Towards Designing Audio Assistance for Comprehending Haptic Graphs: A Multimodal Perspective

Özge Alaçam[1], Christopher Habel[1], and Cengiz Acartürk[2]

[1] Department of Informatics, University of Hamburg, Hamburg/Germany
{alacam,habel}@informatik.uni-hamburg.de
[2] Informatics Institute, Middle East Technical University, Ankara/Turkey
acarturk@metu.edu.tr

Abstract. Statistical graphs, such as line graphs are widely used in multimodal communication settings. Language accompanies graphs and humans produce gestures during the course of communication. For visually impaired people, haptic-audio interfaces provide perceptual access to graphical representations. The local and sequential character of haptic perception introduces limitations in haptic perception of hard-to-encode information, which can be resolved by providing audio assistance. In this article we first present a review of multimodal interactions between gesture, language and graphical representations. We then focus on methodologies for investigating hard-to-encode information in graph comprehension. Finally, we present a case study to provide insight for designing audio assistance.

Keywords: Haptic Graph Comprehension, Audio-Verbal Assistance.

1 Graph Perception and Graph Comprehension

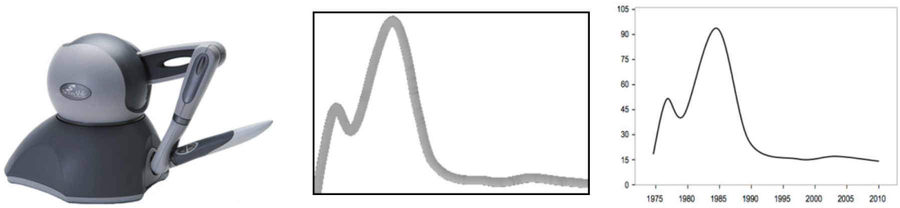
1.1 Presenting Graphs in Different Modalities: Visual vs. Haptic

The primary goal of visualizing data is to (re-)present them in a pictorial format more suitable for using them in thinking, problem solving and communication, namely in the representational modality of graphs [1, 2, 3, 4, 5]. Graphs are successful means to present data, both in tasks of analyzing data and in tasks of communicating data. Communicating visualized data using graphs is used extensively in different types of publications, from scientific journals and textbooks to magazines and newspapers. Line graphs and bar graphs are the dominant, i.e. most frequently used, types of graphs in addressing non-experts in communication through graphs [6]. In addition to text-graphics documents, in many professional communication and classroom settings, graphs, language, and often gestures, accompany each other forming multimodal communication.

Since for blind and visually impaired people the advantages of graphs are not directly accessible, haptic-audio interfaces to graphical representations have been proposed for—partially—substituting vision in the use of graphs [7, 2, 8] and other types of graphics, such as maps, floorplans etc. [9, 10]. Whereas visual perception

C. Stephanidis and M. Antona (Eds.): UAHCI/HCII 2013, Part I, LNCS 8009, pp. 409–418, 2013.

supports comprehension processes which switch between global and local aspects of a graphical representation, haptic perception has a more local and in particular a more sequential character. Thus, compared to visual graphs, one drawback of haptic graphs is the restriction of the haptic sense regarding the possibility of simultaneous perception of information [11]. Comprehension of haptic line graphs is based on exploration processes, i.e. hand movements following the lines with the goal to summarize information of geometrical properties of the line explored; in particular, the detection of shape properties—as concavities and convexities, as well as maxima and minima—are of major importance, see Fig. 1 depicting the Phantom Omni® Haptic Device we use in our studies, as well as an exemplifying haptic line graph and its visual counterpart.



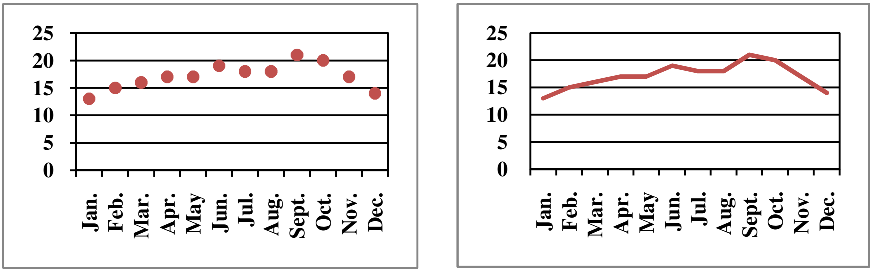
(a) (b) (c)

Fig. 1. (a) Phantom Omni® Haptic Device, sample (b) haptic graph and (c) visual graph

Whereas it is relatively unproblematic to detect haptically the shape of simple graph line with only a single global maximum, graphs with several local maxima require—depending on their complexity—additional assistance for most users of haptic graphs. For resolving some difficulties in haptic exploration of graphs, providing additional information, such as auditory assistance through the auditory channel, has been proved to be helpful [8]. Sonification or speech can support—for example—the detection of local and global extrema of graph lines. The usage of these alternative modalities with the haptic modality arises new research questions, namely, which pieces of content to be presented should be made accessible through the haptic modality, and which pieces should be communicated using language (speech) or sounds (sonification).

1.2 Data Visualization: Visual vs. Haptic

A standard starting point for generating graphs are tables or their computer-science counterpart, relational data-bases [12]. But, visualizing exclusively data points is suitable only in certain cases (see Fig. 2(a), depicting "average daily maximal temperature at San Francisco", data from [13]). Under specific conditions, e.g., if some Gestalt principles are satisfied, human visual processing, leads pre-attentively, to the visual impression of a linear whole, namely a line. Fig. 2(b) depicts a line graph that relieves the perceptual and cognitive load by making the line explicit. This line-graph—independent of whether data points are visually depicted or not—contains elements of the line, which have no origin in the data. The contrast between data-point graphs and line graphs exemplifies how substantially the human perceptual system determines the comprehension of data visualizations. Blind and visually impaired people who can use some types of haptic graphs successfully would have critically more problems in exploring data-point graphs—as depicted in Fig. 1(a)— haptically using an Omni device, since in this case Gestalt constitution is not supported.



(a) (b)

Fig. 2. Average daily maximal temp at San Francisco (a) Data-point graph (b) Line graph

Line graphs, for example, present not only data points explicitly, but in addition, second-order entities, such as *trends*, *local* and *global maxima*, second order properties (e.g., *strength of an upward trend*) and second-order relations (*crossing of lines*) can be also easily detected due to preattentional processes of the human visual system [2]. To design audio-haptic interfaces to graphs, it is important to know, which concepts depicted in graphs, are most relevant to people during graph comprehension. In prior studies we used successfully three empirical paradigms to get insights in human's conceptualizing of graphs: (1) eye tracking in exploring language-graph co-comprehension [1, 14], (2) sketching of graphical cues on graphs under linguistic guidance of written verbal descriptions [15], and (3) combined analyses of gesture patterns and of eye movements during producing spoken descriptions of graphs to investigate the role of gestures in graphical communication [16].

2 Graph Comprehension in Multimodal Contexts

2.1 Language - Graph Comprehension

Graphs are usually accompanied by language in communication settings, either in spoken form, as in presentations or lectures or in written form, as in newspapers, magazine articles and web blogs. In all those settings, a successful communication through graphs and language usually requires the integration of information contributed by both modalities. The pivot of this integration is the construction of reference and co-reference relations between text, the elements of the graph and the entities in the domain of discourse that are referred to by text and the graph [17]. For instance, a specific linguistic constituent, such as “the peak” in the text may refer to a peak in the domain of discourse (e.g., a peak in temperature), which is also referred to by a peak as a graphical entity in the graph. We have investigated various aspects of such integrated conceptual and spatial representations both from a theoretical perspective and in empirical studies [1, 14, 20, 33].

Text and graphical constituents appear in various forms in text-graphics documents. In most cases, the graph and the text are placed side-by-side on the page, thus leading to a separate layout of graph and text. To facilitate the construction of referential links between the text and the figure, different types of cross-reference link are used in the text to provide a link to the graphical material, thus providing a specific type of signaling in verbal form. Figure captions are usually referred to by those explicit cross-reference links, such as "see Figure 1" in the text, thus facilitating the integration of information contributed by the text and the figure [18]. The use of graphical cues on the graph frame, such as an arrow to emphasize an increase or a point-like marker to emphasize a peak, provide another signaling technique that aims to attract the attention of the reader and facilitate the construction of thematic relevance between the information contributed by the text and the information represented by the graphical entities [19]. In addition to those signaling methods, verbal elements, such as verbal annotations on the graph frame may have a facilitating role in comprehension of the text-graphics document [20]. An appropriate methodology for investigating the construction of referential links between language and graphs is the eye-tracking methodology. Eye tracking has been one of the techniques that provide comprehensive information about online cognitive processes of a graph reader since it lets to trace the allocation of attention. The previous empirical studies showed that eye movement parameters such as gaze pattern, transition matrix entropy and fixation rates are useful tools to investigate graph-language comprehension [21].

2.2 Gesture – Graph Comprehension

The studies on language-gesture interaction from the embodied cognition viewpoint are mainly based on the assumption that concepts are sensorimotor, emphasizing that they are based on perceptual experience [26, 27]. If the concepts are concrete and easy to visualize the speaker gestures more; even abstract concepts are grounded in physical terms [25]. There are several frameworks that investigate gestures from various perspectives, but all of them agree on that gestures rely on spatial representations. According to the GSA (Gesture as simulated action) framework [25], one of the frameworks that focus on how gestures are produced, gestures are byproduct of speech. In particular, linguistic planning involves simulation of visuo-spatial events; this activation during articulation is considered as a source of speech accompanying gestures. Another framework, that is closely aligned with the GSA framework and that focuses on how gesture and language production are integrated is the "Interface Hypothesis" [28]. The preparation for language production requires organization of rich and comprehensive information into small packages that contain appropriate amount of informational complexity within a processing unit. According to the "Interface Hypothesis", this processing unit may correspond to a clause for speech production, and the contents of a representational gesture are affected by the organization of these information-processing units, which are prepared for speech production. Therefore this close relationship between gestures and language makes gestures an effective tool in the assessment of the reader`s conceptualizations by means of the analysis of verbal descriptions [29].

Although the interaction between language and gesture has been investigated for the past several decades in a variety of domains [22, 23, 24, 25], specific investigations of graph comprehension—be it based on the visual or the haptic modality—in interaction with language and gesture, has been one of the scarce topics in HCI and relevant disciplines. Gestures provide additional information that enhance comprehension and resolve ambiguities during the course of communication. They are convenient tools to carry spatio-temporal information. Besides, they highlight the information presented by the other modality and convey additional information that is not expressed by the other modality. For instance, within the context of communication through graphs, a fluctuating increase in a line graph may be verbally described by the term "increase" and it may be simultaneously accompanied by a gesture that represents the fluctuation in the increase [16]. Based on these similarities, Tversky [30] proposed that the vocabularies between these three modalities (graphs, language and gesture) can be considered as parallel. One of the studies focused on communication through line graphs [16] showed that the perceptual features of the graphical cues that highlight certain aspects of the visualization (e.g., a graphical cue such as an arrow) influence conceptualization of presented information, and this effect is observable in the gestures produced by graph readers. The results of that study indicated that in order to emphasize process concepts (e.g., *increase*, *decrease*) more vertical and diagonal gestures were produced by humans, whereas more pointing gestures were produced for emphasizing punctual state concepts (e.g., a *peak*). Similar findings were obtained from a comprehension perspective, in the sense that vertical and diagonal gestures were efficient in conveying information about processes. The findings also revealed a low efficiency of non-directional gestures in conveying punctual state information, possibly due to ambiguities between pointing-as-representational gesture and pointing-as-deictic gesture. Analysis on eye movements showed that participants also exhibit back and forth eye movement between the gestures and the face of the narrator, indicating potential source of attention split during the course of comprehension.

To sum up, gestures can be considered as a tool to assess how the graph reader interprets the graph and conceptualizes the processes, events and states represented by the graph, because gestures provide additional information, which is aligned with the visuo-spatial aspects of communication through graphs. Therefore gesture analysis helps to detect the hard-to-encode information and disambiguates, that are generally highlighted with the presence of accompanying gestures.

3 A Case Study on Haptic Graphs and Design Guidelines

Comprehension of haptic graphs is still one of the topics, which has not been comprehensively investigated so far [9, 31]. A systematic investigation of the interaction between modalities in communication through graphs has the potential to contribute to identifying design principles for multimodal communication settings that facilitate efficient and effective communication of information since experiments give evidence about the content relevant to the conveyed information (in particular the question what should be communicated by language). Our particular research focus in the case study [31] is the investigation of the characteristics of hard-to-encode information in graphical communication through gesture and language production. The motivation for analyzing hard-to-encode information is to identify the types of assistance that should be provided to ease comprehension by visually impaired users.

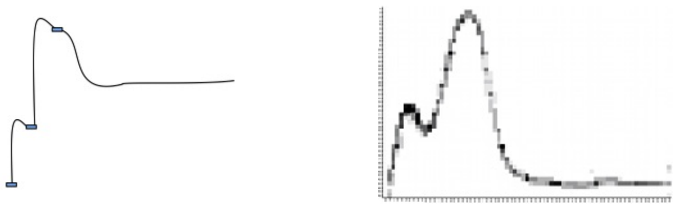
Participants, Materials and Design. The experiment was conducted in two conditions in a within subject design employing a total of twenty participants. In the first condition, the participants (N=9) explored line graphs haptically (see Fig. 1b). In the second condition, the graphs were presented on the computer screen and the participants (N=11) had visual access to the graphs (see Fig. 1c). In both conditions, after the participants explored the graph, they were asked to produce single sentence summaries of bird population graph to the hypothetical audience. As a result of this experimental paradigm, two different types of gesture production occurs; exploration gestures and communicative gestures [32]. The gestures produced during the haptic exploration of the graph are categorized as exploration gestures while the gestures produced during verbal description of the graph are classified as communicative gestures. In this study, Phantom Omni® Haptic Device (Fig. 1a) is used to represent the haptic line graphs. Haptic graph exploration with this device is performed by moving the handle of the haptic device, which can be moved in all three spatial dimensions (with six degree-of-freedom). In haptic graph representation, the graph proper (the line of the line graph) is represented by engraved concavities on a horizontal plane; therefore the graph readers perceive the line as deeper with respect to other area on the surface and trace the line haptically by moving the pen (Fig. 1b).

McNeill's [23] semantic gesture classification and then syntactic classification were used as a coding scheme. In the first classification, the gestures were categorized according to their semantic classifications, such as beat gestures and representational gestures. Then each representational gesture was classified in terms of its directionality: non-directional, and directional. According to this classification, the hand movements conducted in small space without having any directed trajectory were categorized as non-directional gesture, whereas the hand movements with aimed trajectory on the air were classified as directional gestures. Directional gestures were also classified into two categories; (i) single direction, and (ii) multiple directions. The gestures that contained movement in only one direction (such as upward) were classified under the "single direction" category, while category of "multiple directions" covers the gestures formed with the combinations of one-directional gestures (such as movement consisting of upward, downward and upward movements).

Findings. The results—focusing on the communicative gestures—revealed that the gestures produced during the course of verbal description were influenced by the modality (haptic versus visual) of the graphical representation. The results also indicated that graph readers tend to produce multiple directional gestures, as well as one-directional gesture when they explored the graph haptically. On the other hand, the exploration of visual graphs resulted in the production of more one directional gestures compared to multiple directional gestures. This alignment between the psychical properties of the haptic graph (Fig 1b) and communicative gestures produced during verbal description of the graph (see Table 1) is remarkable and the resemblance may be the possible source of this effect that we named as "multimodal carry-over effect". In addition to analysis of communicative gestures (Fig. 3a), the analysis of the relationship between communicative gestures and explorative gestures (Fig. 3b) is considered as another method that can reveal the underlying mechanisms of this effect. While the haptic exploration gestures (most focused regions, back and forth movements) provide evidence for the online cognitive processes in the course of graph comprehension, verbal descriptions and communicative gestures give valuable information about how the graph reader conceptualizes the events represented by the haptic graph. In Table 1 and Fig. 3, the sample verbal description and accompanying gestures for the bird population graph given in Figure 1 were presented in addition to heatmap representation for explorative gestures produced during haptic exploration.

Table 1. Sample Verbal Description produced in one of the protocols

Verbal Description *(translated from German)*
It had a small peak, than a large peak and somewhere in the midfield it levels off.



(a) *(b)*

Fig. 3. (a) Trajectories of the gestures accompany to this description (the rectangles indicate pauses on the gestural movement) (b) Heatmap for Haptic Exploration Gestures in Grayscale (darker areas represents more focused areas)

In addition to providing insight about how graph readers perceive and interpret the haptic graphs, the findings of the case study also give clues about how to utilize these results in the haptic graph design. In the current state of art of haptic line graph design, providing haptic information on the graph axes, such as information for numerical labels and axis title, is hard to implement. This is a potential source of difficulty in haptic comprehension, because the lack of axis information may lead to difficulties in comprehension. As can be seen from Fig. 3a and 3b, although the trend of the trajectory of produced gesture and the graph proper (the line) have quite resemblance, there existed misinterpretation about value, the number of bird population probably due to not having any reference to real values. Such difficulties are good candidates for substitution by other modalities such as verbal or audio assistance. Furthermore, since the visual perception of the graph lets the graph reader to access both local and global information, the information about the start and end point of the graph or steepness/amplitude of the peaks are already accessible to them. However, this information, not provided by haptic modality, is crucial for the haptic graph reader in order to create appropriate mental representation of the graph and content represented by it. The haptic exploration gestures are also used to obtain insight about how the graph reader perceives and comprehend these critical parts of the graphs.

4 Conclusion

To conclude, various modalities are intertwined in communication settings, including communication through line graphs. The investigation of gestures, eye movements, language and haptic exploration in interaction has the potential to provide insights for human interpretation of the represented information that has a direct contribution to multimedia design. One of the fields, profited by these researches, is multimedia design for blind and visually impaired people. The graphical representations are wide spread in both print and electronic media, and they are used as a basic material to elaborate the information, which is hard to express within text only. Therefore, to provide access to graphical representations for blind people is one of the important topics of this field. Haptic graphs are considered as an efficient medium that provides access to the visual representations presented through haptic modality. However haptic representation has lower bandwidth compared to visual modalities, since the haptic exploration is sequential, while visual perception allows the perception of both local and global information about graph at one glance. Therefore, visual representations can be considered as superior in the amount of conveyed information. In order to bridge this gap and present coherent information to the haptic graph readers, haptic graphs should be accompanied by alternative modalities such as verbal/audio modalities. In this study, we aimed to emphasize the multimodal method to be used in the investigation of haptic graph comprehension and detect the content to be provided by haptic modality and to be assisted using alternative modalities. The findings provide insight for the implementation of user interaction design for visually impaired people, by providing guidelines for the design of audiovisual assistance during the course of communication.

Acknowledgment. We thank Matthias Kerzel and our student assistants Gregor Gawellek, Lena Andreeßen and Neele Stoeckler for their valuable effort. The research reported in this paper has been supported by DFG (German Science Foundation) in ITRG 1247 'Cross-modal Interaction in Natural and Artificial Cognitive Systems' (CINACS).

References

1. Acartürk, C.: Multimodal Comprehension of Graph-text Constellations: An Information Processing Perspective. University of Hamburg Dissertation, Hamburg (2010)
2. Habel, C., Acartürk, C.: Towards Comprehension of Haptic Multiple-line Graphs. In: Goncu, C., Marriott, K. (eds.) Proceedings of the Diagrams Workshop on Accessible Graphics: Graphics for Vision Impaired People, Kent, UK (2012)
3. Hegarty, M.: The Cognitive Science of Visual-spatial Displays: Implications for Design. Topics in Cognitive Science 3, 446–474 (2011)
4. Kosslyn, S.M.: Understanding Charts and Graphs. Applied Cognitive Psychology 3, 185–226 (1989)
5. Tufte, E.R.: The Visual Display of Quantitative Information. Graphic Press, Cheshire CT (1983)
6. Zacks, J., Levy, E., Tversky, B.: &Schiano, D.: Graphs in Print. In: Anderson, M., Meyer, B., Olivier, P. (eds.) Diagrammatic Representation and Reasoning, pp. 187–206. Springer, London (2002)
7. Abu Doush, I., Pontelli, E., Simon, D., Son, T.C., Ma, O.: Making Microsoft Excel™ Accessible: Multimodal Presentation of Charts. In: Proceedings Eleventh International ACM SIGACCESS Conference on Computers and Accessibility, Pittsburg, PA, USA, pp. 147–154 (2009)
8. Yu, W., Brewster, S.A.: Evaluation of Multimodal Graphs for Blind People. Journal of Universal Access in the Information Society 2, 105–124 (2003)
9. Sjöström, C., Danielsson, H., Magnusson, C., Rassmus-Gröhn, K.: Phantom-based Haptic Line Graphics for Blind Persons. Visual Impairment Research 5, 13–32 (2003)
10. Yu, J., Habel, C.: A Haptic-Audio Interface for Acquiring Spatial Knowledge about Apartments. In: Magnusson, C., Szymczak, D., Brewster, S. (eds.) HAID 2012. LNCS, vol. 7468, pp. 21–30. Springer, Heidelberg (2012)
11. Loomis, J., Klatzky, R., Lederman, S.: Similarity of Tactual and Visual Picture Recognition with Limited Field of View. Perception 20, 167–177 (1991)
12. Mackinlay, J.D.: Automating the Design of Graphical Presentations of Relational Information. ACM Transactions on Graphics 5, 110–141 (1986)
13. Pearce, E.A., Smith, C.G.: Fodor's World Weather Guide. Random House, New York (1998)
14. Habel, C., Acarturk, C.: Causal Inference in Graph-Text Constellations: Designing Verbally Annotated Graphs. Tsinghua Science & Technology 16, 7–12 (2011)
15. Acartürk, C.: Points, Lines and Arrows in Statistical Graphs. In: Cox, P., Plimmer, B., Rodgers, P. (eds.) Diagrams 2012. LNCS, vol. 7352, pp. 95–101. Springer, Heidelberg (2012)
16. Acartürk, C., Alacam, O.: Gestures in Communication through Line Graphs. In: Miyake, N., Peebles, D., Cooper, R.P. (eds.) Proceedings of the 34th Annual Conference of the Cognitive Science Society, Austin, TX, pp. 66–71. Cognitive Science Society (2012)
17. Habel, C., Acartürk, C.: On Reciprocal Improvement in Multimodal Generation: Co-reference by Text and Information Graphics. In: van der Sluis, I., Theune, M., Reiter, E., Krahmer, E. (eds.) Proceedings of the Workshop on Multimodal Output Generation, MOG 2007, pp. 69–80 (2007)
18. Acartürk, C., Taboada, M., Habel, C.: Cohesion in Multi-modal Documents: Effects of Cross-referencing. Information Design Journal (in press)
19. Lowe, R., Boucheix, J.M.: Cueing Complex Animations: Does Direction of Attention Foster Learning Processes? Learning and Instruction 21, 650–663 (2011)

20. Acarturk, C., Habel, C., Cagiltay, K., Alacam, O.: Multimodal Comprehension of Language and Graphics: Graphs with and without Annotations. Journal of Eye Movement Research 1, 2 (2008)
21. Habel, C., Acartürk, C.: Eye-tracking Evidence for Multimodal Language-graphics Comprehension: The Role of Integrated Conceptual Representations. In: Proceedings of NODALIDA 2009: The Workshop on Multimodal Communication- From Human Behavior to Computational Models, vol. 6 (2009)
22. McNeill, D.: Hand and Mind: What Gestures Reveal about Thought. University of Chicago Press, Chicago (1992)
23. McNeill, D.: Gesture and Thought. University of Chicago Press, Chicago (2005)
24. Goldin-Meadow, S.: Hearing Gesture: How Our Hands Help Us Think. Harvard University Press, Cambridge (2003)
25. Hostetter, A.B., Alibali, M.W.: Visible Embodiment: Gestures as Simulated Sction. Psychonomic Bulletin and Review 15, 495–514 (2008)
26. Barsalou, L.W.: Perceptual Symbol Systems. Behavioral and Brain Sciences 22, 577–660 (1999)
27. Garbarini, F., Adenzato, M.: At the Root of Embodied Cognition: Cognitive Science Meets Neurophysiology. Brain and Cognition 56, 100–106 (2004)
28. Kita, S., Özyürek, A.: What Does Cross-linguistic Variation in Semantic Coordination of Speech and Gesture Reveal? Evidence for an Interface Representation of Spatial Thinking and Speaking. J. Memory and Language 48, 16–32 (2003)
29. Goldin-Meadow, S., Beilock, S.L.: Action's Influence on Thought: The Case of Gesture. Perspectives on Psychological Science 5, 664–674 (2010)
30. Tversky, B.: Visualizing Thought. Topics in Cognitive Science 3, 499–535 (2011)
31. Alaçam, Ö., Habel, C., Acartürk, C.: Investigation of Haptic Line-Graph Comprehension Through Co-Production of Gesture and Language. Tilburg Gesture Research Meeting, Tilburg (2013)
32. Quek, F., McNeill, D., Bryll, R., Duncan, S., Ma, X.F., Kirbas, C., Ansari, R.: Multimodal Human Discourse: Gesture and Speech. ACM Transactions on Computer-Human Interaction (TOCHI) 9, 171–193 (2002)
33. Acartürk, C., Habel, C.: Eye Tracking in Multimodal Comprehension of Graphs. In: Proceedings of the Workshop on Technology-Enhanced Diagrams Research, vol. 887, pp. 11–25 (2012), http://ceur-ws.org