What a line can say: Investigating the semiotic potential of the connecting line in data visualizations

Chapter · April 2020

DOI: 10.5117/9789463722902

1 author:

Verena Lechner
Zurich University of Applied Sciences

20. What a line can say: Investigating the semiotic potential of the connecting line in data visualizations

Verena Elisabeth Lechner

Abstract

The line is a graphical element widely used in data visualizations, its purpose often being to signal a connection between other visual elements. Based on social semiotic theory, this article investigates what semiotic functions connecting lines can have and how these functions can be related to variations in form. The results show that, in addition to the basic function of connecting elements, such lines can also indicate the level of certainty, direct the viewer to read the information either as a narrative or a conceptual claim, indicate patterns of cohesion, and regulate the viewer's position. These findings allow for further empirical research on the formation of visual conventions.

Keywords: Visual variables; Relation; Link; Metafunction; Modality; Arrow

Introduction

New digital forms of data visualization, as they appear for example on online newspaper pages and the webpages of organizations, companies, and private persons, offer the possibility to make data accessible for specialists as well as the broad public. The particular ways in which such graphical forms make meaning to the readers contribute to their social power in society, as Krippendorff states: 'We do not react to the physical properties of things, but act on what they mean to us' (1998, pp. 01_8).

A central task of various types of data visualizations (such as network visualizations, route maps, and others) is to show how different visual

Engebretsen, M. and H. Kennedy (eds.), *Data Visualization in Society*. Amsterdam: Amsterdam University Press, 2020
DOI 10.5117/9789463722902_CH20

elements are connected. Such connections are often represented by lines, which are in the focus of this chapter.[1] This chapter deals with the meaning potential of this basic element of the language of graphics, and it is thus concerned with the detail level of data visualizations. Although several examples will be given, this chapter stays on the theoretical level, whilst also opening up for practical investigations.

The use of graphical lines to represent connections between elements is an old technique (Bertin, 2011; Brinton, 1939) but still ubiquitous in current data visualizations. Nevertheless, their potential for making meaning has changed in the course of time, as the options for visual representation have increased, especially through the advent of digital production techniques and output devices. Today, the fact that connections are often represented by lines in data visualizations can be observed not only in the number of published data visualizations that have this characteristic (see Figure 20.1 as an example). It can also be observed within several tools available for digital creation of data visualizations (e.g. D3.js, Tableau, R). However, what functions these connecting lines have, and what effects different visual appearances of the lines have on their meaning potential, are issues which have not been widely researched.[2]

This chapter asks: What semiotic functions can a connecting line in a data visualization have, in addition to the basic function of indicating a connection between two visual elements?

Raising this question is necessary, not only for the scientific community, in order to generate more knowledge about visual language, but also for practitioners on the production side, in order to raise their awareness on how to communicate as nuanced and clearly as possible with their readers. Thus, a central aim for the chapter is to offer a language for discussing the functions of smaller elements within data visualizations. This is fundamental, because the meaning potential of visual elements informs important decisions in the design process.[3]

1 Besides the connective function, lines may have many other functions, e.g. to define contours, to separate elements, to lead the eye, or to function as a base element for textures and patterns (Poulin, 2012, p. 29).

2 It is possible to find literature about meaning potentials for different visual appearances of lines in general (cf. Habermann, 2015, p. 649; Horn, 1998, pp. 147-148; Ware, 2013, p. 225). However, the suggested meaning potentials are not directly transferable to connecting lines within the context of digital data visualizations.

3 It has to be noted here that other reasons than the desired function (e.g. the circumstances in the data visualization developer tools or aesthetical reasons) can influence the designer's decisions.

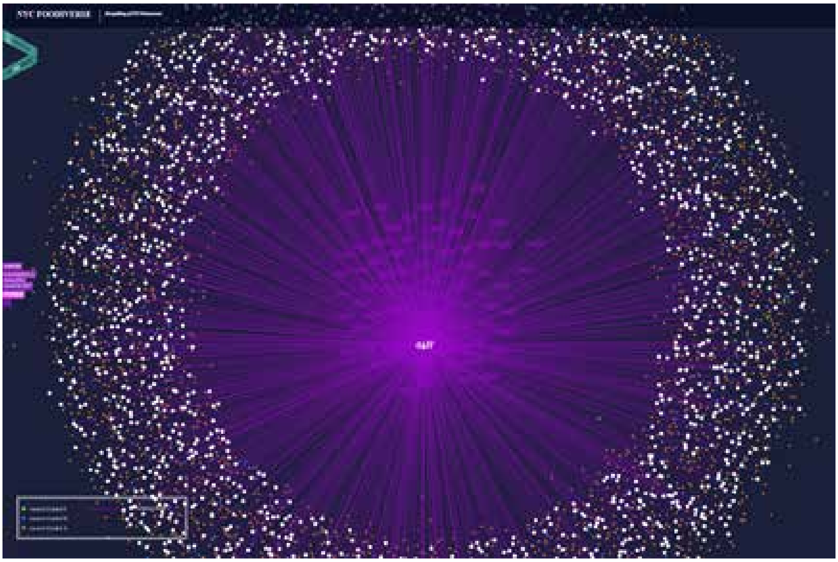

Figure 20.1. Example of a data visualization using lines to represent the connections between sanitary problems (central group of purple letter and number codes) and the restaurants in Manhattan, NYC they occurred in, represented as dots in the outer circle. From 'NYC FOODIVERSE' by W. Su, 2017 (http://nycfoodiverse.com). Copyright 2017 by W. Su. Reprinted with permission.

The approach chosen to answer the question will be presented in three steps. I will start with the central element, the connecting line itself, and describe especially *how* the line connects. In the second step, the elements that are being connected are discussed, in other words, *what* the line connects. In the last part, I describe how the connecting line is *integrated into the whole* data visualization and thus contributes to the creation of larger structures of information. However, before these steps can be taken, I need to make some terminological clarifications and outline the theoretical framework of the discussion.

Terminological considerations

The two terms that identify the object of study, namely *line* and *connection*, are used in many contexts and with a number of different meanings. Within the field of graphics, the French cartographer Jacques Bertin identified the line as one of three basic elements in the language of graphics, together with the point, and the area (2011, p. 271). Decades earlier, the Russian painter Wassily Kandinsky named the point and the line as the two elements that 'constitute the conclusive material for an independent kind of painting—graphic' (1947, p. 20). Concerning the formal characteristics of the line, he stated that the line is a product of the moving point (1947, p. 57). This movement is what provides the line with its main formal characteristics. Wucius Wong points out that the breadth of a line is 'extremely narrow' and 'its length is quite prominent' (1993, p. 45). However, as the gestalt laws state, single elements grouped in a certain way, or incomplete lines, can also be perceived as lines (Lauesen, 2005, pp. 68-69). Thus, in this chapter the term *line* refers to all kinds of visible lines, including incomplete lines and arrangements of visual elements that can be perceived as lines.

Returning to the central term *connecting line*, this refers to lines which have the basic function of establishing or indicating a connection. This means that it must be possible to identify the two parts of a connection: the single *components* and the *relationship*, as Bertin describes it (2011, p. 271), or the *nodes* and the *connector*, to mention the terms used by Engelhardt (2002, p. 40). The terminology used to describe the phenomenon of connections varies, since many synonyms exist, such as *relation*, *relationship*, *link,* and *tie* (Fergusson, 1992, p. 88). The terms are used slightly differently in various corners of the field (Brinton, 1939, pp. 43-72; Engelhardt, 2002, p. 40; Richards, 1984, p. 3/21; Ware, 2013, pp. 221-226). However, Kress and van Leeuwen, who look at graphics from a linguistic perspective similar to this chapter, use the compound term *connecting line* (2006, p. 59), which I have chosen to adopt.

Theoretical framework

Data visualizations, like other types of semiotic material, offer a specific way to communicate meaning. In order to make meaning out of a data visualization, the reader has to apply certain rules, which help him or her to decode what the producer of the data visualization wanted to communicate. On the other side, the producer of the data visualization most likely also had similar rules in mind when deciding on this specific form of visual representation. The meaning potential carried by the visual forms through the application of such shared rules defines the social function of the forms.

In the understanding of how these rules evolve, traditional semiotics and social semiotics differ on certain central aspects. In the former case, rules are seen as predefined and more or less consistent, and the communicating persons have to learn these rules before they are able to apply them, either in production or in interpretation (Hodge & Kress, 1988, p. 12). In contrast to that, in social semiotics, as van Leeuwen (2005, pp. 47-48) describes, people actively participating in social activities are seen as the ones who generate these rules—on the basis of certain culturally shared codes. He further explains that semiosis is an ongoing process, where the sign users themselves have the power to influence and change the rules. This again implies that those rules are seen to be rather unstable and to a high degree dependent on the social situation.

Returning to the case of data visualizations, which are often produced for a large and diverse target group, we can assume that some rules exist, connecting forms to meanings. But they might be somewhat different from how they were years ago and might also be dependent on the social context. M. A. K. Halliday laid the theoretical basis for seeing text as 'a sociological event, a semiotic encounter through which the meanings that constitute the social system are exchanged' (1978, p. 139, emphasis deleted). However, in the centre of his research stands verbal language. In their seminal work *Reading Images*: *The Grammar of Visual Design* (2006, p. 2), Theo van Leeuwen and Gunther Kress state that visual structures, just like linguistic structures, invite a particular interpretation, that is formed by experience and social interaction.[4]

From these theoretical abstractions, we can conclude that the process of meaning-making in contexts involving data visualization is a process where certain culturally formed, relatively stable codes and conventions interplay with a set of more unstable, situated rules concerning the exact meaning of the visual forms displayed. This interplay also defines the meaning of connecting lines, and calls for empirical research to investigate which semiotic functions are conventionalized and which are not.

Halliday defined three universal functions in verbal language, also known as 'metafunction[s]' (2004, p. 30), understood as different aspects of the meaning potentials of a clause. Any clause, any verbal utterance, carries all three metafunctions simultaneously: the 'ideational' (what is said about the world), the 'interpersonal' (how social relations between the participants are constructed), and the 'textual metafunction' (how the parts construct a coherent whole) (pp. 30-31). As Kress and van Leeuwen adapted Halliday's concept of social semiotics to other semiotic modes, they also applied the concept of these metafunctions to the analysis of visual expressions (2006, p. 13), and during the last decades, their work has been adapted by many other researchers. Yet, for every new social

4 It should be underlined that they also reflected critically on this comparison between visual and verbal structures and pointed out that this similarity easily can be overemphasized (see also Kress & van Leeuwen, 2006, p. 76).

semiotic study, the systems of meaning making have to be defined again in order to make a systematic analysis possible. This is the case because different types of visual material offer different semiotic choices and need to be interpreted in different ways. Thus, it is also necessary to define the system of choices activated in the kind of visual material investigated in this book. To develop the basis for that, focusing on connecting lines, is the contribution of this chapter.

Towards an analytical procedure

Functions related to the connecting line itself

Having given a brief insight into terminology and the theoretical framework for the study, I will now present a proposed method for analysing types and functions of connecting lines. Starting with the central element, the connecting line itself, its main function first has to be pointed out. This function is already implied in the word *connecting,* and therefore works as a basic selection criterion for the kind of lines that are to be investigated. As Clive James Richards noted, a line can have a verb-like function, and in the verbal translation of a figure showing two letters with a line in between, he states: 'A is connected to B' (1984, p. 3/21). In the words of Halliday, this corresponds to the ideational meaning, which says something about a process, or 'goings-on' (2004, p. 170). More precisely, the line represents the process itself. Secondly, the connecting line might also say something about the associated circumstances, e.g. whether the connection is strong or weak. The third component of a process—the participants—is determined at the ends of the line, showing what is connected.

Following the proposed analytical structure, we can summarize that a line *connects* certain objects in a *certain manner.* It represents a connection in a specific way, and it can, among other things, also point to the certainty of this connection.[5] In verbal language we have several alternatives to express the certainty of a piece of information, building up the modality system of the language in question (Halliday, 2004, p. 147). This system offers means to express the level of certainty that the speaker wants to give a certain claim—e.g. choosing between *This is probably true* and *This is certainly true.*

5 In his visual grammar of relationship representations Colin Ware named, for example, the strength of a connection as a characteristic that could possibly be expressed by different line weights (2013, p. 225).

Hodge and Kress assumed that modality markers can also appear in other kinds of media, although they considered them to be not so clearly articulated as the ones in verbal language (1988, pp. 121-122). Kress and van Leeuwen further investigated modality in terms of visual communication, including examples like illustrations, photographs, and pieces of art (2006, pp. 159-180). Because modality markers are closely related to the social participants in the communication process and are used to build shared truths, they consider modality as a phenomenon to be categorized as part of the interpersonal rather than the ideational metafunction (2006, pp. 159-160).

Van Leeuwen notes that different coding systems can have different kinds of *coding orientations*—like the 'naturalistic', 'technological', 'sensory', and 'abstract' orientations (van Leeuwen, 2005, pp. 168-170). As he understands it, these different orientations mean that the scales of modality may have different types of markers, or criteria for what is regarded true and realistic, and what is not. A line graph may show little details of the background, compared to a photograph, but that does not indicate that what is represented in the diagram is not true (p. 167). Within the abstract coding system, 'visual truth is abstract truth' (p. 168). 'The more an image [...] represents the general pattern underlying superficially different specific instances, the higher its modality from the point of view of the abstract truth. This is expressed by reduced articulation', he explains further (p. 168). This means that if a data visualization is seen as being part of this coding system, an abstract way of visualizing data conveys an impression of truth.

Visualizations of past and future paths of hurricanes can serve as an example of data visualizations that often contain a degree of uncertainty, such as the data visualization *Irma is following a well-worn path* (Dottle, King, & Koeze, 2017).[6] Here, the future, uncertain path of Hurricane Irma is shown as a dashed line, within a shape surrounded by another dashed line. The interruption of the lines therefore serves as the modality marker within this example. Here, the lack of sufficient data is signalled visually in the data visualization. In other cases, the data available may be precise and sufficient, but for certain reasons (like privacy protection) the visualizations are intentionally made to look imprecise, through the application of uncertainty markers (Dasgupta, Chen, & Kosara, 2012, p. 1022). Yet, although it might appear clear in the example of Hurricane Irma's path, there does not exist any general and recognized description of how certainty is expressed in graphical material through different forms of connecting lines. Sometimes, like in *Musicmap* (Crauwels, 2016, see https://musicmap.

6 See https://fivethirtyeight.com/features/what-lies-in-irmas-path/.

info/), dashed lines are used for purely compositional reasons, in order to separate them from other lines. The same can be said about other visual variables shaping the physical appearance of the line (like colour, shape, etc.). How these visual variables indicate specific types of connections, whether through convention or explicit explanation, is an issue that calls for extensive empirical investigation.

Functions related *to what* the line connects

As stated earlier, to recognize a connecting line as such, it must be possible for the reader to identify not only the line itself, but also the connected components. When looking at different types of data visualizations, it becomes apparent that sometimes lines are used to connect two different elements (as in network diagrams, see Figure 20.1), whereas in other types the lines connect two different states of the same element (as in route maps). In either case, the function type in question belongs to the ideational type, saying something about states in the world.

In order to trace the graphical lines to the natural, non-digital world, a relevant source is Tim Ingold (2007, pp. 41-43), who writes about lines from an ethnological viewpoint. He divides lines into five groups, two of which are called 'threads' and 'traces' (p. 41).[7] In his account, *threads* (such as a washing line, an electrical circuit, a tightrope, or a skein of yarn) seem to correspond to the former group of lines, connecting two different elements. *Traces* (such as a scratched line or the slime trail of a snail), on the other hand, relate to the connecting lines of the second type, connecting different states of the same element.

The way that connecting lines relate the connected components to each other forms their *representational structure*, a concept investigated by Kress and van Leeuwen (2006) within many different visual media. The two main categories into which they divide their investigated material are 'narrative structures' and 'conceptual structures' (p. 79). What is represented in narrative structures are 'unfolding actions and events, processes of change, transitory spatial arrangements' (p. 79). They contain 'vectors' (p. 59) which show a direction. In data visualizations, connecting lines can work as vectors when the direction is made explicit, e.g. by an arrowhead or a tapering body, as in Figure 20.3. Conceptual structures, on the other hand, have no vectors, and represent 'participants in terms of their more generalized

7 The three other groups he calls 'cuts, cracks and creases' (2007, p. 44), 'ghostly lines' (p. 47), and 'lines that don't fit' (p. 50).

and more or less stable and timeless essence, in terms of class, or structure or meaning' (p. 79). In other words, narrative structures always contain a form of action, whereas conceptual structures describe a phenomenon in a certain state. Although originating in different disciplines, it seems obvious that Ingold's traces to a certain degree correspond to the concept of narrative structures, while threads correspond more closely to the concept of conceptual structures.[8] Based on both sources, I suggest that connecting lines have the function of directing the viewer to read the information either as a narrative or as a conceptual claim, and that the way they do this through their visual appearance in current data visualization design is an issue that calls for both theoretical and empirical investigation.

Functions related to the line as part of larger text units

After having started this investigation on the micro level focusing on the connecting line itself, and then extending it to the connected units, it is now time to have a look at the surrounding context, that is, the wholeness of the data visualization. Some data visualization types, like network visualizations or tree diagrams, traditionally contain many connecting lines, often even lines interconnected with each other. Other data visualization types, like flow maps, might either show only one or a few lines, which are not necessarily interconnected (although they may cross each other). Such examples make it obvious that connecting lines contribute to, and are integrated in a bigger whole, a composition of semiotic elements. This observation is a starting point for analyses that focus on the textual, also called the compositional metafunction. At this level, *cohesion* is a core concept.

Linguists working on the discourse level have a long tradition of describing connectedness in verbal texts (Sanders & Pander Maat, 2006, p. 591). For the English language for example, Halliday and Hasan published their pioneering book *Cohesion in English* already in 1976, of which some main ideas shall be explained in the following (1994, pp. 1-4). According to them, what makes a text be regarded as such, is that it forms a recognizable, coherent unit of meaning. For that, it needs to have meaning relations that combine the single text units. These are called 'cohesive properties' (1994, p. 4). These properties come into action when 'the interpretation of some element in the discourse is dependent on that of another' (1994, p. 4, emphasis deleted). In other words,

8 Both Ingold and Kress and van Leeuwen emphasize that it is not always possible to distinguish their research material exhaustively with their categories. They rather see them as a tool for describing what is represented (Ingold, 2007, p. 50; Kress & van Leeuwen, 2006, p. 86).

a cohesive text contains connections between the single elements of the text, which help the reader to understand the meaning of the entire text.

However, the phenomenon of cohesion is manifest not only in verbal texts. Theo van Leeuwen has investigated forms of cohesion in the field of multimodal texts (2005, pp. 179-268), the principles of which will be shortly introduced here. He lists four ways of constructing cohesion, namely: 'composition', 'rhythm', 'information linking', and 'dialogue' (p. 179). *Composition,* he explains, works with the placement of elements in space. For van Leeuwen, whether an element is placed at the bottom or the top of a page, to the left or to the right, in the centre or in the margin, has an impact on its meaning potential. This impact is often based on metaphors from the physical world. Composition, as he continues, is the spatial equivalent to *rhythm.* This in turn is formed by a transition between two opposing states repeated in the dimension of time, such as soft and loud, fast and slow, big and small, and so on. *Information linking* has to do with the ways that one piece of information can be related to another piece. A *dialogic* structure, as the fourth form of cohesion he suggests, appears when more than one voice is perceived either simultaneously or sequentially, like in a spoken dialogue, or in a film track, where the flow of images and the music track may establish a dialogue (pp. 179-268).

When investigating the cohesive functions of connecting lines in digital data visualization, I propose to focus on *composition* and *information linking*, for a number of reasons.

Composition is built up by the spatial arrangement of the constituting elements in the visual object. In the case of data visualizations containing several connecting lines, the conscious placement of such connections, which include the connecting line as well as the connected elements, can be used to imply a certain meaning potential. On a macro level, the composition of these elements can build up specific types of data visualizations and help to define the roles of the connected elements. The configuration of connecting lines can indicate, for example, sequences, hierarchies, or networks, offering very different roles for each of the involved visual elements (see Figure 20.2).

In some types of interactive data visualization, the user is enabled to change the placement of the nodes and its connections manually. The visualizations in the report *Panama papers—the power players* (ICIJ, 2017), for instance, offer this affordance. (Figure 20.3 is a static screenshot of one of them.)[9]

Information linking, as another way to construct cohesion (van Leeuwen, 2005, pp. 219-247) shall here be discussed in further detail. Borrowing the

9 See https://panamapapers.icij.org/the_power_players/

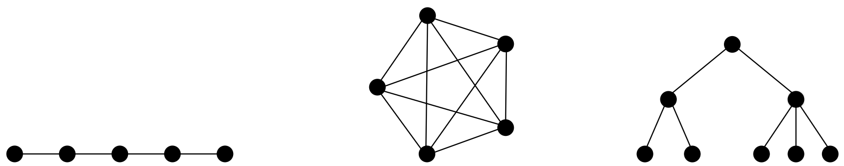

Figure 20.2. Three exemplary compositions of connecting lines and their connected elements. Illustration by V. E. Lechner.

linguistic concept of conjunctions from cohesion within verbal texts (as described in Halliday & Hasan, 1994, pp. 336-338) van Leeuwen states that links are 'temporal, logical or additive' (2005, p. 222). He further explains that if a temporal link occurs, this points to the fact that the two single pieces of information happen either at different points in time or in parallel. A logical link highlights that one of the information pieces 'gives a reason for, a condition of, or a comparison with the information in the first item' (p. 223). If it is not a temporal or logical link, yet the one item adds information to that given by the other, he concludes that additive linking takes place.[10]

Whereas in verbal language, the different linking types can be determined because of explicitly used words like conjunctions, in data visualizations, such 'cohesive tie[s]' (1994, p. 329), as Halliday and Hasan call them, might not always be so obvious. But as van Leeuwen shows with examples of multimodal, non-linear texts (2005, pp. 226-247), they do exist and have to be found by the reader to form the storyline. For this process, the surrounding context plays an important role, as it might influence which linking type might be the most relevant in a specific data visualization.

In Figure 20.3, we see a visualization where the connecting lines indicate a combination of additive and logical linking. The connections around Sigmundur Davíð Gunnlaugsson (former prime minister of Iceland), could be verbally translated to: *Gunnlaugsson is registered in address X and is a shareholder of company Y. Similarly, his wife is also registered at the same address* and *is also a shareholder of the same company* (*which was registered by a consulting firm*). The cohesion markers *and* and *similarly* point to additive respectively logical linking. The example shows that translating the data visualization into text may help to detect the ways in which lines are used as cohesion markers.

As the previous examples show, *composition* and *information linking* are relevant when studying cohesion formed by graphical lines in data

10 Examples for cohesion markers within verbal text: *then, next, meanwhile* (temporal linking); *because, for that reason, otherwise* (logical linking); *and, or* (additive linking). These examples can be found in van Leeuwen (2005, pp. 222-224) as well as in Halliday & Hasan (1994, pp. 336-338).

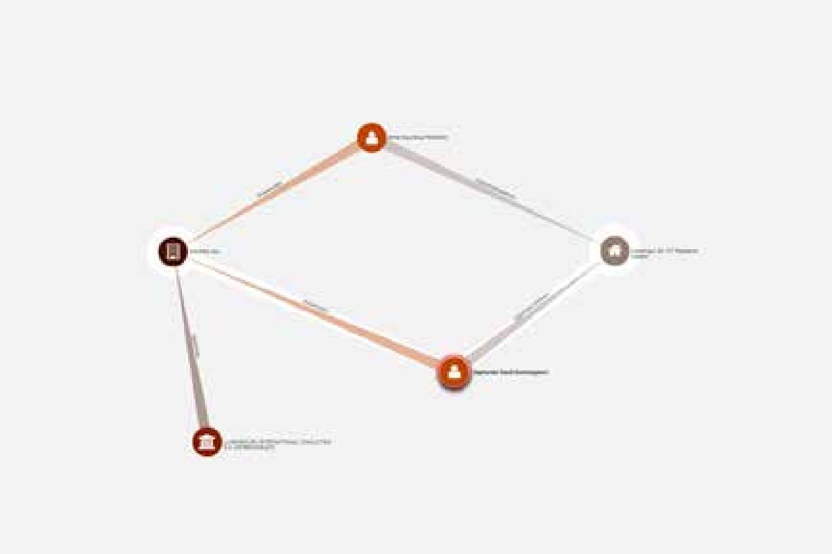

Figure 20.3. Visualization of the connections related to Sigmundur Davíð Gunnlaugsson. From 'Panama Papers—The Power Players' by The International Consortium of Investigative Journalists, 2017 (https://www.icij.org/investigations/panama-papers/the-power-players/). Copyright 2017 by ICIJ. Reprinted with permission.

visualizations. However, *rhythm* and *dialogue* can also form cohesion in such forms of textual expressions.

Rhythm can be established through visual representations of processes occurring over time, analogue to the rhythmic structures perceived in music (van Leeuwen, 2005, p. 182). In data visualizations, this form of cohesive structure can be perceived through observations of visual repetitions and certain patterns of such. It can either be shown by an animation or by presenting these repetitions in a linear sequence in a static presentation. Connecting lines can play a role in such a rhythmic organization of a data visualization. A visualization in the news site of *The Washington Post* shows flight patterns after the Brussels attacks on March 22, 2016 (Muyskens, 2016)[11]. In the early morning, all planes are flying directly to Brussels airport and build up a regular pattern of moving lines. The lines connect the planes to geographical points. But suddenly their flight routes change and the lines representing them develop an irregularity because the planes turn back before entering the airport. This example points to the fact that it is not always the coherent pattern itself that forms the most interesting feature

11 https://www.washingtonpost.com/news/wonk/wp/2016/03/23/watch-what-happened-to-flight-patterns-in-the-moments-after-brussels-attacks/

of a (visual) text, but rather the instances of violation of the pattern. The rhythm is disrupted, and attention is attracted.

In many publicly available data visualizations, a *dialogue* between different semiotic modes appears, between the visual forms, verbal elements, numbers, and sometimes dynamic modes like music or speech. How this dialogue is organized in time and space is interesting to analyse, but not so relevant for the investigation of connecting lines.

Exploring the functions of the line in relation to the total data visualization, we may also look at two aspects of interpersonal meaning, namely the ways in which the reader's position is regulated through *frame size* and *perspective*. According to Kress and van Leeuwen (2006, pp. 124-129) social distance between a human represented on an image (e.g. a photo) and the viewer of the image is managed by different frame sizes, 'close distance', 'middle distance', and 'long distance' (pp. 125-126). They further suggest that a similar set of relations is possible between the viewer and depicted non-human elements. In data visualizations, the connecting lines can be presented from a very far distant position (showing much of the surrounding context) or from very near, just as if the viewer could touch them. In interactive data visualizations, where the user is able to zoom, this could even be changed manually (as in *Musicmap* (Crauwels, 2016)).[12] Such interactive mechanisms offer the reader a position as an active participant in the communication, being able to choose the frame size and therefore also the position from where the data visualization is observed.

Besides frame size, the chosen perspective also influences the relation between the viewer and the represented objects. Data visualizations are often presented in a direct frontal or a top-down angle, which adds to their aura of objectivity, whereas other angles rather indicate subjectivity (Kress & van Leeuwen, 2006, pp. 135-151). Placing the connecting line in a two- or three-dimensional space makes it possible for the connecting line to indicate perspective. In data visualizations, a top-down angle is e.g. often used for route maps, where the movement of certain objects is shown on a map. One such example is the first visualization in the news article *Bussed out: How America moves its homeless* (Outside in America team, Bremer, & Wu, 2017), showing the route of a homeless person relocating in the US.[13] In another visualization from the same article, the same kind of geographical movements are presented as curved lines viewed from a frontal perspective

12 see http://www.musicmap.info

13 see http://www.theguardian.com/us-news/ng-interactive/2017/dec/20/bussed-out-america-moves-homeless-people-country-study

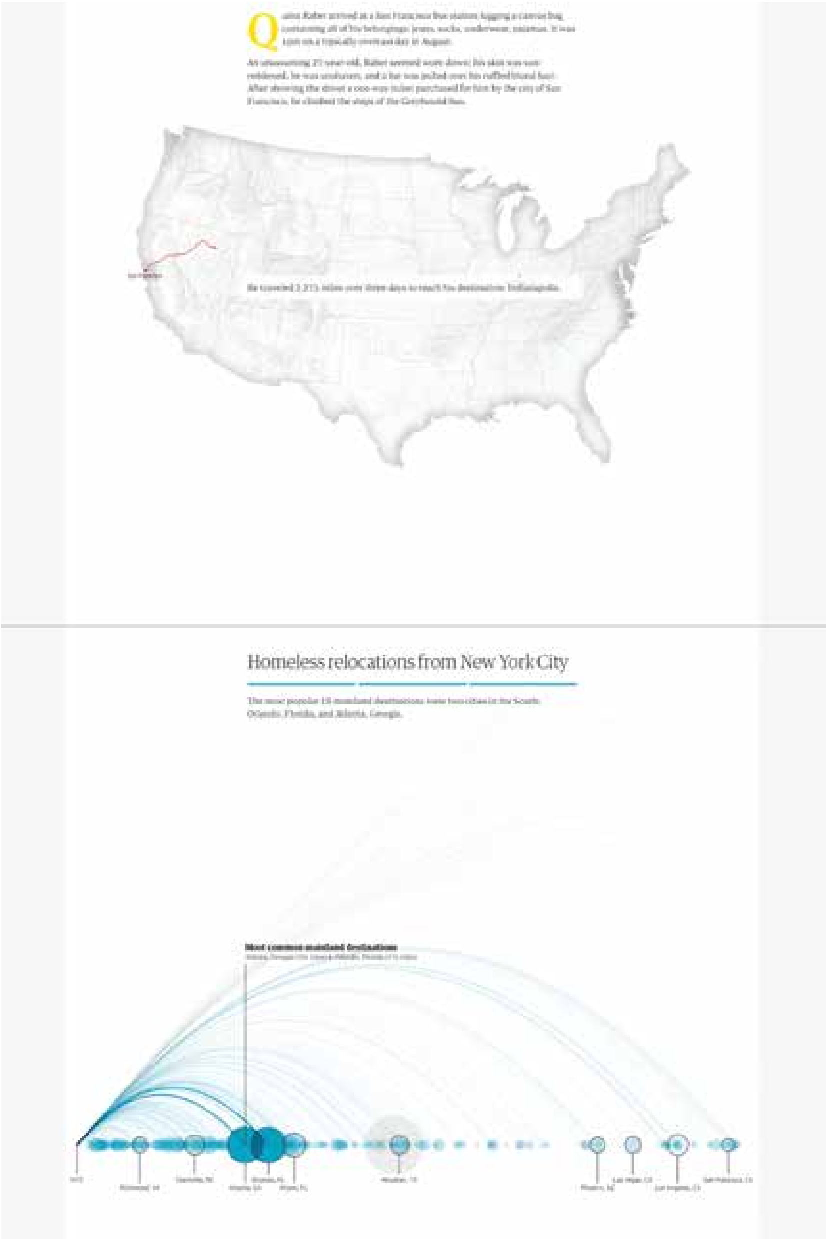

Figure 20.4. Two visualizations of spatial movement, using a top-down angle in a route map (upper picture) and a frontal perspective in an arc diagram (lower picture). From 'Bussed out: How America moves its homeless' by Outside in America team, N. Bremer and S. Wu, 2017 (http://www.theguardian.com/us-news/ng-interactive/2017/dec/20/bussed-out-america-moves-homeless-people-country-study). Copyright 2017 by *The Guardian*. Reprinted with permission.

(see Figure 20.4). If we apply Kress and van Leeuwen's principle, both perspectives indicate an objective representation of reality. Similar to frame size, some interactive data visualizations, such as Kim Albrecht's *Cosmic Web* (n.d.) also include possibilities to manually change the perspective.[14]

Conclusion

Beyond the basic function of connecting, four semiotic functions of connecting lines in the context of data visualizations have been identified and described in this chapter.[15] Connecting lines can potentially be used:

1) To indicate the level of certainty of a specific connection as a modality marker.
   Both the visual scales of the modality markers as well as what the exact values indicate in a certain context need to be investigated further on a corpus of data visualizations before they can be used for analysing single data visualization examples.
2) To direct the viewer to read the information either as a narrative or as a conceptual claim—how things *develop* or how things *are*.
   It should be possible to identify the two types of representational structures with the help of the surrounding context of the data visualization. What kinds of sub-categorization are possible and reasonable can only be discovered with the help of a corpus of data visualizations.
3) To indicate patterns of cohesion in a data visualization, and to indicate the role of particular objects in the context of the whole.
   Composition and information linking are especially relevant when investigating a data visualization as a cohesive textual unit.
4) To regulate the reader's position, by regulating the physical relation between the viewer and the connecting line(s).
   The concepts of frame size and perspective, such as proposed by Kress and van Leeuwen (2006), are directly applicable. However, the effects on the reception side call for further research.

In order to see if and how these potential functions are realized in current data visualization design, empirical research on larger corpora of data visualizations is demanded. Such studies would also offer insights in the

14 See http://cosmicweb.kimalbrecht.com/viz/#1

15 This list is not meant to be exhaustive, but it may serve as a starting point for further investigations.

evolving process of conventionalization—the forming of rules for making meaning through data visualizations, widely shared on both the production side and the viewer side. The stronger the conventions become, the stronger data visualizations' role will be in society, because they will afford a more nuanced communication. The relevance of these fields of knowledge can be demonstrated with the social impact of the following two examples. If a data visualization showing the potential future path of a hurricane (which might be indicated by connecting lines as in *Irma is following a well-worn path* (Dottle, King, & Koeze, 2017)) is interpreted incorrectly, this might have an impact on whether or not people decide to leave their homes. Similarly, if a data visualization about problems of a disadvantaged group in society (such as *Bussed out: How America moves its homeless* (Outside in America team et al., 2017), where connecting lines indicate the journey of homeless people taking part in a relocation programme) provokes a long social distance instead of compassion, social awareness might not be developed, and action might not be taken.

References

Albrecht, K. (n.d.). *Cosmic web* [Project page]. Retrieved March 31, 2018 from http://cosmicweb.kimalbrecht.com/viz/#1

Bertin, J. (2011). *Semiology of graphics: Diagrams, networks, maps* (W. J. Berg, Trans.). Redlands, CA: ESRI Press.

Brinton, W. C. (1939). *Graphic presentation*. New York: Brinton Associates.

Crauwels, K. (2016). *Musicmap: The genealogy and history of popular music genres*. Version 1.0.1. Retrieved April 21, 2018 from http://www.musicmap.info

Dasgupta, A., Chen, M., & Kosara, R. (2012). Conceptualizing visual uncertainty in parallel coordinates. *Computer Graphics Forum, 31*(3), 1015-1024). https://doi.org/10.1111/j.1467-8659.2012.03094.x

Dottle, R., King, R., & Koeze, E. (2017, September 8). What lies in Irma's path [Blog post]. Retrieved from https://fivethirtyeight.com/features/what-lies-in-irmas-path/

Engelhardt, Y. (2002). *The language of graphics. A framework for the analysis of syntax and meaning in maps, charts and diagrams* (Doctoral thesis, University of Amsterdam). Retrieved from http://dare.uva.nl/search?arno.record.id=105970

Fergusson, R. (Ed.). (1992). *The Penguin dictionary of English synonyms and antonyms* (revised). London: Penguin Books.

Habermann, H. (2015). *Grundlagen der Gestaltung: Industrie-Design / Kommunikations-Design*. Göttingen, Germany: Cuvillier Verlag.

Halliday, M. A. K. (2004). *An introduction to functional grammar* (3rd ed). London: Arnold.

Halliday, M. A. K. (1978). *Language as social semiotic: the social interpretation of language and meaning*. London: Arnold.

Halliday, M. A. K., & Hasan, R. (1994). *Cohesion in English* (13th impression). New York: Longman Group Limited.

Hodge, R. I. V., & Kress, G. (1988). *Social semiotics*. New York: Cornell University Press.

Horn, R. E. (1998). *Visual language: Global communication for the 21st century*. Bainbridge Island, WA: MacroVU.

ICIJ. (2017). Panama Papers—The Power Players [Investigation project page]. Retrieved July 30, 2018 from https://panamapapers.icij.org/the_power_players/

Ingold, T. (2007). *Lines: A Brief History*. Abingdon: Routledge.

Kandinsky, W. (1947 [1926]). *Point and line to plane: Contribution to the analysis of the pictorial elements*. (H. Dearstyne & H. Rebay, Trans.). New York: Solomon R. Guggenheim Foundation. Retrieved from http://archive.org/details/pointlinetoplaneookand

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). London: Routledge.

Krippendorff, K. (1998, December). *Design discourse: a way to redesign design*. Keynote speech presented at the Society for Science of Design Studies, Tokyo, Japan. Retrieved from https://repository.upenn.edu/asc_papers/227

Lauesen, S. (2005). *User interface design: A software engineering perspective*. Harlow: Pearson Education.

Muyskens, J. (2016, March 23). Watch what happened to flight patterns in the moments after Brussels attacks. *The Washington Post.* Retrieved from https://www.washingtonpost.com/news/wonk/wp/2016/03/23/watch-what-happened-to-flight-patterns-in-the-moments-after-brussels-attacks/

Outside in America team, Bremer, N., & Wu, S. (2017, December 20). Bussed out: how America moves its homeless. *The Guardian.* Retrieved from http://www.theguardian.com/us-news/ng-interactive/2017/dec/20/bussed-out-america-moves-homeless-people-country-study

Poulin, R. (2012). *The language of graphic design: An illustrated handbook for understanding fundamental design principles*. Beverly, MA: Rockport.

Richards, C. J. (1984). *Diagrammatics: An investigation aimed at providing a theoretical framework for studying diagrams and for establishing a taxonomy of their fundamental modes of graphic organization* (Doctoral thesis). Royal College of Art, London, United Kingdom.

Sanders, T., & Pander Maat, H. (2006). Cohesion and coherence: Linguistic approaches. In: *Encyclopedia of Language & Linguistics* (2nd ed.). (pp. 591-595). https://doi.org/10.1016/B0-08-044854-2/00497-1

Su, W. (2017a). NYC FOODIVERSE. Retrieved July 31, 2018 from https://jiahao01121.github.io/Restaurant_data_viz/second_draft_mar_25/index.html

Su, W. (2017b). NYC FOODIVERSE—Storytelling of NYC Restaurant. Retrieved April 25, 2018 from http://nycfoodiverse.com/

van Leeuwen, T. (2005). *Introducing social semiotics*. London & New York: Routledge.

Ware, C. (2013). *Information visualization: Perception for design* (3rd ed.). Waltham, MA: Morgan Kaufmann.

Wong, W. (1993). *Principles of form and design*. New York: John Wiley & Sons.

About the author

Verena Elisabeth Lechner is currently a PhD student at the University of Agder, Norway. She holds an MA in InterMedia from the University of Applied Sciences Vorarlberg, Austria and is especially interested in the visual appearances and meaning potential of graphic elements within data visualizations.