiter, 2005), which produces timely appropriate assisting utterances. Previously, the suggested system has not been implemented.

The goal of this paper is to show that the ideas of Lohmann, Kerzel, and Habel (2010) and Lohmann, Eschenbach, and Habel (2011) can be implemented in a prototype which is able to generate helpful assisting utterances; that is, to show that the language-generation components of VAVETaM are technically possible. The remainder of the paper is structured as follows: We first briefly survey some related work in Section 2, and then describe the overall structure of VAVETaM in Section 3. We then present a description of our system in Section 4 paying special attention to the input to natural language generation (Subsection 4.1) and the generation component itself (Subsection 4.2). We show the appropriateness of the approach by discussing an example input, the pro-



Figure 1: The Sensable Omni Haptic Device and a Visualization of a Virtual Tactile Map.

cesses performed, and the automatically generated output in Section 5 before we close with concluding remarks in Section 6.

2 Related Work

To make maps more accessible for visually impaired people by overcoming drawbacks of uni-modal tactile maps, a number of multi-modal systems that combine haptics and sound have been developed. An early system is the NOMAD system. It is based on a traditional physical tactile map, which is placed on a touch pad. The system allows for the association of sound to objects on the map (Parkes, 1988, 1994). The approach to use traditional physical tactile maps as overlays on touch pads has been used in various systems that were developed subsequently (e.g., Miele, Landau, & Gilden, 2006; Wang, Li, Hedgpeth, & Haven, 2009). Overviews of research on accessible maps for blind and visually impaired people can be found in Buzzi, Buzzi, Leporini, and Martusciello (2011) and in De Almeida (Vasconcellos) and Tsuji (2005). Other researchers have advanced the way haptic perception is realized by using more flexible human-computer-interaction systems that do not need physical tactile map overlays. For example, Zeng and Weber (2010) have proposed an audio-tactile system which is based on a large-scale

braille display and De Felice, Renna, Attolico, and Distante (2007) presented the Omero system, which makes use of a virtual haptic interface similar to the interface used in our research.

Existing systems work on the basis of sounds or canned texts that are associated to objects or areas on the map. Sound playback starts when the user touches a map object or, in some systems, by clicking or tapping on it. Yet, when humans are asked to verbally assist a virtual tactile map explorer, they produce assisting utterances in which they make much more use of spatial language and give brief augmenting descriptions of the objects that are currently explored and their surroundings (Lohmann et al., 2011). Based on this, Lohmann and colleagues suggest which informational content should be included in assisting utterances for a tactile-map reading task. Among the types of information that are suggested for verbal assisting utterances is information allowing for identification of objects, e.g., by stating its name (e.g., ‘This is 42nd Avenue’); information about the spatial relation of objects (‘The church is above the museum’); and talking about the ends of streets that are explored (‘This street is restricted to the left by the map frame’).

Empirical (Wizard-of-Oz-like) research with 24 blindfolded sighted participants has concerned an audio-tactile system that makes use of assisting utterances containing the information discussed above and shown its potential. Different outcome measures, among them sketch maps and a verbal task, showed an improved knowledge acquisition with verbal assisting utterances compared to a baseline condition in which participants verbally only received information about the names of objects (Lohmann & Habel, forthcoming). Empirical research with blind and visually impaired people is ongoing. Data from the ongoing experiment with blind and visually impaired participants is used to show the function of the system in Section 5.

3 The Structure of VAVETaM

In this section we will recap the overall structure of VAVETaM as presented by Habel et al. (2010) and Lohmann et al. (2010). Figure 2 depicts the relevant parts of the structure.

The *Virtual-Environment Tactile Map* (VETM)

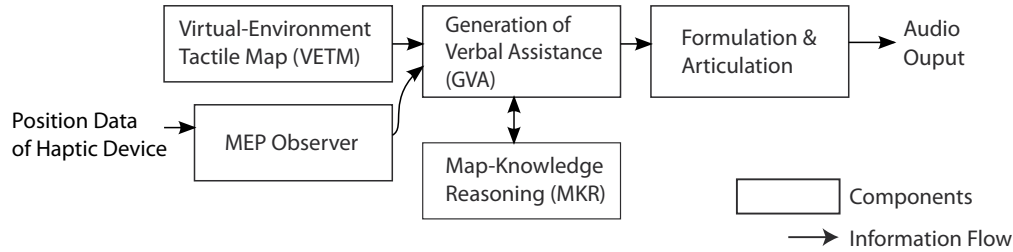


Figure 2: The Interaction of the Generation Components with Other Components of VAVETaM (modified version following Habel et al., 2010).

knowledge base forms the basis for rendering the tactile map, for analyzing movements, and for verbalizing assistive utterances forming the central knowledge component in the architecture.

Knowledge needed for natural-language generation is represented in a propositional format which is linked to knowledge needed for movement classification and for the haptic presentation of the map. The latter is stored in a spatial-geometric, coordinate-based format. The knowledge for assistance generation is represented using the Referential-Nets formalism developed by Habel (1986) and successfully used by Guhe, Habel, and Tschander (2004) for natural language generation. Knowledge for verbalization is organized by interrelated Referential Objects (RefOs), which are the potential objects of discourse. A referential object consists of an identifier for the object (an arbitrary string, for example *pt3*), additional associated information such as the sort of the object, and associated propositional information that can be verbalized (such as the name of the object and relations to other objects, e.g., that the object is ‘left of’ another object). Important sorts of objects in the map domain are potential landmarks, regions, the frame of the map, and tracks and track segments¹. See Lohmann et al. (2011) for a discussion of the propositional layer of the VETM knowledge base.

The *Haptic Device* provides a stream of position data. This stream of data is the input to the *Map-Exploratory-Procedures Observer* (MEP Observer) component and its subcomponents which analyzes the movements the map user performs. By categorizing the movements and specifying them with identi-

fiers of the objects currently explored by the user, a conceptualization of the user’s movements is created that is suitable as input to the component dealing with assisting-utterance generation. For the case of tactile-map explorations, different circumstances affect which information shall be given via natural language in an exploration situation: (a) what kind of information is the user is trying to get (exploration category), (b) about which object the user is trying to get information, and (c) what has happened before (history).

The *Map-Knowledge Reasoning* (MKR) component serves as memory for both the MEP Observer and the GVA component by keeping track of verbal and haptic information that has been presented to the user. This component hence helps to avoid unnecessary verbal repetitions.

The *Generation of Verbal Assistance* (GVA) component, which is at the core of the prototype that we will present in Section 4, solves the central task of natural language generation. This component selects the knowledge that is suitable for verbalization in an exploration situation from the VETM knowledge base and prepares it in a way appropriate for further output. It sends *preverbal messages* (PVMs, see Levelt, 1989), propositional representations of the semantics of the planned utterance, to the *Formulation & Articulation* components for the generation of a surface structure and final utterance.

4 Description of the Prototype

In order to show how an artificial system is able to generate situated assistance in a well-formed fashion, we present a prototype implementation of the core components for natural language generation in the VAVETaM system.

¹A track is a structure enabling locomotion, such as a street. The meaning of the term is similar to the meaning of the term ‘path’ introduced by Lynch (1960).

We implemented dummy components in place for the Map-Knowledge Reasoning (MKR) and MEP Observer components to allow us to test the natural language output. The *MKR Simulator* provides basic functions sufficient to avoid unnecessary repetitions of utterances by preventing production of the same message for a defined time period. An exception to this rule are those messages that are needed to identify an object on the map, such as ‘This is Dorfstraße’, which are given every time the user touches an object.² The *MEP Simulator* generates input to the component as the MEP Observer is planned to do (see Kerzel & Habel, 2011, for a discussion of a possible technical realization).

In the following subsection, we will discuss Map-Exploratory Procedures (MEPs), which are output by the MEP Observer and form the basic input to the generation component (GVA), which we then discuss in Subsection 4.2. Finally, we present the inner workings of the Formulation & Articulation components in Subsection 4.3.

4.1 Conceptualization of the User’s Movements

One of the core challenges for situated natural language generation is to timely connect the user’s percepts (in the case of virtual-tactile-map exploration indicated by movements that the user performs with the device) to symbolic natural language (Roy & Reiter, 2005). The task to be solved is to have a well-specified conceptualization of exploration situations. An *exploration situation* is constituted by the kind of movements the user performs, the map objects the user wants to gain knowledge about (which constitutes the haptic focus (Lohmann et al., 2011)), and the haptic exploration and verbalization history. In the structure of the VAVETaM system, the MEP Observer fulfills the task of categorizing the user’s movements and detecting objects in the haptic focus. Lohmann et al. (2011) discuss how *Map-Exploratory Procedures* (MEPs), a specialization of Exploratory Procedures, introduced as categories of general haptic interaction by Lederman and Klatzky (2009), can be used to categorize the map user’s movements. MEP types are shown in Table 1.

For example, a *trackMEP* is, straightforwardly,

²User studies showed that the verbal identification is necessary to recognize the haptic objects.

MEP Type	Indication
trackMEP	Exploration of a track or track segment object
landmarkMEP	Exploration of a potential landmark object
regionMEP	Exploration of a region object
frameMEP	Exploration of a frame object
stopMEP	No exploration

Table 1: Map Exploratory Procedures (MEPs).

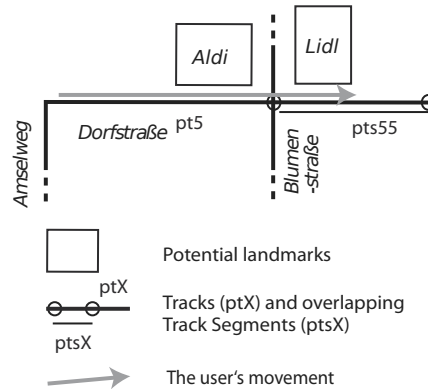


Figure 3: Visualization of a Part of a Virtual Tactile Map.

characterized by a track-following movement indicating that the user wants to know something about a track object. MEPs are (optionally) specified with identifier(s) that link objects on the propositional layer of the VETM knowledge base as belonging to the haptic focus of the MEP.

In this work, we extend the concept to be able to cope with multiple objects or parts of objects that can simultaneously be in the user’s haptic focus. The following example illustrates overlapping haptic foci (see Figure 3). Consider the track with the name ‘Dorfstraße’ being represented as track object *pt5* on the propositional layer of the VETM knowledge base. If the track *pt5* forms a dead end, this dead end can additionally be represented as a unique track segment object (*pts55*). When the user explores the track *pt5* from the left to the right, at a certain point, both *pt5* and *pts55* are in the haptic focus.

Since the user is exploring a track, the movement is characterized by a *trackMEP* which is specified by the objects *pt5* and *pts55* and either will be in the *primary haptic focus*.³ Thus, in this case, *pts55* is in

³Notice that the decision whether in fact *pt5* or *pts55* are

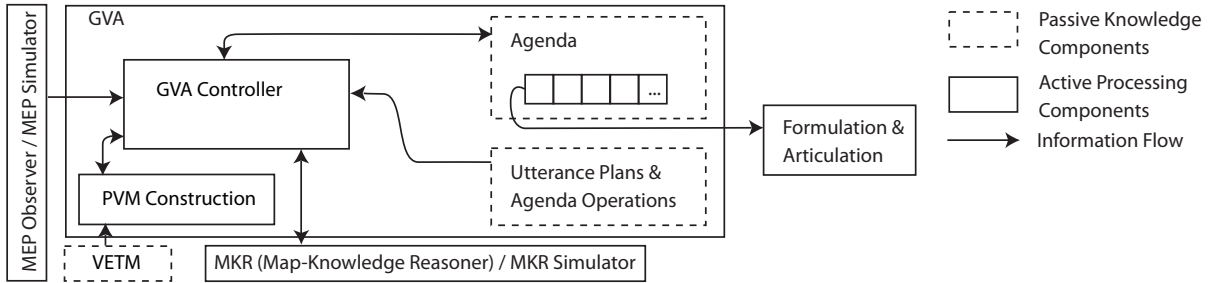


Figure 4: The Architecture of the Generation of Verbal Assistance Component.

the *secondary haptic focus*. It is reasonable to talk about both, the track and the dead end itself. As a notational convention, we denote MEPs by their type, the object in primary focus (if available), and a (possibly empty) list of objects in secondary focus. For the example above, we write `trackMEP(pt5, [pts55])`.

4.2 Structure of the Generation Component

The focus of our prototype is on the GVA component of the VAVETaM system that solves the ‘*What to say?*’ task, the task of determining the content appropriate for utterance in an exploration situation (Reiter & Dale, 2000; De Smedt, Horacek, & Zock, 1996). This component interacts with different components introduced above (see Section 3): (1) it receives the conceptualization of the user’s movements (MEPs and specifications) from the MEP Observer; (2) it accesses the propositional layer of the VETM knowledge base in order to retrieve information about the objects that is suitable for verbalization; (3) it interacts with the MKR component, which keeps track of the exploration and verbalization history; and (4) it then sends semantic representations in the form of preverbal messages (PVMs) to the Formulation & Articulation components.

The GVA component consists of several subcomponents which are visualized in Figure 4. The *GVA Controller* controls the execution of other processes through controlling the *Agenda*, which is an ordered list of preverbal-message representations of utterances.⁴ Once the GVA component receives a (specified) MEP describing the user’s movements from the MEP Observer, it looks up *Utterance Plans* & focused primarily upon is up to the MEP Observer component.

⁴The term ‘Agenda’ is used in a similar context in the Collage system (Rich, Sidner, & Lesh, 2001).

Agenda Operations that specify which information to express is suitable in the given exploration situation and where it should be placed on the Agenda. The *PVM Construction* component searches an utterance plan that allows to construct a preverbal message that contains this information. The top element of the Agenda is passed on to the Formulation & Articulation component as soon as that component has finished uttering the previous element.

In the current implementation of the GVA, utterance plans are stored as lists of potential messages and construction rules. For example, with a *trackMEP*, associated knowledge is stored that the object shall first be identified (by either stating the name associated to that object, e.g., ‘Dorfstraße’ or choosing a referring expression that allows for definite identification). Then, if available, information about geometric relations such as parallelism with other linear objects on the map is selected from the VETM knowledge base, followed by information about spatial relations with other map objects. Subsequently, the construction of a preverbal message that informs the user about the extent of the track in the haptic focus is tried, followed by information about crossings the track has. For each of these construction rules is tested whether the VETM knowledge base contains suitable information. If it does, a preverbal message is generated and added to the Agenda unless the MKR component rejects the message because this utterance is inappropriate given the exploration and verbalization history, which prevents unnecessary repetitions of information. For example, if the user has previously explored the track `pt5` and already received the information that the buildings ‘Lidl’ and ‘Aldi’ (cf. Figure 3) are above the track a short time before, the articulation of this information is pre-

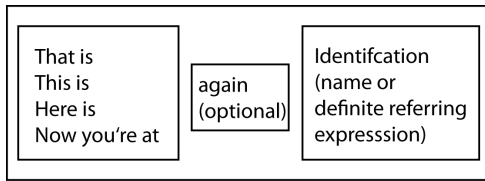


Figure 5: Literal Translation of the Template for a German Identification Message.

vented and the user is given other information (or none, if no more suitable information is available).

4.3 Formulation and Articulation

In the prototype system presented, formulation is implemented in a template-based approach (Reiter & Dale, 2000). The Formulation component uses a set of sentence templates which consist of partial lexicalizations and gaps to fill with information for the exploration situations. Additionally, a lexicon stores knowledge about natural language expressions that can be used to express spatial situations. Figure 5 shows a simple template used for the generation of identification messages.⁵ Of the four utterance parts depicted in the left box, one is chosen randomly enabling some variation in the utterances. If the MKR component has marked the preverbal message as a repetition of a previously articulated utterance, a marker word is placed in the sentence (here: ‘again’). Then, the sentence is completed by either selecting the name of the object in focus from the VETM knowledge base or by selecting a referring expression. The former results in utterances such as ‘This is Dorfstraße’. This text is then synthesized using text-to-speech (TTS) software.

5 A Worked Example

As described, the development of the component that conceptualizes the user’s movement is not yet finished. Therefore, to show the function of the implementation, we used example inputs that were derived by manually annotating screen-records from experimental data that was previously collected in Wizard-of-Oz-like experiments with blindfolded sighted, blind, and visually impaired people. In these ex-

⁵Note that the system is implemented in German; the ordering of elements indeed leads to grammatically correct German sentences.

periments, participants received pre-recorded verbal assisting utterances that were selected by the experimenter using a custom-built software tool based on a visualization of the user’s movement on a computer screen (Lohmann & Habel, forthcoming, and Section 2). Using video records of the visualizations of the user’s movements, the first author manually annotated the relevant MEPs and their specifications that, in the VAVETaM structure, the MEP Observer component should output. These manually annotated MEPs form the input to test the prototype system.⁶

In order to exemplify the function of the generation system, a small part of one of the annotated inputs is detailed in this section.⁷ Figure 6 visualizes a part of the movement of a visually impaired map explorer and the corresponding names and identifiers of the objects used for the specification of the MEPs in the VETM knowledge base. As the figure shows, the map explorer touches the track *pt3*, coming from the left. The track is explored for a while with small movements. (This position is remained for a relatively long time, maybe listening to the ongoing utterances.) Then, the map explorer proceeds to the bottom end of the track before following the track upwards. Figure 6 shows that the bottom end of the track is conceptualized as distinct track segment, track segment *pts33*, which is part of the track *pt3*.

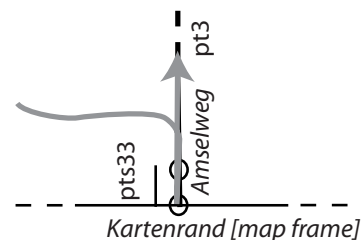


Figure 6: Example Movement a Visually Impaired Map Explorer Performed in an Ongoing Experiment.

The annotated MEPs and their specification of this small exploration movement are shown in Table 2. The GVA component and the Formulation & Articulation components generate detailed log files that in-

⁶Detecting MEPs is an instance of event detection in virtual haptic environments (Kerzel & Habel, 2011), which showed its applicability for the task in an early prototype (M. Kerzel, personal communication).

⁷We also tested other annotated inputs; this example is representative of the behavior of the prototype.

Time in Seconds	Input to the GVA
...	...
33.0–54.0	trackMEP(pt3)
54.0–57.0	trackMEP(pt3, [pts33])
57.0–57.8	trackMEP(pt3)
...	...

Table 2: Manually Categorized MEPs and Specifications for the Exploration Depicted in Figure 6.

dicating which information has been selected from the VETM knowledge base, which preverbal messages (PVMs) are put onto the Agenda, and how utterances are articulated. Based on the log files, we detail the processes performed by the GVA component and the resulting verbal output in Table 3.

During the user’s long first exploration movement of the track pt3 from seconds 33 to 54, which is conceptualized by trackMEP(pt3), the GVA component expresses all the information that is associated with the track pt3 in the VETM. The first message informs the user about the identity of the track by stating the identifying utterance ‘This is Amselweg’. Then, the user is informed about geometric relations of this track to other tracks. In the present case, information about parallelism with the track pt4 is available in the VETM and a corresponding utterance is produced. Subsequently, the user is informed about the extent of the track, i.e., where it ends. Then, information about the intersections the track has is uttered. These are all assisting utterances that are possible given the current MEP and the knowledge base.⁸

Next, the user moves downwards resulting in the distinct track segment pts33 coming into secondary focus. All PVMs about the object in primary focus (pt3) are blocked by the MKR component, as they have just been uttered. Thus, a message that informs the user about his or her position on the track segment is formulated, resulting in a message such as ‘Here, Amselweg is restricted by the map frame’. When the user leaves the track segment pts33, no further assisting utterances are given as all information associated with the track pt3 has been expressed recently.

⁸Note that the order in which information is given is fixed in the current system as explained in Subsection 4.2. Whether giving the messages in another order, which is potentially more flexible, is more helpful, has to be further evaluated.

...
33.0–54.0 s
MEP Simulator fires trackMEP(pt3)
GVA receives: trackMEP(pt3)
GVA clears agenda due to MEP change
PVM Construction is able to generate PVMs of class: Identification, Geometric_Relation, Extension, Junctions
PVMs Identification, Geometric_Relation, Extension, Junctions, are put on the Agenda (0 prohibited by Map-Knowledge Reasoner)
Formulation getting Identification PVM for the RefO pt3: the following aspects have been chosen by PVM Construction: name ‘Amselweg’
Speechout: “Dies ist der Amselweg.” [<i>“This is Amselweg.”</i>]
Formulation getting Geometric_Relation PVM for the RefO pt3: the following aspects have been chosen by the PVM Construction: IS_PARALLEL_TO with the arguments [pt3, pt4]
Speechout: “Parallel zu ihm verläuft die Blumenstraße.” [<i>“... which is parallel to Blumenstraße”</i>]
Formulation getting Extent PVM for the RefO pt3: the following aspects have been chosen by the PVM Construction: predicate HAS_UPPER_LIMIT with the arguments [pt3, ptco1]; predicate HAS_LOWER_LIMIT with the arguments [pt3, pfr3]
Speechout: “. . . er mündet nach oben in die Dorfstraße und endet unten am Kartenrand.” [<i>“... it forms a corner with Dorfstraße at the top and at the bottom is restricted by the map frame.”</i>]
Formulation getting Junctions PVM for the RefO pt3: the following aspects have been chosen by the PVM Construction: predicate IS_IN_TRACK_CONFIG with the arguments [pt3, ptco4]
Speechout: “Außerdem hat er eine Kreuzung mit der Hochstraße.” [<i>“Furthermore, the street crosses Hochstraße.”</i>]
54.0–57.0 s
MEP Simulator changes MEP specification to trackMEP(pt3, [pts33])
GVA receives: trackMEP(pt3, [pts33])
GVA detects secondary focus change
PVM Construction is able to generate PVMs of class: Identification
Identification-class PVM is put at the front of the Agenda (0 prohibited by Map-Knowledge Reasoner)
Formulation getting Identification PVM for the RefO pts33
...

...
Speechout: "Hier endet der Amselweg am Kartenrand." ["Here, Amselweg is restricted by the map frame."]
57.0–57.8 s
MEP Simulator changes MEP specification to track-MEP(pt3)
GVA receives: trackMEP(pt3)
Nothing happens, primary focus not new
...

Table 3: The Processes and Output (German and Translated) of the GVA and the Formulator.

6 Conclusion

We presented a prototype system that generates situated assisting utterances for tactile-map explorations to ease tactile map learning. The prototype is based on an earlier concept. We focussed on the GVA component in the system, which solves the ‘What to say?’ task of natural language generation, taking into account the situated context. We exemplified the working of the component in a testing environment based on a conceptualization of a part of a real tactile-map exploration, for which it generates plausible and timely output that is comparable to assisting utterances that were in previous research tested in Wizard-of-Oz-like experiments with blindfolded sighted people and in ongoing experiments with blind and visually impaired people. Therefore, we conclude that a generation system working in the manner described is technically possible. We also explained in detail the structure and implementation of MEPs, which are the basis for categorization of the user’s movements and, with additional specification, the input to the GVA component.

More fine-grained analysis is needed to gain knowledge (1) about how much information should be given via the verbal channel to maximize efficiency, and (2) whether the system can be improved by using more flexible Utterance Plans.

7 Discussion and Outlook

One problem which became apparent in the experiments and also in preliminary tests of the fully integrated prototype system is the fact that the user’s exploration movements on the map may be very quick. In these cases, the information to be delivered may already be outdated when the assistive utterance con-

veys this information. This is partly due to the German word order, as can be seen in Figure 5, which shows the template for identification messages.

Problems can occur in cases where an utterance is verbalized shortly before the user starts exploring another map object. In this case, the exploration situation changes during articulation. Currently, the components concerned with language generation work in a modularized sequential manner without feedback. If an utterance was sent to formulation, it cannot not be changed anymore. Hence, it can happen that assisting utterances and the user’s exploration are not in all cases timely.

One possible remedy to this problem is to extend the formulation to work in an *incremental* fashion such that it explicitly handles situations in which a currently articulated utterance is outdated (e.g., an identification utterance that is no longer valid because the object to be identified has gone out of focus) and by altering it to a new utterance of similar structure (i.e., an identification utterance for a different object which just came into the haptic focus). In this case, it could *adapt* the ongoing utterance (if it is still in an early stage of production) to replace the previous identifying word (e.g., ‘Amselweg’) with the new word (i.e., ‘Dorfstraße’). Of course, this is only possible if the articulation (text-to-speech synthesis) works in an incremental fashion (i.e., it is able to change yet unspoken parts of an ongoing utterance). Such work is currently ongoing and we plan to integrate this functionality in our future work.

Acknowledgments

The research reported in this paper has been partially supported by DFG (German Science Foundation) in IRTG 1247 ‘Cross-modal Interaction in Natural and Artificial Cognitive Systems’ (CINACS). We thank the anonymous reviewers for their highly useful commentaries.

References

- Buzzi, M., Buzzi, M., Leporini, B., & Martusciello, L. (2011). Making visual maps accessible to the blind. *Universal Access in Human-Computer Interaction. Users Diversity*, 271–280.
- De Almeida (Vasconcellos), R. A., & Tsuji, B. (2005). Interactive mapping for people who are

- blind or visually impaired. In *Modern cartography series* (Vol. 4, pp. 411–431). Elsevier.
- De Felice, F., Renna, F., Attolico, G., & Distanto, A. (2007). A haptic/acoustic application to allow blind the access to spatial information. In *World haptics conference* (pp. 310 – 315).
- De Smedt, K., Horacek, H., & Zock, M. (1996). Architectures for natural language generation: Problems and perspectives. *Trends in Natural Language Generation An Artificial Intelligence Perspective*, 17–46.
- Espinosa, M. A., Ungar, S., Ochaita, E., Blades, M., & Spencer, C. (1998). Comparing methods for introducing blind and visually impaired people to unfamiliar urban environments. *Journal of Environmental Psychology*, 18, 277 – 287.
- Guhe, M., Habel, C., & Tschander, L. (2004). Incremental generation of interconnected preverbal messages. In T. Pechmann & C. Habel (Eds.), *Multidisciplinary approaches to language production* (pp. 7–52). Berlin, New York: De Gruyter.
- Habel, C. (1986). *Prinzipien der Referentialität*. Berlin, Heidelberg, New York: Springer.
- Habel, C., Kerzel, M., & Lohmann, K. (2010). Verbal assistance in Tactile-Map explorations: A case for visual representations and reasoning. In *Proceedings of AAAI workshop on visual representations and reasoning 2010*.
- Kerzel, M., & Habel, C. (2011). Monitoring and describing events for virtual-environment tactile-map exploration. In M. F. W. A. Galton & M. Duckham (Eds.), *Proceedings of workshop on 'identifying objects, processes and events', 10th international conference on spatial information theory*. Belfast, ME.
- Lederman, S., & Klatzky, R. (2009). Haptic perception: A tutorial. *Attention, Perception, & Psychophysics*, 71(7), 1439–1459.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.
- Lohmann, K., Eschenbach, C., & Habel, C. (2011). Linking spatial haptic perception to linguistic representations: Assisting utterances for Tactile-Map explorations. In M. Egenhofer, N. Giudice, R. Moratz, & M. Worboys (Eds.), *Spatial information theory* (pp. 328–349). Berlin, Heidelberg: Springer.
- Lohmann, K., & Habel, C. (forthcoming). Extended verbal assistance facilitates knowledge acquisition of virtual tactile maps. Accepted for presentation at Spatial Cognition 2012.
- Lohmann, K., Kerzel, M., & Habel, C. (2010). Generating verbal assistance for Tactile-Map explorations. In I. van der Sluis, K. Bergmann, C. van Hooijdonk, & M. Theune (Eds.), *Proceedings of the 3rd workshop on multimodal output generation 2010*. Dublin.
- Lynch, K. (1960). *The image of the city*. Cambridge, MA; London: MIT Press.
- Miele, J. A., Landau, S., & Gildea, D. (2006). Talking TMAP: automated generation of audio-tactile maps using Smith-Kettlewell's TMAP software. *British Journal of Visual Impairment*, 24(2), 93–100.
- Parkes, D. (1988). "NOMAD": An audio-tactile tool for the acquisition, use and management of spatially distributed information by partially sighted and blind people. In *Proceedings of the 2nd international conference on maps and graphics for visually disabled people*. Nottingham, UK.
- Parkes, D. (1994). Audio tactile systems for designing and learning complex environments as a vision impaired person: static and dynamic spatial information access. *Learning Environment Technology: Selected Papers from LETA*, 94, 219–223.
- Reiter, E., & Dale, R. (2000). *Building natural language generation systems*. Cambridge: Cambridge University Press.
- Rich, C., Sidner, C., & Lesh, N. (2001). Collagen: applying collaborative discourse theory to human-computer interaction. *AI magazine*, 22(4), 15–26.
- Roy, D., & Reiter, E. (2005). Connecting language to the world. *Artificial Intelligence*, 167(1-2), 1–12.
- Ungar, S. (2000). Cognitive mapping without visual experience. In R. Kitchin & S. Freundschuh (Eds.), *Cognitive mapping: Past, present and future* (pp. 221–248). London: Routledge.
- Wang, Z., Li, B., Hedgpeth, T., & Haven, T. (2009). Instant tactile-audio map: enabling access to digital maps for people with visual impairment. In *Proceeding of the 11th international ACM*

SIGACCESS conference on computers and accessibility (pp. 43–50). Pittsburg, PA.

Zeng, L., & Weber, G. (2010). Audio-haptic browser for a geographical information system. In K. Miesenberger, W. Zagler, & A. Karschmer (Eds.), *Computers helping people with special needs, part II* (pp. 466–473).