Verbally Assisted Haptic Graph Comprehension: The Role of Taking Initiative in a Joint Activity

Özge Alaçam
Department of Informatics
University of Hamburg
Hamburg/Germany
alacam@informatik.uni-hamburg.de

Christopher Habel
Department of Informatics
University of Hamburg
Hamburg/Germany
habel@informatik.uni-hamburg.de

Cengiz Acartürk
Cognitive Science
Middle East Technical University, Ankara/Turkey
acarturk@metu.edu.tr

Abstract

Statistical graphs are tools for multimodal communication in daily life settings. For visually impaired people, haptic interfaces provide perceptual access to the information provided by the graph. Haptic comprehension is facilitated by audio and verbal assistance. We investigate the circumstances under which verbal assistance facilitates haptic comprehension of graphs. For this, we focus on cases where unaided haptic graph comprehension has limitations, such as when describing a global (rather than a local) maximum. In an experiment that employed a joint activity setting, we observed two major aspects of verbal assistance for haptic graph comprehension, as indicated by post-exploration sketch production of the participants: First, the results revealed significant impact of the explorer as a dialogue initiator during the course of haptic exploration. Second, verbal assistance led to more successful graph comprehension when it was enriched by modifiers.

Keywords: haptic line graph comprehension; sketching, verbal assistance, turn-taking

1 Haptic Audio Line-Graph Exploration

Presenting and representing information in visuo-spatial formats, such as graphs, maps or diagrams, is important, as well as successful, for thinking, problem solving and communication (Hegarty, 2011). Their usage covers science and education settings, and also the news media and economy bulletins. There have been continuous efforts for the inclusion of blind and visually impaired people in using these visuo-spatial representations. Users can explore haptic graphs (Fig. 1.a) by hand movements following graph lines engraved in a physical plane (Fig. 1.b) or by using a force-feedback device, for instance a Phantom Omni® (Fig. 1.c), to explore virtual graphs lines, i.e. lines engraved in a virtual plane. Comprehension of haptic line graphs is based on exploration processes with the goal to collect information provided by the geometrical properties of the line explored; in particular, shape properties and shape entities—as concavities and convexities, maxima and minima, corners and smooth turning points—have to be detected.

(a) (b) (c)

Figure 1: (a) Sample haptic graph, (b) Exploration of a physical haptic map and (c) Phantom Omni® device and visualization in the domain of geometry

In addition to pure haptic graphs, in the past decade, haptic-audio interfaces have been developed to provide perceptual access to spatial representations, thus facilitating comprehension of spatial displays by visually impaired (see, e.g. Yu and Brewster, 2003; Zhao et al., 2008; Abu Doush et al., 2010). But there still remains much need for further development of specific types of haptic-spatial interfaces and the need for research that focus on peculiar characteristics of the interface design.

Statistical graphs do not only present data but also provide perceptual access to second-order entities, such as extrema, trends and trend changes. In particular, the properties of the line shape allow distinguishing a global maximum from a set of local maxima, or detecting inflection points that depict trend changes. Haptic graphs may be conceived as efficient interfaces in providing access to those spatial and conceptual structures through haptic sensory modality. On the other hand, haptic information intake has a lower bandwidth compared to visual information intake, since haptic exploration is sequential, while visual perception allows the perception of both local and global information about graph at one glance. In order to bridge this bandwidth gap between haptic exploration and visual exploration of graphs, and in order to provide sufficient information access to haptic graph readers, haptic graphs should be accompanied by alternative modalities, such as modalities that provide additional verbal or audio information. Accordingly, a challenge is faced for designing haptic graphs: It is necessary to determine which information depicted by the graph—or by the segments of the graph—are appreciated as *important*. This challenge becomes significant when designing haptic line graphs with several local maxima, as opposed to simple graph lines with a single global maximum, because in contrast to visual exploration, a local maximum cannot be recognized as a global maximum during the course of haptic exploration.

We propose that verbal assistance may facilitate overcoming the problem of distinguishing local maxima from the global maximum, and similar problems to the local maxima problem, by providing the necessary information through the auditory channel. To investigate the circumstances under which verbal assistance facilitates haptic comprehension of graphs, we designed an experimental setting, in which two participants perform joint activity for graph exploration, thus performing verbal assistance for haptic graph exploration (Alaçam et al, 2013a). It is a task-oriented joint activity (Clark, 1996) of two agents, a (visually impaired or blindfolded) explorer (E) of a haptic graph and an observing assistant (A) providing verbal assistance, as depicted in Figure 2.

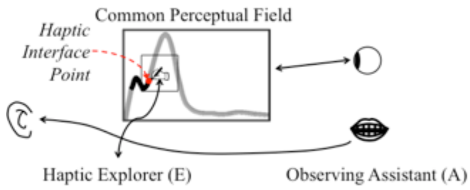


Figure 2: Assisted haptic graph exploration, a joint activity

A successful communication through graphs and language usually requires the integration of information contributed by both graphical entities and verbal entities so that the reader arrives at integrated conceptual and spatial representations. We have investigated various aspects of such integrated conceptual and spatial representations both from a theoretical perspective and in empirical studies (Habel and Acartürk, 2007, Acartürk, 2010). In the present paper we focus on haptic graph exploration as a collaborative activity between two humans, A and E: They share a common field of perception, namely the haptic graph, but their perception and comprehension processes differ significantly. For example, while (visually impaired or blindfolded) E explores the highlighted, black segment of the haptic graph (Fig. 2), A perceives the global shape of the graph. In particular, A is aware of shape landmarks and line segments. Similarly, when E starts exploring the first local maximum followed by a local minimum, E does not have any information about the global maximum, which is already part of A's knowledge. Thus, the haptic explorer E and the assistant A have different internal representations of the graph line, and A's referring to the graph could augment E's internal model substantially. At this moment (corresponding to the position depicted in Fig. 2) A can take the initiative, and starts uttering "Now you have reached the heights of the last peak" to provide E with additional information. Another suitable comment would be "You are in the increase to the population maximum", or even "You are in the increase to the population maximum of about 90, that was reached in 1985". Alternatively, E can take the initiative before A, by asking for advice. Initiating a dialogue has a significant role in constructing the alignment in the dialogue because it aims at making explicit the missing information (or information that is difficult to comprehend during the course of exploration), which is necessary for achieving the alignment. The joint activity in our case can be considered as an asymmetric dialogue, namely haptic-explorer-dominant, because the participants in the experiment (see below) were instructed that for efficient and effective verbal assistance systems, the haptic explorer initiates the help request and the verbal assistant provides help based on the explorer's need. This asymmetry is due to the asymmetry of participants' roles during the course of joint activity: assisted person (E) – assisting person (A) [corresponding, e.g. to 'stranger' – 'local' in route-instruction dialogues].

The success of the joint activity of the explorer E and the observing assistant A in general, and in particular the success of their task-oriented dialogue, depend on the alignment of the interlocutor's internal models, especially on building implicit common ground (Garrod and Pickering, 2004). E's internal model of the activity space, i.e. the haptic graph and E's explorations, is perceived via haptic and motor sensation, whereas A's internal model of the same space is built up by visual perception. Therefore similarities and differences in their conceptualization play a central role in aligning at the situation-model level. To be assistive, A should provide E verbally with content which is difficult to acquire haptically. This—haptically difficult to be built up—content has to be combined with haptically-explored content in the same sentence (or phrase) to fulfill the given-new contract (Clark and Haviland, 1977). In our study, verbal assistant is expected to provide most helpful and relevant information for haptic explorer at that particular moment among all-possible information that can be derived from the representation, by taking into account haptic explorer´s previous actions on the graph and previous utterances. The motivation that underlies this expectation is that the content of the verbal assistance has the potential to influence the alignment process, thus leading to a better or worse comprehension of the haptic graphs. As described in more detail below, we focus on the role verbal assistance content, as well as the role of dialogue initiating in the present study.

In the investigation of haptic graph comprehension, we have employed various methodologies including analysis of gestures, referring expressions and haptic exploration movements. Our previous research (Alaçam et al., 2013b) on accompanying gestures produced during verbal description of haptic graphs showed that in haptic graph comprehension, speech-accompanying gestures are usually produced with expressions that highlight relevant shape properties. Moreover, the analysis of referring expressions in the dialogues provides insight about how the haptic explorer comprehend the data, parse it for naming, and recognize which graphical elements are salient and which are hard to distinguish from the others. In addition to the analyses of gesture and referring expression production, sketch analysis of explored graphs is an appropriate methodology to evaluate the conceptualization of the events represented by the graph. As stated by Tversky (1999), "drawings reveal people's conceptions of things, not their perceptions of things". In particular, sketches provide complementary data for the analysis, because they can reveal details that the graph reader skips in verbal description for various reasons (e.g., the concept may be hard to express verbally or it may be considered as redundant by the reader). Accordingly, we employ the analysis of post-exploration sketches in the present study.

To sum up, we focus on two aspect of designing verbal assistance for haptic exploration of line graphs: (1) the role of the haptic explorer as dialogue-initiator (or no-initiator) (2) the role of the content of the verbal assistance. The empirical study presented in the paper was conducted with blindfolded participants to investigate modality-dependent differences by keeping other variables (i.e. the modality used to acquire prior graph-domain knowledge) constant and also to clarify which pieces of content should be provided by the modalities.

2 Experiment

In order to investigate the contribution of verbal assistance in haptic graph comprehension, we conducted the experiment in two conditions and performed a comparative analysis of the results. The first condition examined haptic exploration of line graphs by single, blindfolded participants, in the absence of verbal assistance. In the second condition, participant pairs (a blindfolded haptic explorer and a verbal assistant who was able to observe the haptic exploration) collaborated in exploring the haptic line graphs of Condition 1. In both conditions, each single session took approximately one hour. Haptic explorers were presented a warm-up session to get familiarized with the equipment (in this case, the Phantom Omni®, Fig. 1). Then they were presented the stimuli. The stimuli included five haptic line graphs with smooth edges (Fig. 3, two additional graphs were employed for familiarization with haptic line graphs). The participants were informed that they were presented bird-population graphs. The graphs were presented in randomized order. During the course of the experiment session, the explorer-participants performed haptic graph exploration by moving the handle of the haptic device, which can be moved in three spatial dimensions (with six degree-of-freedom). The line (proper) in the graphs was represented by engraved concavities on a horizontal plane; therefore the graph readers perceived the line as deeper than the other regions of the surface. Due to the interface limitations in the haptic representation, the numerical data labels were not presented. After the exploration of each graph, the participants produced single-sentence verbal descriptions of the graphs to a hypothetical audience; their spontaneous speech accompanying gestures produced during verbal descriptions were also recorded (these data are not in the focus of the present study). After the verbal description, they produced a sketch of the graph with paper and pencil. Two raters scored the sketches (all raters were blind to the goals of the study) for their similarity to the stimulus-graphs by using a 1 (least similar) to 5 (most similar) Likert Scale. The line graph stimuli were selected to represent a variety of patterns in terms of the number and polarity of curvature landmarks, length and direction of segments (Fig. 3).

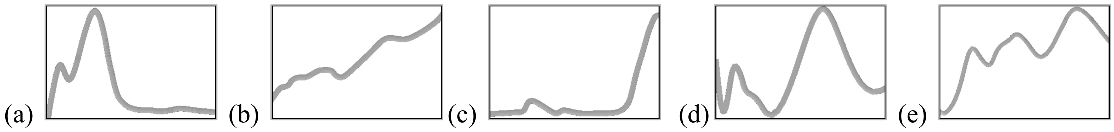


Figure 3: Stimulus-graphs (from graph-a to graph-e)

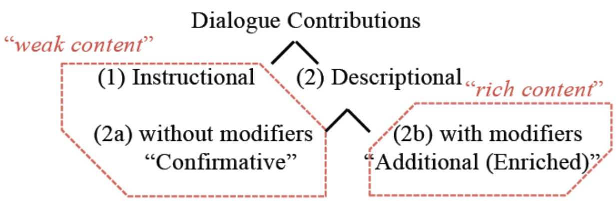
In Condition 1, nine university students (four female, *Mean age* = 25.0, *SD* = 6.3) performed haptic exploration of the stimuli without verbal assistance, and they completed the experimental protocol presented above. The participants of Condition 2 were pairs of students (13 pairs, 11 female, *M*=25.3, *SD*=3.27). Each pair was composed of a haptic explorer (E) and a verbal assistant (A). The haptic explorers were informed that the goal of the study was to design efficient and effective verbal assistance systems and the haptic explorer initiates the help request. The task of the verbal assistant was to provide the necessary information in a short description when required by the haptic explorer. The haptic explorer and the verbal assistant were located in separate rooms, so they communicated through speakers. While E explored the graph haptically throughout the experimental session, A had visual access to both the graph and E's exploration process (i.e. the current position E on the graph), which was displayed, on the computer screen together with the graph. Both E and A followed the experimental protocol employed in Condition 1 (i.e., a single-sentence verbal summary and sketch production). The results of Condition 2 showed that out of 65 experimental protocols of graph stimuli, 28 protocols involved at least one request from the verbal assistant. In other words, the haptic explorers were dialogue initiators in 28 of 65 experiment protocols of the graph stimuli. Since this corresponds to almost half of the pairs in Condition 2, we decided to conduct a further analysis of data by dividing the protocols of Condition 2 into two groups: (1) Dialogue-initiated protocols (henceforth, dialogue-initiator protocols) (2) The protocols that involved no dialogue initiative, thus no verbal assistance (henceforth, no-initiator protocols). In this study, we focused on the content of the dialogue and the similarity of participants' sketches to the stimulus-graphs as a performance measure. For the analysis of the sketches, inter-rater reliability between the two raters was assessed using a two-way mixed, consistency average-measures ICC (Intra-class correlation). The resulting ICC was in the "good range", identified by ICC=.69 (Cicchetti, 1994).

2.1 Utterances by the Verbal Assistant

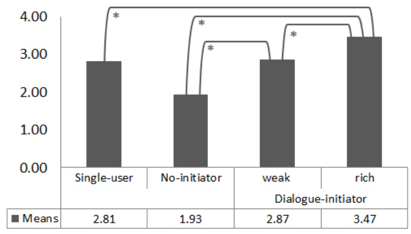
Providing information in response to E´s question or statement has the potential to enhance E's comprehension of the graph at that particular moment of verbal assistance. In other words, E makes an inference by using the information provided by A's utterance and forms a potentially more correct representation of the graph or a graph segment by combining this information with his/her exploration. During the course of the dialogue between the haptic explorer and the verbal assistant, each of the explorer's utterances can affect the explorer's mental representation of the graph. The assistant's contributions to the dialogue can be classified as follows: (1) instructional (i.e. navigational, such as 'go downward from there'), or (2) descriptional. Descriptional utterances include, (2a) confirmative assistance (specifying exploration events or graph entities without using modifiers - such as 'there is a decrease'), and (2b) additional assistance (specifying properties of exploration events or graph entities using modifiers, such as 'there is a steep decrease'). Based on this scheme (see Fig. 4a), we classified the dialogues into two major groups. Firstly, we identified *weak content* dialogues, which were less informative. These were the dialogues that contained assistance focused on (or restricted to) 'basic spatial properties' of the currently-explored region (i.e. the location or polarity of the graph segments). Secondly, we identified *rich content* dialogues, in which the verbal assistant also provided additional properties about the region explored (e.g., information about the steepness or length of the graph segments). Two coders performed the classification task. Interrater reliability between coders was calculated by Cohen's kappa. The results revealed a value of .80 that indicates substantial interrater agreement.

2.2 Analysis of Post-Exploration Sketches

We analyzed participants' sketches in terms of their similarity to the stimulus-graphs. A statistical analysis using Kruskal-Wallis test revealed a significant difference between the ratings, $\chi 2(2, N=108)=23.3$, $p<.01$, among single-user protocols, no-initiator protocols, and dialogue-initiator protocols. Post-hoc testing of contrast using Mann-Whitney by using Bonferroni correction (so all effects are reported at a .0167 level of significance) showed that the sketches in no-initiator protocols ($M=1.93$, $SD=0.90$) received lower similarity scores both compared to the sketches in the single-user protocols ($M=2.81$, $SD=1.16$), $U=410.0$, $p<.01$ and the dialogue-initiator protocols ($M=3.17$, $SD=1.14$) $U=170.5$, $p<.01$, without a significant difference between the latter two. A further Mann-Whitney test (with Bonferroni correction) was conducted by taking into account the information content of the utterances. The results showed that the utterances that contained *rich content* resulted in higher similarity scores for the sketches ($M=3.47$, $SD=.72$) in the dialogue-initiator protocols than the sketches in the single-user protocols ($M=2.81$, $SD=1.16$), $U=236.50$, $p<.05$ and the other conditions, see Fig. 4b. This indicated that the dialogues that contained more specific information (such as slight increase, biggest curve etc.) resulted in better sketch production as an indicator of more complete conceptualization of the event, see Fig. 5 for sketch samples.



(a)



(b)

Figure 4: (a) Classification scheme for dialogue contribution and
(b) Ratings for sketches in five point Likert Scale (1: least similar and 5: most similar)

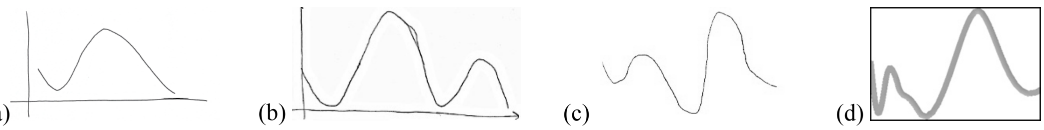

(a) (b) (c) (d)

Figure 5: Sketches after the protocols (a) without verbal assistance, (b) with *weak content* verbal assistance (c) with *rich content* verbal assistance and (d) Haptic graph-stimuli.

3 Discussion

In order to generate automatic verbal assistance with adequate content, it is necessary to identify the types and roles of individual utterances, as well as the structure of the dialogue content by using empirical studies. We conducted an experiment with two-conditions to investigate the contribution of verbal assistance in haptic graph exploration. The results of the first condition (single–user protocol without verbal assistance) and the second condition in two protocols (no-initiator protocols without assistance and dialogue-initiator protocols with assistance) were compared in terms of the analysis of sketches produced by the participants. High similarity ratings of the sketches were correlated with the richness of the content provided by verbal assistance. The sketches for dialogue-initiator protocols were significantly more similar to stimulus-graph compared to single-user and no-initiator protocols. The results also demonstrated considerable effect of verbal assistance content. In other words, the dialogues that contained modifiers (cf. *rich content*) were helpful to the explorer. Modifier presence made the assistance more elaborate; it helped the participant to notice the features of the event, which were currently explored (e.g. steepness of the curve and length, relation with another curve).

4 Conclusion

In this paper, we have investigated the role of the haptic explorer as dialogue-initiator (or no-initiator) and the role of the content of the verbal assistance in a collaborative activity. Although haptic explorer and visual assistant share a common field of perception, their perception and comprehension processes differ significantly. The empirical results pointed out that haptic graph readers benefit from the verbal assistance to achieve more successful conceptualization of the events that are represented by graph lines. This is because the verbal assistant has a more complete mental representation of the graph (both global and local information on the graph) from the onset of the partner's haptic exploration. The results also revealed valuable insights about how the comprehension is affected by the provided language content. Combining these information with haptic exploration patterns before/during/after the explorer's help request will provide a concrete base to build up automatic detection of the need for verbal assistance. The detection of what a graph reader wants to know at a particular time during the course of exploration, by means of the analysis of his/her current position of exploration, previous exploration movements, and referring utterances (the referred locations and how these regions were referred) would yield a more effective design of (learning) environment for the graph reader compared to presenting all possible information to the graph reader at once. In addition, the information content and the need highlighted by the assistance request is another crucial topic that we addressed in this study. To sum up, our results indicated that taking initiative for requesting help and having adequate verbal assistance enriched by modifiers, rather than just confirmation of the basic spatial properties in response, seems a superb combination for a successful joint activity that inherently requires asymmetric dialogues between two users with different roles; haptic explorer and verbal assistant. Future work will address designing an automated verbal assistance based on the experimental findings.

Acknowledgments. The study reported in this paper has been partially supported by DFG (German Science Foundation) in ITRG 1247 'Cross-modal Interaction in Natural and Artificial Cognitive Systems' (CINACS).

Reference

Iyad Abu Doush, Enrico Pontelli, Tran C. Son, Dominic Simon, and Ou Ma. 2010. *Multimodal presentation of two-dimensional charts: an investigation using Open Office XML and Microsoft Excel*. ACM Transactions on accessible computing, 3(2), Article 8.

Cengiz Acartürk, C. 2010. *Multimodal comprehension of graph-text constellations: An information processing perspective*. Doctoral dissertation, University of Hamburg, Hamburg.

Özge Alaçam, Christopher Habel, and Cengiz Acartürk. 2013a. *Towards designing audio assistance for comprehending haptic graphs: A Multimodal Perspective*. In Constantine Stephanidis and Margherita Antona (Eds.), Universal Access in Human-Computer Interaction. Heidelberg: Springer.

Özge Alaçam, Christopher Habel, and Cengiz Acartürk. 2013b. *Investigation of haptic line-graph comprehension through co-production of gesture and language*. Tilburg Gesture Research Meeting, Tilburg.

Domenic V. Cicchetti. 1994. *Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology*. Psychological Assessment, 6(4), 284–290.

Herbert H. Clark, and Susan E. Haviland. 1977. *Comprehension and the given-new contract*. In Roy O. Freedle (Ed.), Discourse production and comprehension. Hillsdale, NJ: Erlbaum.

Herbert H. Clark. 1996. Using language. Cambridge: Cambridge University Press.

Simon Garrod, and Martin J. Pickering. 2004. *Why is conversation so easy?* Trends in Cognitive Sciences, 8. (1), 8-11.

Christopher Habel, and Cengiz Acartürk. 2007. *On reciprocal improvement in multimodal generation: Co-reference by text and information graphics*. In Ielka van der Sluis, Mariet Theune, Ehud Reiter and Emiel Krahmer (Eds.), Proceedings of Workshop on Multimodal Output Generation, 69-80.

Christopher Habel, Özge Alaçam, and Cengiz Acartürk. 2013. *Verbally assisted comprehension of haptic line-graphs: referring expressions in a collaborative activity*. In Proceedings of the CogSci 2013 Workshop on Production of Referring Expressions, Berlin.

Mary Hegarty. 2011. *The cognitive science of visual-spatial displays: implications for design*. Topics in Cognitive Science, 3, 446–474.

Barbara Tversky. 1999. *What does drawing reveal about thinking?*. In J.S. Gero, B. Tversky (eds.) Visual and Spatial Reasoning in Design (VR99), Key Centre of Design Computing and Cognition, University of Sydney.

Wai Yu, and Stephen A. Brewster. 2003. *Evaluation of multimodal graphs for blind people*. Journal of Universal Access in the Information Society, 2, 105-124.

Haixia Zhao, Catherine Plaisant, Ben Schneiderman, and Jonathan Lazar. 2008. *Data sonification for users with visual impairment: a case study with georeferenced data*. ACM Transactions on Computer-Human Interaction (TOCHI), 15(1).