Verbally Assisted Comprehension of Haptic Line-Graphs: Referring Expressions in a Collaborative Activity

Christopher Habel, Özge Alaçam
Department of Informatics, University of Hamburg, Germany

Cengiz Acartürk
Informatics Institute, Middle East Technical University, Turkey

Abstract

Graphs are successful means to present data, in both tasks that of analyzing and that of communicating data. In addition to text-graph documents, in many professional communication and instruction settings, graphs, language, and often gestures accompany each other forming multimodal communication. In the present study, we investigate referring expressions used in a collaborative joint activity (Clark, 1996), namely in acquiring knowledge from graphs and language. Our long term goal is to realize a cooperative system that provides blind graph readers with verbal assistance.

Verbally Assisted Haptic Graphs

For visually impaired people the haptic modality provides a suitable means to acquire knowledge from graphs. Comprehension of haptic line graphs is based on exploration processes. Using a force-feedback device, which yields information about geometrical properties of the lines, users explore haptic graphs by hand-controlling a stylus (Fig. 1).

Figure 1: Sample haptic graph and Phantom Omni® device

Whereas visual perception supports comprehension processes that switch between global and local aspects of a graphical representation, haptic perception has a more local and sequential character. In contrast to simple graph lines with a single global maximum, haptic exploration of graphs with several local maxima requires additional assistance for comprehension. Providing verbal assistance using the auditory channel has been demonstrated to be helpful (Yu & Brewster, 2003).

When users interact with a haptic graph by means of a haptic device interface, their exploration movements can be observed by the system's assistance module, in a similar way to a human observing a blind person's finger movements in exploring an embossed graph for giving additional assistive verbal information during the course of the knowledge acquisition task.[1]

Referential Links in Text-Graph Documents

In comprehending expository text-graph documents *reference links* and *co-reference links* of different type have to be established. We exemplify this by an excerpt from a waterbird census report[2]: For each observed bird species, the report contains a line graph (see Fig. 2) and accompanying text. The text covers five topics. One of these topics, *Bolinas Lagoon Population Trends*, summarizes the change in population. A sample excerpt is given in (1).

(1) *Bolinas Lagoon Population Trends*

From a peak of about 60 wintering birds in 1976, numbers have declined to about 20 birds currently.

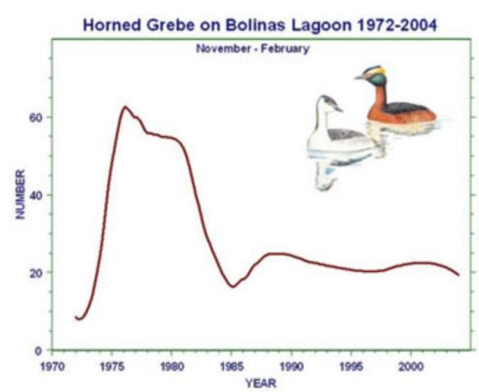

Figure 2: The population graph of wintering birds.

Referential expressions, such as 'peak of about 60', can be interpreted as referring to the domain of discourse (waterbirds at Bolinas Lagoon) or as referring to the domain of external representations, in particular the graph, which depicts some data about the domain of discourse. In processing documents in which two communicational modalities, such as text and graphics, contribute to a common conceptual representation, speaker and hearer have to establish *inter*- and *intra-representational coherence* by

[1] A parallel verbal-assistance approach in the domain of tactile maps has been realized successfully (Lohmann & Habel, 2012).

[2] "Waterbird Census at Bolinas Lagoon" by Point Reyes Bird Observatory: http://www.prbo.org/cms/print.php?mid=370, retrieved on 20 May 2013.

employing internal conceptual representations (Habel & Acartürk, 2007; Acartürk, 2010). Sentence (1) characterizes the population trend of Bolinas Lagoon's Horned Grebes verbally by specifying it as a specific change, namely a *decrease_of_value* change, lexicalized by *decline,* with a *begin-value* (realized by a *from_PP*) and an *end-value* (realized by a *to_PP*). The graphical counterparts of these verbal items are the *global maximum* and the *right end-point* of the graph line. From the perspective of a seeing human, who produces a trend description depicted by a line graph, salient parts of the graph line are primary candidates to be referred to.[3] In other words, the referring expressions, which were used inside the change-characterizing PPs (as discussed above) are evoked by visually salient graph entities. Whereas in general, i.e. both in static as well as in dynamic scenes, *color* and *motion* play major roles as visually salient attributes for identifying entities using referring expressions (see, e.g., Koolen, Goodbeek & Krahmer, 2013; Carmi & Itti, 2006), in the static domain of line graphs *shape* is the dominant saliency attribute.

The conceptual inventory for verbalizing line-graph descriptions, as well as trend descriptions, has to fulfill requirements from language and from perception. Since graph lines can be seen as a specific type of contours, we include some concepts proved as successful in shape segmentation (see, Cohen & Singh, 2007), into the inventory of spatial concepts, namely *curvature landmarks,* such as, *positive maxima*, *negative minima*, and *inflections.* Moreover, in contrast to general shape segmentation, where no specific frames of reference are presupposed, line graphs provide mostly a distinguished frame of reference, explicitly via axes or implicitly through the medium of presentation (such as printed page or computer screen). In addition to Cohen-Singh landmarks, this frame of reference is employed for specifying further types of contour landmarks: Since graph lines are finite and not closed, two types of *endpoints* (left vs. right) can be distinguished.[4] Furthermore, in contrast to Cohen-Singh's *curvature maxima / minima*, a second type of extrema comes into the play, namely those with respect to the y-values, due to the ordering with respect to the vertical axis. These *value-extrema* can be global or local. (See Acartürk, 2010, on gaze studies in comprehension of value-extrema of line graphs.)

To illustrate the effectiveness of our inventory of shape concepts for language production, we come back to Fig. 2. The prominent, i.e., highly salient, shape landmarks (*psl*):

psl_1 – left-endpoint
psl_2 – glob-val-max & positive-curv-max
psl_3 – loc-val-min & neg-curv-min
psl_4 – right-endpoint

In producing basic path descriptions the *goal* has the prominent role with the *source* as secondary player (see Lakusta & Landau, 2012). Choosing the more prominent psl_2 for the source role and neglecting psl_3 leads to the core structure of sentence (1). The production of appropriate referring expression depends, inter alia, on fine-grained conceptual representations of the selected prominent shaped landmarks.[5]

The identification of the referent properties, such as the prominent shaped landmarks in the present study has been conceived as an important phase in generation of referring expressions (Reiter & Dale, 2000). The identification phase, however, is insufficient alone to provide a complete specification of the generation process (van Deemter, Gatt, van Gompel, & Krahmer, 2012). Since the past several decades, research on referring expressions have focused on various communicative aspects that have gone beyond the identification phase, such as overspecification (Pechmann, 1989) and forming conceptual pacts during the course of interaction in a collaborative environment (Brennan & Clark, 1996). One important aspect in daily-life communication is the sensory modality that the communication takes place. Previous research has shown that not only saliency in the domain of discourse via the linguistic context but also saliency in the visual context influences humans' choice referring expressions (Fukumura, van Gompel, & Pickering, 2010). Nevertheless, as of our knowledge, no study have been conducted to investigate how saliency in the haptic context influences the choice of referring expressions. In the present study, we report a comparative investigation of human referring expression generation in the visual context and in the haptic context, in the domain of line graphs. The following section introduces verbally assisted haptic graphs as the haptic context in which the referring expressions are produced.

Language Production for Verbally Assisted Haptic Graphs

Verbally assisted haptic graph exploration is a task-oriented collaborative activity between two partners, a (visually impaired) explorer (E) of a haptic graph and an observing assistant (A) providing verbal assistance (Fig. 3). Although in our current system and our empirical investigations so far the haptic explorer (E) does not communicate with the observing assistant (A) verbally[6], assisted haptic exploration has a dialog-like character:

(2.a) A has to synchronize language production with E's hand-movements in a turn-taking manner.

[3] See Habel & Acartürk (2007, 2009) and Acartürk (2010) on details of these semantic and conceptual analyses.

[4] In this paper we discuss line graphs with orthogonal directed / ordered axes. The axis of the independent variable is horizontal. Due to the direction / ordering of this axis, it is also suitable to distinguish start-point and end-point without using a common concept as border-point.

[5] The computational verbal assistant currently under development will use the *incremental conceptualizer* INC (see Guhe, Habel & Tschander, 2003) for producing such referential expressions.

[6] Investigating and realizing task-oriented dialogs between explorers and assistants, in particular with an assistive system, is planned for future research.

(2.b) The quality of verbal assistance depends on establishing appropriate referential and co-referential links.

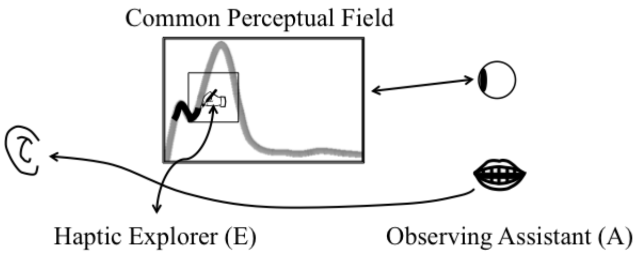

Figure 3: Assisted haptic graph exploration, a joint activity

A and E share a common field of perception, namely the haptic graph, but their perception and comprehension processes differ dramatically. For example, while E explores the highlighted, black segment of the haptic graph, A perceives the global shape of the graph, in particular, A is aware of shape landmarks and line segments. For example, when E explores the first local maximum followed by a local minimum (see Fig. 2), E does not have information about the global maximum, which is already part of A's knowledge. Therefore, the explorer and the assistant have different internal representations of the graph line, and A's referring to the graph could augment E's internal model substantially. For example, uttering (at the position depicted in Fig. 3) *Now you have reached the heights of the last peak.* would 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.* Since verbal assistance is a type of *instruction,* overspecified referring expressions are adequate for our domain (see Koolen, Gatt, Goudbeek, and Krahmer, 2011).

The success of the joint activity of explorer E and observing assistant A in general, and also the success of A's utterances in particular, depend on the alignment of the interlocutor's internal models, especially on building *implicit common ground* (Garrod & Pickering, 2004). Since 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 build up by visual perception, similarities and differences in their conceptualization play the central role in aligning on the situation-model level.

To be assistive, A should provide E verbally—in particular—with content, i.e., conceptual representations, that 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 & Haviland, 1977). Below, we present an experimental study that aimed at detecting the similarities and differences between the internal representations induced by the sensor modality used in graph comprehension, in terms of an evaluation of referring expressions produced by human participants. The results justify the conceptual analyses concerning the inventory of spatial concepts for characterizing graph lines and the content they provide. In addition, the studies give a good empirical basis for natural line graph descriptions.

Referring Expressions in Verbal Descriptions of Graphs

Towards designing verbal assistance for haptic graphs, we conducted experiments with (sighted) participants, who had no visual access to graphs while exploring haptic line graphs and participants perceiving the same graphs visually. All groups produced spoken descriptions after the presentation. Both verbal descriptions and speech-accompanying gestures were analyzed. In the present study, we only focus on the verbal descriptions regarding referring expressions (for gesture analysis, see Alacam, Habel & Acartürk, 2013). The experiment was conducted in three conditions in a between-subject design (31 participants). In the first condition (9 participants), the participants explored line graphs haptically (Fig. 1). In the recent state of the art of realizing haptic graphs, reading haptic data labels is a challenge for untrained user; therefore the graphs in the haptic condition did not contain data labels. In the second condition (11 participants), graphs with data labels (Fig. 4a) were presented on a computer screen, thus the participants had visual access to the graphs. In the third condition (11 participants), which served as a control condition, the participants inspected the visual graphs without data labels (Fig. 4b). In all conditions, after the participants explored the graph, they were asked to produce single sentence summaries of the graphs to a hypothetical audience.

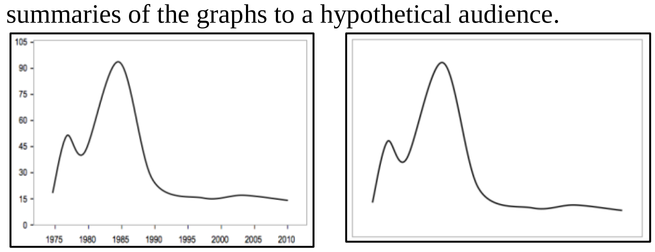

Figure 4: (a) Sample visual graph with data labels, (b) sample visual graph without data labels

Verbal Coding Graph-line segments were evaluated under two main categories; (i) curvature landmarks that emphasize a salient change in the pattern of the graph line and (ii) line segments. Since start points and end points carry semantically different information than other curvature landmarks on the graph-line, the category "landmarks" were also split into two subtypes; (i) start/end landmarks and (ii) intermediate curvature landmarks (see Fig. 5). Besides these categories, three additional categories were used. If the verbal expression conveyed information about the whole graph line, it was evaluated under the category "general", if the shape of the line was described, then it was evaluated under the "shape" category. Additionally, there are some regions that have slight changes, such as fluctuations. Those terms that refer to small changes were evaluated under "fluctuations" category.

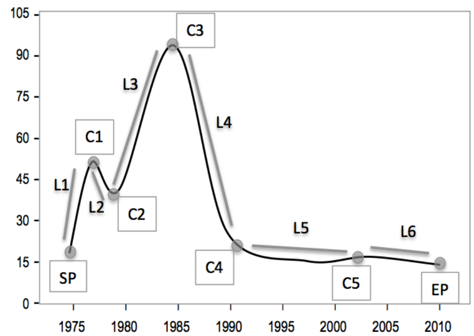

Figure 5: A sample graph segmentation (L: line segment, C: curvature landmark, SP: start point, EP: end point)

Furthermore, verbal expressions that refer to "landmarks" on the graph were also classified into three groups according to the content of the referred information. If the verbal expression contained words such as *peak*, *minimum* etc., they were evaluated under the "term" category. The category "value" classified the expressions that referred to the quantity of the bird population. Phrases such as *high value*, and *low value* were also classified in this category. The third category, "date" covered time related expressions.

Results A two-way ANOVA was conducted to investigate the differences between referring to the highlighted graph segments (curvature landmarks versus line segments) across modality. The results showed that the participants referred to line segments more than curvature landmarks for all conditions; visual graph with data labels (χ^2= 17.4, p< .05), visual graph without data labels (χ^2= 64.9, p< .05) and haptic graph condition (χ^2= 22.0, p< .05).

Moreover, Pearson chi square tests revealed that the production rate for curvature landmarks and line segments for visual graphs without data labels are significantly lower than that for visual graph with data labels (χ^2= 16.7, p< .05) and for haptic graphs (χ^2= 9.9, p< .05), without a difference between the latter two conditions, χ^2= .7, p> .05 (Fig. 6). This may indicate that the absence of data labels makes the detection of curvature landmarks harder in both modalities. On the other hand, haptic perception recovers some of the information mainly about these landmarks, which are also easily accessible via data labels.

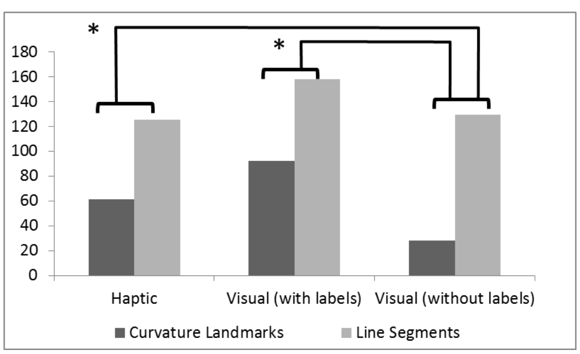

Figure 6: Production rate for curvature landmarks versus line segments

As mentioned in the coding section, the category "landmarks" has two sub-groups: start/end landmarks and intermediate curvature landmarks. A Pearson chi-square test revealed that while there was a significant difference in the mention rate for start/end landmarks and intermediate curvatures in the visual graphs with data labels (χ^2= 0.0, p< .05), the differences both in visual graphs without data labels (χ^2= 19.3, p> .05), and the haptic condition (χ^2= .5, p> .05), were not significant, see Table 1. Furthermore, more intermediate landmarks were mentioned more frequently in the descriptions of visual graphs with data labels, compared to both the haptic condition (χ^2= 5.2, p< .05) and the visual graphs without data labels (χ^2= 5.6, p< .05). However, the mention rate between haptic condition and the visual graphs without data labels were not significant (χ^2= .2, p> .05).

Table 1: Average number of referred curvature landmarks

	Haptic	Visual with data labels	Visual without data labels
Intermediate	3.60	* 6.40	1.50
Start/end	3.00	2.30	1.50

In all conditions, for the descriptions of start landmarks and end landmarks, the participants preferred using the "value" category. Moreover, intermediate curvature landmarks were mostly mentioned in the expressions under the "term" category for both haptic graphs and visual graphs without data labels. On the other hand, the participants in the visual graph condition with data labels also preferred using expressions regarding date and values represented in x and y axes, as well as the expressions in the "terms" category. Table 2 shows that the absence of data labels in the visual modality resulted in a significant decrease in the production of expressions under the "term" category (χ^2= 9.1, p< .05). However, the results indicate that haptic modality recovers that information and enhance the conceptualization of terms, such as *peak*, *minimum* etc. compared to visual graphs without data labels (χ^2= 8.4, p< .05). This analysis also indicates that expressions relating "date" and "value" are closely related to the presence data labels because they give direct access to that information.

Table 2: Average mention rate of intermediate curvature landmarks

	Haptic	Visual with Data Labels	Visual without Data Labels
Term	3.13	3.20	1.20
Value	0.25	1.30	0.10
Date	0.25	1.90	0.20

One of the major differences between the referring expressions under the three conditions was about curvature minimums, as exemplified in Fig. 7(a). The curvature minimum landmarks were mostly highlighted by the participants in the condition, in which the visual graph had data labels. The mention of these landmarks in other two conditions was seldom. Moreover, small line segments that corresponded to variation against the general pattern (Fig. 7b) were detected easily and frequently mentioned during verbal description in the haptic condition. However, these regions were usually ignored in the visual conditions. This can be due to saliency originated from friction provided by force feedback mechanism of haptic device. Small changes congruent with the general pattern can be traced very easily, but the landmarks that contain significant changes, also require physical effort for tracing since haptic representation of this kind of points can be relatively difficult.

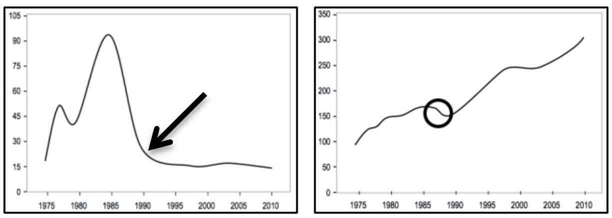

Figure 7: Example for (a) differentiating landmarks and for (b) line segments

In addition to analysis on the mention rate of curvature landmarks and line segments, the effect of modality in the production of modifiers (adjectives and adverbs) that mostly accompanied to these expressions was also investigated. The results of a Mann Whitney U test showed that the participants in the haptic condition tended to use more modifiers in their descriptions ($M= 3.65$, $SD= 1.33$) than the visual graphs with data labels ($M= 2.62$, $SD= 1.22$), $U= 584$, $p< .05$ and the visual graphs without data labels ($M= 2.33$, $SD= 1.59$), $U= 543$, $p< .05$. Moreover, the difference between the visual graphs with and without data labels was not significant ($U= 1113$, $p> .05$).

A Kruskal-Wallis test indicated a significant main effect of modality on the production of adjectives, $\chi^2= 13.8$, $p< .05$. However, modality did not have an effect on the production of adverbs, $\chi^2= 4.39$, $p< .05$. More detailed analysis was conducted on participant`s preference of modifiers (adjectives or adverbs). A Wilcoxon Signed Rank test indicated that while there was no significant difference between the number of adjectives ($M= 1.5$) and adverbs ($M= 1.2$) produced during verbal description of visual graphs with data labels ($Z= -1.05$, $p>.05$), participants produced more adjectives ($M= 1.5$) than adverbs ($M= .9$) in visual condition without data labels ($Z= -2.53$, $p< .05$). Similar to visual graphs without data labels, for haptic graphs, the production of adjectives ($M= 2.3$) was significantly higher than adverbs ($M= 1.3$), $Z= -2.62$, $p< .05$. This indicates that the graphs without data labels resulted in production of more adjectives, mainly shape based adjectives. Especially, the expressions observed in the haptic condition (e.g., *like staircase shape*, *exponential increase*, *wave shaped*, *valley phase*, *damped cosine*, etc.) were highly distinguishing from the adjectives produced in visual conditions in which the trends are mostly described according to direction of line segment (such as increasing/decreasing trend).

Discussion

Comparative analyses of the verbal descriptions produced by haptic and visual explorers exhibited both similarities and differences between haptic graphs and visual graphs. Certain aspects of graph segments turned out to be more difficult to acquire in the haptic modality than the visual modality, largely due to the sequential and local perception with a narrow bandwidth of information in the haptic modality. Due to those differences between the modalities the referring expressions that were produced by the participants exhibited diversity for the same type of graph entities under different sensory modalities, as reflected by the comparative analysis between the three groups of participants in the experiment.

A major finding in the present study was that the participants preferred to highlight line segments more frequently than curvature landmarks. This indicates that in communication through line graphs, line segments are more prominent source of information compared to curvature landmarks. Accordingly, the shape of the graph line is the property that identifies the referents by distinguishing it from its distractors. Moreover, the tendency to refer to line segments (rather than curvature landmarks) was more salient in visual graphs without data labels than haptic graphs.

Another finding that was obtained in the reported experiment was that the presence of data labels (in the visual modality) resulted in an increase in reference to the intermediate landmarks, whereas reference to the start/end landmarks was influenced neither by the modality nor the presence of data labels. A further analysis of the type of the referent (i.e., “value”, “date” and “term”) revealed that referring expressions that referred to quantity and time (i.e., “value” and “date” categories) were more frequently used in the presence of data labels. As expected, quantitative relations were difficult to acquire in the haptic condition without verbal assistance. On the other hand, the production of referring expressions under the “term” revealed a different picture: The haptic graphs and the labeled, visual graphs resulted in a similar number of “term” referring expressions, both being more frequent than the visual graphs without data labels. The similarity in the production of “term” referring expressions between non-labeled haptic graphs and labeled, visual graphs may indicate that the participants who inspected the haptic graphs were able to conceptualize the represented information in a similar way that the other participants did in the visual graphs with data labels (rather than the participants did in the visual graphs without data labels). Further study is necessary to investigate the validity of this proposal.

Finally, the participants, who verbalized the haptic graphs produced more modifiers (mainly, adjectives) compared to the visual graphs. Through the haptic modality, the participants produced shape adjectives that possibly facilitated perceiver’s memory of the represented information.

Conclusions

The study of production of referring expressions in different sensory modalities is a necessary step for the development of verbally assisted haptic exploration systems. Verbally assisted haptic exploration has a dialog-like character, even if—as in our empirical studies so far—the haptic explorers do not communicate verbally with the assistant. First of all, the synchronization between the explorer (E) and the assistant (A) should be provided in a turn taking manner, with respect to the hand movements of the explorer. Additionally, for efficient interaction, the quality of verbal assistance highly depends on establishing appropriate referential and co-referential links. In the course of the development of fully automatic verbal assistance for haptic graphs, it is necessary to investigate how the graph readers conceptualize the events represented by the graph segments (curvature landmarks or line segments), under both sensory modalities (in this case, visual and haptic).

Based on the analyses of verbal descriptions produced in the graph description tasks, we have found systematic relationships in production of referring expressions under different experimental conditions. The findings indicate that both the sensory modality and the presence of data labels influence the production of referring expressions by humans. These findings are relevant for designing an NLG system that produces adequate referring expressions for verbal assistance and for aligning the interlocutors' internal models (Garrod & Pickering, 2004) during verbally assisted graph exploration. The future work will address investigating and realizing task-oriented dialogs between explorers and assistants, in particular together with the verbal assistance system.

Acknowledgments. We thank Lena Andreeßen, Gregor Gawellek, Matthias Kerzel, and Neele Stoeckler for their assistance and are grateful to Carola Eschenbach for her comments. 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).

References

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

Alacam, Ö., Habel, C., & Acartürk, C. (2013). Investigation of haptic line-graph comprehension through co-production of gesture and language. In *Proceedings of Tilburg Gesture Research Meeting*, Tilburg, Netherlands.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology, 22*(6), 1482–1493.

Carmi, R., & Itti, L. (2006). Visual causes versus correlates of attentional selection in dynamic scenes. *Vision Research, 46*. 4333–4345.

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. In R. O. Freedle (Ed.), *Discourse production and comprehension* (pp. 1–40). Hillsdale, NJ: Erlbaum.

Cohen, E., & Singh, M. (2007). Geometric determinants of shape segmentation: Tests using segment identification. *Vision Research, 47*, 2825-2840.

Fukumura, K., van Gompel, R., & Pickering, M. J. (2010). The use of visual context during the production of referring expressions. *Quarterly Journal of Experimental Psychology, 63*, 1700–1715.

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences, 8*(1), 8-11.

Guhe, M., Habel, C., & Tschander, L. (2003). Incremental production of preverbal messages with INC. In F. Detje, D. Dörner, & H. Schaub (Eds.), *Proceedings of the 5th International Conference on Cognitive Modeling (ICCM-5)*, (pp 123–128). Bamberg University Press.

Habel, C., & Acartürk, C. (2007). On reciprocal improvement in multimodal generation: Co-reference by text and information graphics. In I. van der Sluis, M. Theune, E. Reiter & E. Krahmer (Eds.), *Proceedings of Workshop on Multimodal Output Generation*. (pp. 69-80).

Habel, C., & Acartürk, C. (2009). Eye-tracking evidence for multimodal language-graphics comprehension: The role of integrated conceptual representations. In C. Navarretta, P. Paggio, J. Allwood, E. Alsén & Y. Katagiri (Eds.), *Proceedings NODALIDA 2009 Workshop. Multimodal Communication* (Vol. 6, pp. 9-14). Odense.

Koolen, R., Gatt, A., Goudbeek, M., & Krahmer, E. (2011). Factors causing overspecification in definite descriptions. *Journal of Pragmatics, 43*, 3231-3250.

Koolen, R., Goudbeek, M., & Krahmer, E. (2013). The effect of scene variation on the redundant use of color in definite reference. *Cognitive Science. 37*, 395–411.

Lakusta, L., & Landau, B. (2012). Language and memory for motion events: Origins of the asymmetry between source and goal paths. *Cognitive Science*, *36*, 517-544.

Lohmann, K., & Habel, C. (2012). Extended verbal assistance facilitates knowledge acquisition of virtual tactile maps. In C. Stachniss, K. Schill & D. Uttal (Eds.). *Spatial Cognition VIII*. Berlin: Springer.

Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics*, *27*, 98-110.

Reiter, E., & Dale, R. (2000). *Building natural language generation systems*. Cambridge University Press.

van Deemter, K., Gatt, A., van Gompel, R. P. G., & Krahmer, E. (2012). Toward a computational psycholinguistics of reference production. *Topics in Cognitive Science*, *4*(2), 166-183.

Yu, W., Brewster, S.A. (2003). Evaluation of multimodal graphs for blind people. *Journal of Universal Access in the Information Society, 2*, 105-124.