Referring expressions in communication through line graphs: A comparative analysis of verbal descriptions

Cengiz Acartürk
Middle East Technical University

1. Introduction

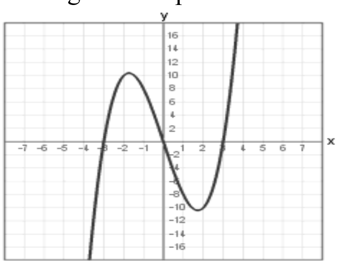
Language users identify objects that are relevant to their ongoing interaction by means of referring expressions. The use of referring expressions in communication facilitate identifying what other speakers talk about by directing attention to relevant objects in the visual world (Gundel & Hedberg, 2008). The investigation of referring expressions focuses on the interaction between language and cognition. In particular, the study of the use of referring expressions in practice has been conceived as essential for understanding how language is manifested in cognition (Hanks, 2009; Çakır & Acartürk, 2012). In linguistics, various types of referring expressions have been studied, including the study of pronoun resolution and anaphoric adverbials. Recently, referring expressions has gained increased interest in cognitive science, as well (e.g., Gray, 2012). In Turkish, the studies on referring expressions have involved pronoun disambiguation and anaphora resolution by means of natural language processing techniques, and a computational model of contextually appropriate anaphor and pronoun generation for Turkish (Kılıçaslan et al., 2009; Tin & Akman, 1994; Yüksel & Bozşahin, 2002), all focusing on linguistic context.

The present study focuses on referring expressions that are used for identifying graphical entities in the domain of statistical graphs. Statistical graphs (also called information graphics) are depictive representations for representing data. Graphs comprise a specific type of representational modality, which is different than both diagrammatic pictures (e.g., cartoons and maps) and non-diagrammatic pictures (e.g., pictorial illustrations, photographs; Bernsen, 1994). Statistical graphs are different than pictures in terms of the relationship to their referents. In particular, pictures and their referents share spatial similarity of the layout,whereas graphs are abstract representations that have an internal syntax (cf. representational formats, Kosslyn, 1980, p. 31). Moreover, in a similar way to natural language syntax, the syntactic properties of statistical graphs provide the basis for a systematic analysis at semantic and pragmatic levels (Kosslyn, 1989; Pinker, 1990). Graphs are frequently used means of communication in daily life. The communication settings thorough graphs involve written settings in newspaper articles, magazines and instructional material, as well as spoken communication settings, such as instructional presentations in lectures or in meetings. Despite their frequent use, little is known about how humans interpret graphs. The processes of conceptualization and verbalization of graphs are quite different than the graphical methods that are employed in designing and generating graphs by computer software tools. As a consequence, when designers aim at developing intelligent artificial systems that are able to verbalize graphs for specific purposes, such as verbalization generation systems for visually impaired people, the knowledge of a graph verbalization corpus is needed. Of particular importance is the set of referring expressions that are employed in verbalizing the graphs by human participants.

The ultimate goal of the present in-progress study is a multimodal corpus of graphs and their verbal descriptions. The graphs that convey trend information, in particular, line graphs are most frequently used graph type in communication settings (Zacks et al., 2002; Alaçam, Habel, & Acartürk, 2013). Therefore, in its recent form, the corpus design focuses on graphs that represent trend information, in particular, *line graphs* and *bar graphs* in time domain. In the remaining sections of the present article, a preliminary study was reported, which aimed at evaluating the role of two major factors that influence verbal descriptions: the type of the trend graph and the presence of a graphical cue.

2. Comprehension of Data Graphs

From a mathematical perspective, a line graph (or a bar graph that represents trend information in time domain) represents a mapping between two variables. If it is a function graph, each and every point on the graph line (proper) in a line graph represents a value of the function, given its domain value (Fig.1a). If it is a representation of statistical data (viz., a data graph), the lines serve for connecting data points, without a representative role. For example, in Fig. (1b), the graph represents the mean temperature in Ankara for five consecutive days in November. The data points are shown by black squares. The lines that connect the data points do not represent the data; instead, they provide perceptual continuum that aims at facilitating the interpretation of the graph by readers (Acartürk, to appear).



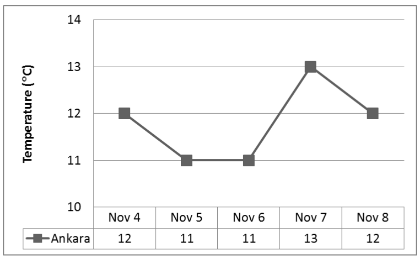


Figure 1. A line graph as a representation of a mathematical function, $y = x^3-9x$ (left).A line graph as a representation of data points, the average temperature in Ankara, °C (right).

From the perspective of human cognition, when it represents data in time domain, a line graph is a depictive representation of states and processes that unfold in time, in addition to being a representation of the mapping between time and the domain variable. Accordingly, humans verbalize specific processes, such as an increase or a decrease, when they are asked to verbalize data graphs, rather than generating a verbal description of values of the variables only.Those cognitive aspects in graph comprehension and verbalization also become important when the authors of graphs—e.g., professional graph designers, Internet blog users, etc.—use software tools to produce graphs of different types. Authors design and produce graphs based on their communicative purposes. They specify the visual techniques that connect the data points. They modify graphs by adding graphical cues, such as arrows and dot circles, which aim at foregrounding certain aspects in graphical entities (or the referents of the graphical entities in the domain of discourse). The particular aspect that is at the focus of the present study is the use of graphical cues, such as a dot circle or an arrow (Fig. 2).

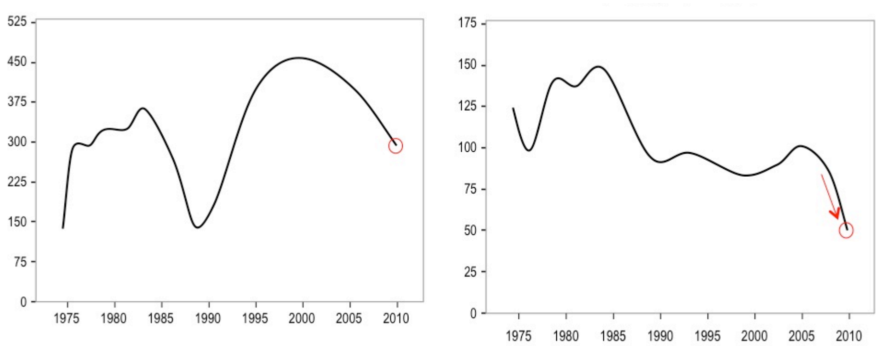
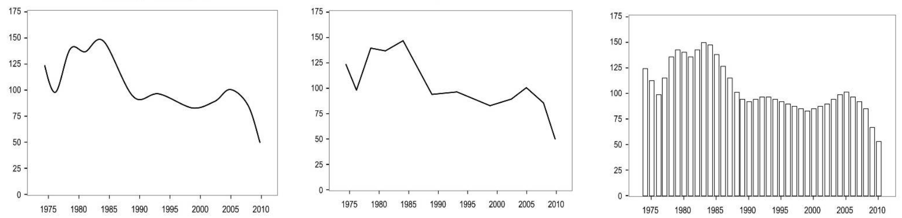


Figure 2. Two line graphs and graphical cues on the graph lines (i.e. a dot circle on the left graph, a dot circle and an arrow on the right graph).

Different graphical cues lead to different conceptualization of graphical entities (Acartürk, 2012, to appear). For instance, the graphical cues on the graph lines in Fig. 2 bring into foreground the decrease and the domain value at the endpoint of the line. The present study reports an experimental investigation of the production of referring expressions by human participants under a set of experimental conditions, as specified by different types of modifications in graphs. In Section 3, a theoretical framework is presented, which is based on a linguistic and topological analysis of graph comprehension. The experiments, reported in Section 4, present a complementary perspective to the theoretical framework.

3. The Theoretical Framework

In communication through line graphs or in their verbal descriptions, the comprehension and conceptualization of processes, events and states that are represented by the graph line leads to a referring expression vocabulary that consists of *spatial terms*. Communication through line graphs is a specific case of communication about physical space, where a set of abstract structures constructs the building blocks of conceptual representations. (Habel & Eschenbach, 1997; Eschenbach et al., 1998; Tschander et al., 2003). In particular, the vocabulary of communication through line graphs consists of shape nouns, spatial prepositions, adverbial modifiers, and verbs of change in space (Acartürk, 2010). For example, the term “peak” refers to an entity in the domain of discourse. The entity in the domain of discourse is also referred to by the peak in the graph (Habel & Acartürk, 2007). In topology, directed, linear, and bounded entities are called PATH (Eschenbach et al., 2000). The line (proper) in line graphs in time domain is a representation of the PATH concept. The peculiar aspects of the PATH concept can be referred to by graphical entities or by linguistic entities.To sum up, communication through line graphs is a specific case of communication about physical space, where the abstract structures involve process concepts such as DECREASE, based on the concept of PATH OF MOTİON, represented by the graph line (proper). The way that the graph represents the PATH OF MOTİON is specified by a set of assumptions—mostly, involuntarily—made by the author of the graph, e.g. the author may or may not use a smoothing algorithm. It is also likely that the author uses a bar graph, instead of a line graph. The three graphs in Fig. 3 exemplify the three ways that a graph represents the population data of a bird species.



(a) (b) (c)

Figure3. (a) A smooth-line graph that represents the population of a bird species in a lagoon, between 1970 and 2010. (b) A straight-line graph that represents the same data (c) A bar graph that represents the same data. Excerpted and redesigned based on PRBO, (2003).

In the experimental investigations, reported in the next section, comparative analyses of verbalizations are presented in terms of the three types of trend graphs, and in terms of the presence and absence of graphical cues, as exemplified in Fig. 2.

4. The Experimental Investigation

Six groups of participants (125, in total) produced verbal descriptions of graphs of three different type under two experimental conditions: the presence or absence of a graphical cue on the graph, as shown in Table 1.

Table 1: The distribution of experiment stimuli into six groups of participants.

	Type of the trend graph	Graphical cue available?
Group 1	Smooth line	No (e.g., Figure 4a)
Group 2	Straight line	No (e.g., Figure 4b)
Group 3	Bar	No (e.g., Figure 4c)
Group 4	Smooth line	Yes (e.g., Figure 5a)
Group 5	Straight line	Yes (e.g., Figure 5b)
Group 6	Bar	Yes (e.g., Figure 5c)

The graphs were presented in printed form and the participants produced written descriptions of the presented stimuli. The participants were asked to imagine themselves in a seminar environment where the presenter tells a single sentence for each graph. They were asked to predict and to write down the sentences. The stimuli involved eight graphs of different perceptual properties (e.g., a continuous increase, a decrease followed by a stable state, a stable state followed by an increase, etc.).

As the first step of the analysis of the collected data, the verbalizations were analyzed according to the number of alphanumerical words (including words and numerals, such as “1970”) produced by the participants. The participants produced a total of 960 sentences in verbal descriptions, which consisted of 11,751 alphanumerical words (henceforth, words). Below are sample descriptions from the data.

(1) *Siyah karınlı yağmurkuşu 1975'ten 1980'lere kadar artmış iken, 1985'ten sonra 1990'a kadar azalma gösterip sonrasında tekrar artış göstermiştir.*
‘The black bellied plover exhibited a decrease after 1985 until 1990 and then increase again, whereas it increased from 1975 to 1980s.’

(2) *Siyah karınlı yağmurkuşu nüfusu 2010 yılında 300 civarındadır.*
‘The population of black bellied plover is about 300 in the year 2010.’

The sample (1) contains 18 words, four of which are numerals, and the sample (2) includes eight word tokens, of which two are numerals. Table 2 shows the mean number of words and the mean number of numerals within the words, for each group.

Table 2: The mean number of words produced by each participant. The numbers in parenthesis show standard deviation.

	Type of the trend graph	Graphical cue available?	Mean number of words (SD)	Mean number of numerals (SD)
Group 1	Smooth line	No (e.g., Figure 4a)	12.2 (1.83)	1.29 (0.16)
Group 2	Straight line	No (e.g., Figure 4b)	13.8 (2.19)	1.52 (0.41)
Group 3	Bar	No (e.g., Figure 4c)	12.3 (1.52)	1.36 (0.22)
Group 4	Smooth line	Yes (e.g., Figure 5a)	11.4 (1.95)	1.58 (0.29)
Group 5	Straight line	Yes (e.g., Figure 5b)	11.0 (2.63)	1.49 (0.38)
Group 6	Bar	Yes (e.g., Figure 5c)	11.4 (2.42)	1.59 (0.38)

A repeated measures ANOVA was conducted to evaluate the mean number of words produced by the participants. The analysis revealed a significant main effect for the presence of a graphical cue, Wilks's $\Lambda = .45$, $F(7,105) = 18.0$, $p < .01$ partial $\eta^2 = .55$, showing that the verbal descriptions were longer (i.e., the sentences consisted of a higher number of words) in the absence of a graphical cue compared to the verbal descriptions in the presence of a graphical cue. On the other hand, no significant difference was obtained between the three types of graphs, in terms of the mean number of words produced by the participants. Moreover, the analysis of the mean number of numerals revealed significant effects neither for the presence or absence of a graphical cue, nor for the graph type. To sum up, the major finding in terms of the analysis of the number of words revealed that, when graphical cues are absent, the participants produced a longer description for the graph compared to the graphs with graphical cues.

A qualitative comparison of the verbalizations for non-cued graphs and cued graphs showed that non-cued graph verbalizations were mostly *holistic*, as exemplified in sentence (1) above, whereas cued graph verbalizations were *specific* to the aspect that was foregrounded by the graphical cue, as exemplified by the sentence (2) above. In the stimuli exemplified in Fig. 2 (left), the specific emphasis was put on the punctual state at the final year of population, which was represented by a dot circle at the endpoint of the graph line.

For a better understanding of the difference between verbalizations of cued graphs and non-cued graphs, a follow-up analysis was conducted in terms of the frequency of referring expressions in verbalizations. For this purpose, a sample stimuli (Fig. 3, left) was selected. The verbalizations of both cued and non-cued versions of the smooth-line, the straight-line and the bar-graph forms of the stimuli were analyzed. The participants in the six groups produced a total of 1358 words for the stimuli. The mean number of words per verbal description per participant was

14.0 in non-cued graph groups (i.e., Group 1, 2 and 3), whereas the mean was 8.97 in the cued-graph groups (i.e., Group 4, 5 and 6). The analysis of the word tokens was performed by grouping similar words into a set of predefined word token classes. For this, 32 word classes were identified. The predefined classes were able to cover the majority (1024 words) of the 1358 words, as shown in Table 3. The remaining 334 words were remained unclassified.

Table 3: The percentage coverage of the predefined word classes.

	Type of the trend graph	Graphical cue available?	Coverage % of the word token classes
Group 1	Smooth line	No (e.g., Figure 4a)	73.3
Group 2	Straight line	No (e.g., Figure 4b)	66.9
Group 3	Bar	No (e.g., Figure 4c)	74.0
Group 4	Smooth line	Yes (e.g., Figure 5a)	77.3
Group 5	Straight line	Yes (e.g., Figure 5b)	82.4
Group 6	Bar	Yes (e.g., Figure 5c)	83.4

For all graph types, the coverage percentages were higher in cued graphs compared to non-cued graphs. This difference shows that a narrower vocabulary was employed in verbal descriptions of the cued graphs compared to non-cued graphs.

A further analysis of the word token classes also showed that in all groups of participants, the four word token classes, shown below, were the top four classes, with minor variation in the order of occurrence.

- The name of the domain variable (i.e., the name of the bird species)
- Various numerals that represent time (e.g., 1990, 2000, etc.)
- The unit of the time variable (e.g., *yıl*, *sene* ‘year’)
- The unit of the domain variable (e.g., *sayı* ‘number’, *popülasyon,* nüfus ‘population’)

Independent of the presence or absence of a graphical cue, and independent of the graph type, the participants employed referring expressions that belong to those four word token classes with top frequency in their verbalizations (constituting about 50% of all the words in non-cued graph verbalizations and about 60% of all the words in cued graph verbalizations). The remaining word classes, however, reveal the picture more differently between non-cued graphs and cued graphs, as outlined below:

- The verbal descriptions of non-cued graphs involved more nominal forms of referring expressions for the processes represented by the graph (e.g., *decrease*, *increase*), whereas the verbal descriptions of cued graphs involved more verbal forms of referring expressions for the processes represented by the graph. This finding reveals another aspect of the holistic descriptions, which may be independent of the form of the graphical cue.

- The words *sonra* 'after then' and *kadar* 'until' play an important role in holistic descriptions of non-cued graphs, whereas in the cued graphs of the N1 stimuli, they played a minor role in descriptions.
- The participants who verbalized non-cued graphs produced more adverbial modifiers, such as *ani* 'sudden' and *önemli* 'major' to modify process verbs or nominals (e.g., *artış* 'increase'), compared to the participants who verbalized cued graphs. This finding indicates that adverbial modifiers play a more important role in the holistic verbal description of graphs compared to the verbal descriptions that emphasize a specific aspect of the graph (in this case, the emphasis on a punctual state for the analyzed stimuli).
- Cued-graph participants produced numeral labels for the domain values (e.g., 40) and approximation words (e.g., *dolaylarında, civarında* 'about, around') more frequently compared to the non-cued graph participants. The emphasis of the domain value, together with its modifiers, is the outcome of the punctual-state graphical cue, as well.

In summary, the partial analysis of the verbal descriptions revealed the four frequently-used word token classes at the top of the list: the name of the domain variable, time numerals (e.g., 2010), the unit of time (e.g., *yıl* 'year') and the unit of domain value (e.g., *sayı* 'population'). Therefore, a typical verbal description template for an automatic verbal generation system would consist of those four components. The remaining expressions in verbal descriptions revealed the differences in verbalizations between non-cued graphs and cued graphs. In particular, a verbal description template for a holistic description generation system would employ more nominal referring expressions, such as *düşüş* 'a fall', adverbial modifiers, such as *ani* 'sudden' in *ani düşüş* "a sudden fall", as well as temporal and value modifiers *sonra* 'after then' and *kadar* 'until'. The use of a graphical cue, however, would lead to a more specific description. The present article reported an investigation of the use of a graphical cue in the form of a dot circle. The use of the dot circle led to more verbal forms of referring expressions that refer to the processes represented by the graph (e.g., *azalmıştır* "decreased") compared to nominal forms. Therefore, a verbal description template for those specific descriptions would consist of verbal forms, as well as numerals for values and their modification by adverbials (e.g., *civarında* 'about, around').

5. Conclusion and Future Work

The present study aimed at contributing to the in-progress development of a multimodal corpus of referring expressions in communication through line graphs. In particular, a partial comparative analysis of non-cued and cued graph verbalizations was presented. A multimodal corpus of referring expressions would serve for automatic generation of verbal descriptions for statistical graphs. Accordingly, the present article reported a set of systematic observations in holistic verbal descriptions of non-cued graphs, as well as more specific verbal descriptions of cued graphs. Future work will complete the analysis so that it covers the other graphical cues types that were included in the experimental investigation. Future work will also address developing a naturalistic corpus of referring expressions, which will be comprised of publicly available written material, such as newspaper articles that involve statistical graphs and their verbal descriptions.

References

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

Acartürk, C. (2012). Points, lines and arrows in statistical graphs. In P. Cox, P. Rodgers & B. Plimmer (Eds.), *Diagrams 2012, Lecture Notes in Artificial Intelligence* (Vol. 7352, pp. 95-101). Berlin Heidelberg: Springer-Verlag.

Acartürk, C. (in press). Towards a systematic understanding of graphical cues in communication through statistical graphs. Journal of Visual Languages & Computing.

Acartürk, C., & Çakır, P. M. (2012). Towards building a corpus of Turkish referring expressions. In S. Demir, I. D. El-Kahlout & M. U. Dogan (Eds.), *Proceedings of the Workshop on Language Resources and Technologies for Turkic Languages. LREC 2012 Language and Resources Evaluation Conference*. (pp. 1-5). Istanbul, Turkey.

Alaçam, Ö., Habel, C., & Acartürk, C. (2013). Towards designing audio assistance for comprehending haptic graphs: A multimodal perspective. In C. Stephanidis & M. Antona (Eds.), *UAHCI/HCII 2013, Part I, LNCS Lecture Notes in Computer Science* (Vol. 8009, pp. 409-418). Heidelberg: Springer.

Bernsen, N. O. (1994). Foundations of multimodal representations: A taxonomy of representational modalities. *Interacting with Computers, 6*(4), 347-371.

Eschenbach, C., Habel, C., Kulik, L., & Leßmöllmann, A. (1998). Shape nouns and shape concepts: A geometry for ‚corner‘. In C. Freksa, C. Habel & K. F. Wender (Eds.), *Spatial cognition* (pp. 177-201). Berlin: Springer.

Eschenbach, C., Tschander, L., Habel, C., & Kulik, L. (2000). Lexical specifications of paths. In C. Freksa, W. Brauer, C. Habel & K. F. Wender (Eds.), *Spatial cognition II* (pp. 127-144). Berlin: Springer.

Gray, W. D. (2012). Introduction to Volume 4, Issue 2 of topics: Production of referring expressions: Bridging the gap between computational and empirical approaches to reference. *Topics in Cognitive Science, 4*, doi: 10.1111/j.1756-8765.2012.01190.x.

Gundel, J. K., & Hedberg, N. (2008). *Reference: Interdisciplinary perspectives*. New York, NY: Oxford 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 MOG 2007: The workshop on multimodal output generation* (pp. 69-80). University of Aberdeen, UK.

Habel, C., & Eschenbach, C. (1997). Abstract structures in spatial cognition. In C. Freksa, M. Jantzen & R. Valk (Eds.), *Foundations of computer science – potential – theory – cognition* (pp. 369-378). Berlin: Springer.

Hanks, W. (2009). Fieldwork on deixis. *Journal of Pragmatics*, 41, 10-24.
Kılıçaslan, Y., Güner, E. S., & Yıldırım, S. (2009). Learning-based pronoun resolution for Turkish with a comparative evaluation. *Computer Speech and Language, 23*(3), 311-331.
Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA.: Harvard University Press.
Kosslyn, S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology, 3*(3), 185-226.
Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73-126). Hillsdale, NJ: Erlbaum.
PRBO Conservation Science (2003). Website of PRBO Conservation Science. http://www.prbo.org/cms/366, retrieved on August 21, 2012.
Tın, E., & Akman, V. (1994). Situated processing of pronominal anaphora. In *Proceedings of Second Conference for Natural Language Processing (KONVENS'94)* (pp. 369-378).
Tschander, L., Schmidtke, H., Habel, C., Eschenbach, C., & Kulik, L. (2003). A geometric agent following route instructions. In F. Freksa, W. Brauer, C. Habel & K. F.Wender (Eds.), *Spatial Cognition III* (pp. 89-111). Heidelberg: Springer.
Yüksel, Ö., & Bozşahin, C. (2002). Contextually appropriate reference generation. *Natural Language Engineering*, 8, 69-89.
Zacks, J., Levy, E., Tversky, B., & Schiano, D. (2002). Graphs in print. In M. Anderson, B. Meyer & P. Olivier (Eds.), *Diagrammatic representation and reasoning*. London: Springer-Verlag.