




# Modality and Uncertainty in Data Visualizations: A Corpus Approach to the Use of Connecting Lines

Verena Elisabeth Lechner<sup>(✉)</sup> 

University of Agder, 4630 Kristiansand, Norway  
verena.lechner@uia.no

**Abstract.** In data visualizations, connecting lines may have various semiotic functions, including the semiotic potential of indicating modality and uncertainty. The goal of this article is to find out how this semiotic potential is realized in current best practices of data visualizations and what conventions exist for the visual manifestations of these functions. This issue is addressed by using a corpus-based approach and a two-level analysis method within a social semiotic framework. First, the article offers a theoretical discussion on how the concepts of modality and uncertainty interrelate. Second, a method for investigating how these concepts are visualized at different levels is presented. Third, a corpus analysis including 163 award-winning data visualizations is presented. The results indicate the existence of certain conventions for visual modality markers, and thus offer new insights relevant for both design theory and practice.

**Keywords:** Uncertainty · Modality · Social semiotics · Visual variables · Line · Connection · Reliability · Probability

## 1 Introduction

Data visualization (further abbreviated as *DV*) is a rapidly developing visual means of communication, strongly influenced by the advent of new digital technologies [1]. The amount of accessible data is greater than ever [2] and the forms of representation are constantly being developed further [3]. Consequently, the desire to express a variety of meanings through *DV* is increasing, and so are the graphical opportunities to do it. Such ‘new modes of production bring with them new affordances’ [4], which means that the conventions that connect visual expressions to culturally shared meanings are constantly under development. A specific graphical element, namely the line that is used to connect two entities (further referred to as *connecting line*), appears in many different

---

<sup>1</sup> The Peutinger Map, known as the first route map (366-225 BC) [3] and a multiple time-series graph showing planetary movements over time (10th century) [5] are historic visualizations that use connecting lines as their central graphical element.

---

This work was supported by the Research Council of Norway under Grant number 259536.

© The Author(s) 2020

A.-V. Pietarinen et al. (Eds.): Diagrams 2020, LNAI 12169, pp. 110–127, 2020.

[https://doi.org/10.1007/978-3-030-54249-8\\_9](https://doi.org/10.1007/978-3-030-54249-8_9)

visualization types (such as line charts, network diagrams, route maps etc.) and not just since the digital age.<sup>1</sup> However, the possibilities to signify specific meanings with such connecting lines have increased, because the application of transparency, interaction effects, animations etc., has become much easier in the digital era.

In the practice field of DV, *uncertainty* is a much-disputed topic<sup>2</sup>. One reason for that is that many big and complex datasets available today include elements of uncertainty related to confidence, variability, trends etc. [2]. The consequence for DV designers (here used as a collective term for all persons included in the DV production process) is that they have to find ways to visualize this uncertainty. Especially when designing visualizations for lay audiences, depicting uncertainty still remains a challenge, so big that it sometimes is not visualized at all [10].

*Modality*, as investigated by linguists and semioticians in verbal and visual text, is a concept that to some extent overlaps with uncertainty, as it is discussed in the practice field of DV. However, modality has not been brought into that field so far.

In the following, I will present the different perspectives on modality, both from a functional grammar point of view [following 11] and from a multimodality point of view [following 4 and 12] and relate the concept of modality to the concept of uncertainty. This is done because the linguistic concept of modality is very elaborated, and it is a hypothesis underlying this study, that it is relevant and useful also in the investigation of DVs. Following this theoretical trajectory, a two-level analysis method of modality and uncertainty in DVs will be presented.

This method is applied in the second, analytical part of the article. A corpus analysis of 163 award-winning DVs that include connecting lines, is presented. The focus of analysis is whether and how modality and uncertainty are expressed through connecting lines. Summing up, this article aims for three goals: (1) to clarify the relation between the concepts of modality and uncertainty in the field of DV, (2) to present a two-level method of analysis of modality and uncertainty in this text type, (3) to reveal graphical variations and conventions concerning the expression of modality and uncertainty through connecting lines within a corpus of award-winning DVs.

Corpus-based studies on current digital DVs in general and particularly those focusing on single graphical elements are still rare. Possible reasons for that may be the low availability of ready-to-use corpora and that the methods for accurately and time-effectively analyse such material are still at beginning stages [13–15].

## 2 Theoretical Perspectives on Modality and Uncertainty

### 2.1 Modality in Verbal Language

As a discussion capturing the breadth of works around the concept of modality is well outside the scope of this article I shall here only briefly introduce the work of the linguist Michael Halliday, who extended the system of modality with several aspects [16] relevant in this context.

<sup>2</sup> This can be stated because uncertainty was subject to many presentations at the IEEE Vis conference 2018 and 2019 [6, 7] as well as by the large number of publications published lately [additionally to the ones already cited in this article [e.g.: 8, 9].

Halliday sees modality as ‘the speaker’s judgement, or request of the judgement of the listener, on the status of what is being said’ [11]. Modality ‘construe[s] the region of uncertainty that lies between “yes” and “no”’ [11] and is therefore an ‘expression of indeterminacy’ [11]. His system of modality – as applied on the clause level – includes four variables: the ‘modality type’, ‘value’, ‘polarity’ and ‘orientation’ [11]. I will further go deeper into the first two of these variables. The *modality type* ‘modalization’ [11] is most relevant in the context of DVs because it counts clauses that indicate some degree of a proposition’s probability or usuality. Indications of probability, verbally expressed, for example, with adverbs (modal disjuncts) like *certainly*, *probably* or *possibly*, play an important role in some DVs, where an element of uncertainty is aimed to be communicated. They express a high, median and low *modality value*, respectively, which are the three modal judgement options suggested by Halliday [11]. Demonstrated with an example of Halliday [11]: *It certainly is* expresses a higher probability of this proposition than *It possibly is*, but both lie in between *It is* and *It isn’t*.

## 2.2 Modality in Visual Material

As seen above, Halliday looked at the ways in which single words or word groups can express different degrees of probability. When it comes to modality in visual material, Gunther Kress and Theo van Leeuwen have borrowed the basic concepts from Halliday’s functional grammar [12, 16]. The different levels of modality (modality value) are defined on scales of modality markers, such as colour saturation [12]. What constitutes a modality marker and where exactly on the scale the highest or lowest modality value is determined, is dependent on the ‘coding orientation’ [12]. Coding orientation, as Kress and Van Leeuwen further explain, refers to what counts as real in different social practices. Four types are named: the ‘technological’, ‘sensory’, ‘abstract’ and ‘naturalistic coding orientation’ [12].

In contexts where the semiotic content is a ‘general pattern’ or a ‘deeper “essence” of what it depicts’ [16] (as it often is in DVs), an abstract coding orientation will be applied. In such cases, semiotic reduction is crucial. This means a DV is valued as realistic if the most ‘reduced articulation’ [16] possible is used. A photo, on the other hand, is, according to Van Leeuwen [16], judged realistic if the colours, the articulation of depth, light and shadow, detail and background etc. are natural. Thus, a naturalistic coding orientation is applied.

By introducing the concept of coding orientation to different types of visual material, the issue of the ‘construal and evaluation of the reliability of messages’ [4] is focused. This constitutes a different aspect of a statement than probability. Thus, expressions of probability (*it will probably rain tomorrow*) and reliability (*you can believe me when I say that it will rain tomorrow*) have to be considered separately. However, especially in the context of statements realized by DVs, expressions of probability and reliability may be combined (*you can believe me when I say that it will probably rain tomorrow*). Moreover, it should be noted that the exemplary visual analyses carried out by Kress and Van Leeuwen regard the visual representation mainly as a whole [17] and therefore evaluate whether it, in its entirety, represents the ‘given “proposition” (...) as true or not’ [12]. In contrast to that, Halliday [11] looks at modality expressions on the clause level which means that the modality value of single sentences within a verbal text can vary.

### 2.3 Relating Modality Theory to the Analysis of Data Visualizations

As Halliday's statement that modality 'construe[s] the region of uncertainty that lies between "yes" and "no"' [11] implies, modality and uncertainty are intertwined. However, uncertainty is not only a research object for linguists and semioticians, but also widely disputed within the practice field of DV. Uncertainty is in that context related to different stages of the DV communication process. As a basis for the production of a DV, the designer has collected data about an aspect of the world that is either certain or uncertain. If the data is uncertain, this is what Dasgupta et al. [18] call data uncertainty. Data uncertainty may be caused by several reasons, like measurement imprecision, incompleteness of data (including missing values, sampling, aggregation), inference (including predictions, modeling and describing past events), disagreement and data incredibility [19].

During the design phase, the designer must decide what level of certainty that is most expedient to signal. The designer can decide to signal a high or low level of probability and reliability – or not to signal modality at all. After that decision, visual techniques for intendedly signalling a certain level of probability and reliability are chosen and applied by the designer. The results can be seen as visual expressions of modality. In most cases, what Dasgupta et al. [18] call *visual uncertainty* correlates with an intention to express a lowered level of probability (based on *data uncertainty*) or reliability. But it can also be a result of an unintended or unconscious application of visual forms that by convention or by earlier experience are associated with uncertainty by readers.

Summing up, *uncertainty* is a wider concept than *modality*, because it includes all factors causing uncertainty on the side of the reader, whether or not intended by the producer. In the present study, I am only interested in the visual expressions of modality that relate to lowered probability or reliability.

This comparative discussion of uncertainty (as discussed in the practice field of DV) and modality (as discussed by linguists and semioticians) allows for applying a more nuanced vocabulary when talking about uncertainty in DVs. It also allows for developing a detailed analysis method of modality in DVs, as presented below. The method is designed to answer the following research questions:

- How is lowered probability and reliability expressed by connecting lines in a corpus of award-winning, digital DVs?
- Does the corpus indicate any clear conventions concerning this issue?

## 3 A Two-Level Analysis Approach to Modality in Data Visualizations

### 3.1 Visual Segmentation

I will in the following propose a two-level approach to the investigation of modality in DVs. The two levels, further called *detail* and *global level*, refer to what parts of the DV that are in focus. How this visual segmentation is done, is inspired by Morten Boeriis' [17] dynamic functional rank scale. In *Tekstzoom* [17] he claims that a visual text can have several modality profiles on different text levels, and differentiates between

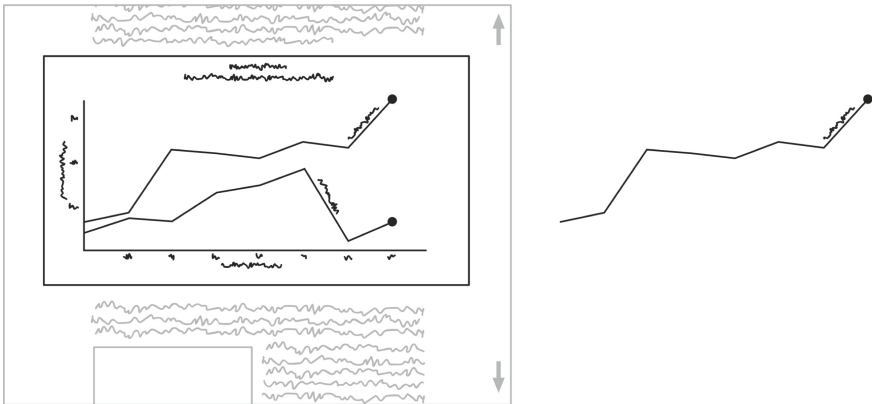
four different text levels. For analysing modality in DVs, I propose that distinguishing between two zoom levels is sufficient.

At the detail level, only single graphical elements, like single lines or points, and the associated words, are considered (see right part of Fig. 1). This unit is comparable to a verbal sentence, as a part of a whole text. Here, we are interested in how these graphical elements – together with associated words – signal a certain level of probability and reliability, related to the detail statement they represent.

At the global level, the whole visualization (which may be integrated into a larger multimodal text including more verbal text or other visualizations) is focused (see left, the black part of Fig. 1). The pertinent question on this level is whether and how the choice of visual style signals that the visualization is a true reflection of an aspect of the world or not. The issue of coding orientation is here central, considering e.g. the effect that a hand drawing might have, compared to a digitally produced DV, regarding reliability. However, it may also be possible to find verbal hints of data uncertainty (expressing lowered probability) that concern not only the detail statement, but also the global statement of the whole DV. These verbal hints may be found within the global level, or in the surrounding co-text, as it might exist e.g. in a news article (see the grey area in Fig. 1).

Such a separation into two text levels allows for the investigation of whether and how single graphical elements, as well as the visual style of the whole visualization, signal modality.

Although this study focuses on the detail level, due to the connecting line constituting the study object, it is important to understand this model as a holistic concept. Boeriis claims that the overall modality of a text is a product of all modality profiles on all levels [17]. In other words, modality expressions on different text levels influence each other. However, how exactly this influence takes place and what effect it has on the overall modality profile is not a focused issue in this study.



**Fig. 1.** Left: abstract representation of a line graph (black part = global level) and the co-text (grey) within a website; Right: only the detail level.

### 3.2 Operationalizing the Theory

Based on the two proposed levels for the DV analysis, Table 1 introduces concrete questions for an analysis of modality in DVs as well as the answer options. It should be understood as an extensible method offer, that may be adjusted to fit also analyses of other semiotic material or other research foci.

**Table 1.** Questions and answer options for an analysis of modality in DVs.

Nr.	Question	Answer options
Questions on the detail level:		
1	Only focus on the graphical elements that convey the main statement of the DV. If there are more of the same kind, decide for one exemplary unit, comparable to a clause. How could this clause be formulated?	Verbal statement
2	What kind of graphical element(s) represent(s) this statement?	Description of the visual element(s)
3	Which coding orientation needs to be applied when viewing this/these visual element(s)?	Abstract-, naturalistic-, technological-, sensory coding orientation [16] or none of these
4	What do(es) the graphical element(s) look like?	Description of the visual appearance of the visual element(s) in focus, see Table 2
5	Does the visual appearance of the graphical element(s) indicate any form of modality in terms of lowered probability or lowered reliability?	Yes or no
6	If yes, how is this lowered probability and/or lowered reliability signalled visually?	Description of the visual variables used
7	Is it explicitly verbally stated on detail or on global level of the DV, or in the co-text, that data uncertainty is represented within this detail statement of the DV?	Yes or no
8	If it is explicitly verbally stated that data uncertainty is represented, how?	Concrete formulation
9	Based on the former questions, do(es) the graphical element(s), signal modality in terms of lowered reliability?	Yes or no

*(continued)*

**Table 1.** (continued)

Nr.	Question	Answer options
10	Does the verbal sentence from question 1 need a reformulation, considering modality in terms of lowered probability and reliability, if any of them are expressed? If yes, which?	Verbal statement
Questions on the global level:		
11	What is the overall statement of this DV?	Verbal statement
12	Which coding orientation needs to be applied to the visual expression of the overall statement?	Abstract-, naturalistic-, technological-, sensory coding orientation [16] or none of these
13	Does the choice of visual style signal any sense of modality in terms of lowered reliability?	Yes or no
14	If yes, which modality value is signalled?	High, medium or low modality value
15	If yes, how does it signal a high, medium or low modality value?	Description of the visual clues that underly this decision [based on the descriptions of the coding orientations of 16]
16	Is it explicitly verbally stated on the detail or global level of the DV, or in the co-text, that data uncertainty is represented in the overall statement of the DV?	Yes or no
17	If it is explicitly verbally stated that data uncertainty is represented, how?	Verbal statement

Note. The verbal statements that are the answers to questions 1, 10, and 11 are only meant to serve as proposals. Visual language cannot be directly translated to verbal language or the other way around [12, 20]. However, the formulation of these sentences is considered helpful for investigating the semiotic potentials.

### 3.3 Description of the Visual Appearance of Connecting Lines

In Table 2, I suggest a set of visual variables and manifestation categories that can be used when focusing on connecting lines on a detail level. They are based on the system of ‘visual variables’ suggested by Jacques Bertin [21], as well as other scholars [22–27], who developed Bertin’s visual variables further or contributed to a nuanced description of the visual appearance of lines. Figure 2 shows some visual examples to each visual variable of Table 2.

**Table 2.** The visual variables a line can have, and a suggestion of manifestation categories.

Nr.	Variable	Manifestation categories					
1	Position	Left out because dependent on DV specifications like scales, dimension of the DV, screen and window					
2	Orientation						
3	Size						
4	Colour	No variation (single - coloured)	Abrupt variation			Smooth transition	
5	Clarity: crispness	High				Lowered	
6	Clarity: transparency	None				Low, medium or high	
7	Clarity: resolution	High				Lowered	
8	Pattern	Continuous line	Dashed /dotted line	Irregularly dashed/dotted line	Change between continuous line and dashe / dotted line	Change between continuous line and large interruptions(s)	Change between continuous line, dashed/dotted line and large interruption(s)
9	Shape: forces	One force (straight)	Two forces, curved	Two forces, bent		Three or more forces, curved	Three or more forces, bent
10	Shape: line pressure	Consistent			Inconsistent		
11	Shape: form of extremities	No and direction-signifying extremity			Direction-signifying extremity (like e.g., an arrowhead)		
12	Interaction	Not possible			Possible		
13	Dynamics	Not available			Available, in one or more other visual variables (1-12)		

### 3.4 How to Identify Visual Indications of Lowered Probability and Reliability

Based on existing literature, we can assume there are three ways to identify visual indications of lowered probability and reliability in DVs. First, some visualization types are specifically developed to represent data uncertainty. Second, users may judge a visual element as an indication of uncertainty based on an analogy to the ‘experiential world’ [12]. Third, the user judgement may be based on criteria for what is real in the coding orientation applied.

Within the field of statistics, error bars and several newer visualization types, like gradient plots, violin plots or fan plots, are designed for indicating data uncertainty [28]. Also other visualization types can express data uncertainty, as is the case e.g. in various kinds of weather forecasts (see Hullman et al. [29] for other examples). However, in this study, focusing on the semiotic functions of connecting lines, it is most relevant to consider ways to identify signals of lowered probability and reliability on the detail level.



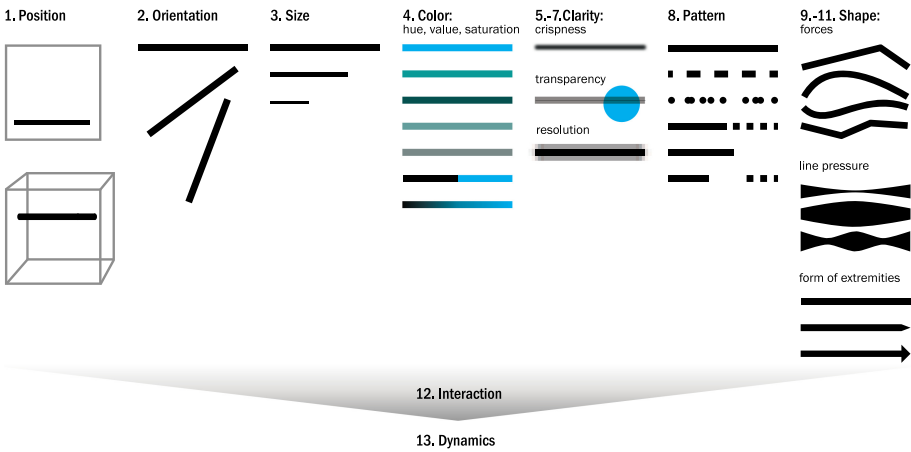


Fig. 2. Some examples to each of the visual variables from Table 2.

A first hint of potentially signalled lowered probability or reliability (referring to question 5 of Table 1) can be found when the line in focus resembles directly or metaphorically what the uncertainty indicates [30]. Analogies to our ‘experiential world’ [12] can be the reason why certain characteristics intuitively are interpreted as signs of uncertainty. The *sketchiness* of hand-drawn lines may metaphorically signal uncertainty [30], as well as the visual degradation of the line (through blur), since ‘the harder it is to see ..., the more uncertain it appears’ [31]. Thus, blurry, sketchy, animated lines or lines with a pattern that leads to interruption (e.g., dashed lines) and lines with certain colour characteristics (e.g., low saturation) can indicate uncertainty [30, 31]. Also, if the visual appearance of the line changes along the length, this may be a hint of an indication of uncertainty. To an analyst, these aspects have to be considered, together with the coding orientation in use.

Given that one needs to apply an abstract coding orientation when analysing a line in a DV, the question to ask is: Is this the most ‘reduced articulation’ to represent the ‘general pattern’ or ‘the deeper “essence” of what it depicts’ [16] or not? Depending on the DV type and context, a line with the characteristics of the 3<sup>rd</sup> column of Table 2 (a straight, single-coloured, continuous, non-transparent etc. line) is counted as using the most reduced articulation. Whenever a more elaborated visual appearance is used, and other reasons behind this specific visual appearance can be ruled out, the line visually signals lowered probability or reliability. Such reasons can be: a) the intention to differentiate between different categories by different kinds of lines (as seen in Fig. 7); b) the intention to create a certain aesthetic effect, or c) the technical production tools favouring that kind of visual appearance.

In order to differentiate between signalled lowered probability and lowered reliability, it is often helpful to observe clues in the verbal text. If the visually depicted modality represents data uncertainty (and therefore lowered probability), the visual signal will normally be accompanied by explicit verbal clues (e.g., *forecast*, *scenario*, *95% confidence*). If that is not the case, and yet, the line visually signals some kind of modality, the analyst can conclude that the line signals lowered reliability. This conclusion can be

based on the existence of ‘intermodal tension’ [32], i.e. that the verbal and the visual modes offer different, incompatible information. Engebretsen also states, that the conventions within ‘genres focusing on informativity and fact-oriented learning ... points [*sic*] toward a rhetoric of clarity and unambiguousness’ [32]. Thus, unclarity and polysemy within visualizations have a negative impact on the reliability. In practical analysis, it can be difficult to judge whether incidents of such tension represent an intended use of modality or an unintended visual uncertainty expression.

## 4 Corpus Analysis

### 4.1 Data Selection and Database Setup

The method suggested in the previous section was applied to a corpus of 163 DVs. Due to the focus on the connecting line in this study, only the detail level was included in the analysis. The DVs were collected from the winner lists of the 2015, 2016 and 2017 Kantar Information is Beautiful Awards<sup>3</sup> [33] and the Malofiej Awards number 24, 25 and 26<sup>4</sup> [34]. All DVs but one were targeted to the general public and were published in online news media or other channels of public information. All winners with publicly available digital DVs (at the date of data collection) that contained one or more central DVs with one or more connecting line(s) in the leading role of communicating the DV’s meaning were selected. The result of this filtering process was 163 single DVs stemming from 105 award-winning websites<sup>5</sup>. To establish a stable data basis for the analysis, over 400 screenshots, PDF documents and screencasts were created and organized in a relational database.

Due to the nature of the World Wide Web, it is impossible to claim that this corpus is a representative sample of the whole population of DVs with the characteristics mentioned above. Thus, the results of this analysis can by no means be used to generate valid statements about the whole population. However, this corpus contains a broad variety of DVs produced during the named timeframe in the western world, and the results of the research based on this material can be seen as a good approximation of how DVs in these countries have been developed in this specific time frame. Moreover, such awards raise publicity, and these DVs are judged by experts as ‘best practices’ and viewed by a broad audience, including practitioners. Therefore, they are expected to serve as models and to have strong convention forming abilities.

### 4.2 Method

Each DV was coded according to the method proposed in Section A *two-level analysis approach to modality in data visualizations*, using a detailed coding scheme. The detailed coding scheme contained the same questions and answer options as those in Table 1, with a description of criteria for choosing each option. Before that, an inter-rater reliability

<sup>3</sup> This includes projects created between May 2014 and September 2017.

<sup>4</sup> These projects were published between 2015 and 2017.

<sup>5</sup> A full list of all included DVs, the coding instructions and results of the inter-rater reliability study and the final coding scheme can be requested from the author.

study of a random sample of 25 DVs (approx. 15% of the corpus) was performed for the questions that contain judgement variables. This was necessary in order to ‘estimate how reliable the categorisation (coding) is’ [35], and therefore make sure that the stated questions and offered answer categories are precise and adequate. Two coders (a second coder and me) used the same coding instructions and worked independently. With the answers of both coders, Gwet’s  $AC_1$  and Gwet’s  $AC_2$  coefficient [36] were calculated. Results showed that for all questions, the coders had either substantial agreement or higher when analysed according to Gwet’s benchmarking method [36]<sup>6</sup>. This amount of agreement was deemed sufficient and the coding method was generally approved.

However, follow-up discussions between the two raters after the pre-test and also during the start of the single-coded analysis revealed that a few small adjustments of the coding scheme would still improve the rating process. Following an iterative method, these changes were made, resulting in the final coding scheme, that was then applied to the whole corpus. In instances of doubt, the second rater of the inter-rater reliability test was contacted to discuss the final codings. The (single-coded) analysis of the whole corpus then made it possible to generate frequency counts of whether and how modality is signalled in this corpus with connecting lines.

### 4.3 Analysis Findings

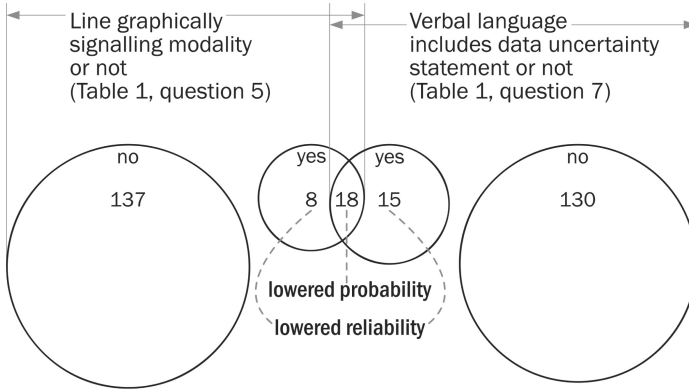
This section presents the results of the analysis on the detail level, using question 1 to 10 in Table 1. Due to the selection criteria for this corpus, the main statement in each DV is represented through graphical lines. For each DV, only one line is focused in the analysis. For all except two of the 163 lines in focus, an abstract coding orientation needs to be applied. For the final two, a naturalistic coding orientation is the most suitable.

As shown in Fig. 3, the connecting line in the focus of 26 (18 + 8) DVs out of 163 are found to indicate modality (lowered probability or lowered reliability). Within 33 (15 + 18) DVs, it is explicitly stated verbally that data uncertainty is represented within the detail statement represented through the focused connecting line. However, in only 18 DVs modality is signalled *both* visually through the connecting line in focus *and* through a corresponding verbal clue for data uncertainty. These 18 lines are therefore considered to signal lowered probability, while the reliability is not reduced. Under the earlier presented assumption that intermodal tension causes lowered reliability, this means that, on the detail level, the focused lines of 23 DVs (8 + 15) are found to be included in an instance of lowered reliability.

68% of the 41 DVs that verbally and/or visually signal modality on the detail level, are either route maps (41%) or line graphs (27%). The high occurrence of these two DV types also reflects the fact that these two types are the most common ones in this corpus (36% route maps, 23% line graphs).

I will now look at what visual variables of the connecting lines signal what kind of modality. When the lines in focus signal lowered probability (based on data uncertainty), different manifestation categories of the visual variable pattern were the most commonly used – especially those with pattern changes (see column three Table 3).

<sup>6</sup> The benchmarking method proposed in Gwet (2014: 173–181) was complemented by information given by Gwet in personal mail correspondence on 9–17 July 2019.



**Fig. 3.** Distribution of visually and/or verbally signalled modality.

Changes between a continuous line and large interruption(s) and between a continuous line and a dashed/dotted line are used 6 and 5 times respectively. Figure 4 presents an example of the latter. However, also other visual characteristics, namely transparency, lowered crispness, colour variations, inconsistent line pressure, three or more forces (curved) or dynamics in the size are used for that purpose.

**Table 3.** Distribution of visual characteristics of connecting lines used to signal modality (lowered probability and reliability) and the distribution of the same characteristics being used for other purposes. Note that in some DVs, several visual characteristics are used simultaneously to signal modality. N for each line in this table is 163 connecting lines from 163 DVs.

Visual variable used for signalling either lowered probability (data uncertainty) or reliability (answer to question 6 of Table 1)	Manifestation category used (answer to question 4 of Table 1)	Instances of visual characteristics signalling lowered probability (based on data uncertainty)	Instances of visual characteristics signalling lowered reliability	Instances of visual characteristics used for other purposes than signalling lowered probability or reliability
Pattern	Change between continuous line and dashed/dotted line	5	2	4
Pattern	Change between continuous line and large interruptions(s)	6	–	4

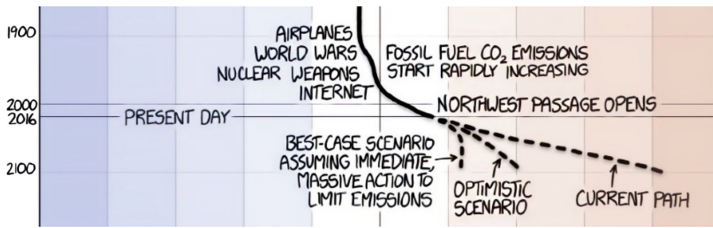
(continued)

**Table 3.** (continued)

Visual variable used for signalling either lowered probability (data uncertainty) or reliability (answer to question 6 of Table 1)	Manifestation category used (answer to question 4 of Table 1)	Instances of visual characteristics signalling lowered probability (based on data uncertainty)	Instances of visual characteristics signalling lowered reliability	Instances of visual characteristics used for other purposes than signalling lowered probability or reliability
Pattern	Change between continuous line, dashed/dotted line and large interruption(s)	2	–	–
Pattern	Dashed/dotted line	2	–	9
Pattern	Irregularly dashed/dotted line	–	1	4
Clarity: transparency	Low, medium or high	3	1	35
Clarity: crispness	Lowered	1	–	–
Colour	Abrupt variation	2	–	8
Colour	Smooth transition	1	–	15
Shape: forces	Two forces, curved	–	2	18
Shape: forces	Three or more forces, curved	1	1	47
Shape: line pressure	Inconsistent	3	–	13
Dynamics in size	Yes	1	1	50
Dynamics in orientation	Yes	–	1	16
Dynamics in position	Yes	–	1	15

When we look at how lowered reliability is signalled by the focused connecting lines (see column four Table 3), the results reveal that the visual variable *pattern* does not have such a prominent role. The pattern and the curviness of the lines signal lowered reliability three times. However, dynamics (in size, orientation and position) and transparency are also found once each.

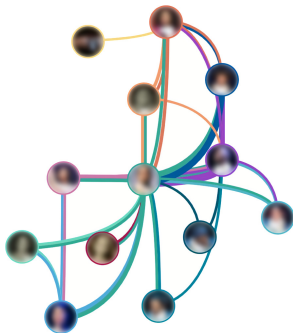
An example of a visualization where curvature signals lowered reliability can be found in *An interactive visualization of every line in Hamilton* [38, see Fig. 5]. Here, the



**Fig. 4.** Screenshot from *A timeline of earth's average temperature*, indicating data uncertainty by pattern change to a dashed line. © Randall Munroe [37]. Distributed under CC BY-NC 2.5.

semiotic motivation behind some connecting lines being curved, while others are straight, is not clear. Because the DV does not use the most reduced articulation possible (while applying an abstract coding orientation), it is rated as expressing lowered reliability.

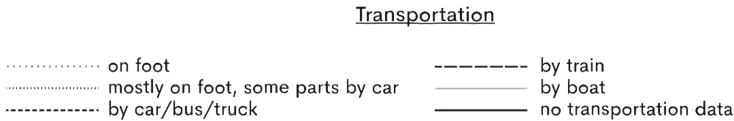
As shown in Table 3, most of the visual characteristics of lines used to express modality, are not used exclusively for that purpose. Column five shows how many times the visual characteristics highlighted in column three and four are used for other purposes. For instance, in the visualization *The Stories Behind a Line* [39], different categories of transport means are visualized through different dashed/dotted lines (see Fig. 7). Another example of dashed lines not signalling modality is found in *Syrian war explained in 5 min* [40: 5:00, see Fig. 6]. Here, the animated dashes iconically represent moving bombs.



**Fig. 5.** Screenshot from *An interactive visualization of every line in Hamilton*, where curvature indicates lowered reliability. © Shirley Wu [38]. Reproduced with permission. Photos are blurred for copyright reasons.



**Fig. 6.** Abstract representation of a film frame of *Syrian war explained in 5 min* [40: 5:00]. The dashes move towards the square field named 'rebels', iconically representing moving bombs.



**Fig. 7.** Screenshot of the legend of *The Stories Behind a Line*, using interrupted lines for different categories of transport means. © Federica Fragapane, designed in collaboration with Alex Piacentini [39]. Reproduced with permission.

#### 4.4 Limitations of the Results

Because most of the DVs were only single-coded and some questions contain judgement variables, it has to be kept in mind that my cultural background and previous knowledge might have influenced the interpretation. To counter this, the coding instructions were developed as detailed as possible, strictly followed and the inter-rater reliability study was performed.

Moreover, since only one connecting line was focused on the detail level of each DV, even if sometimes one DV contained more connecting lines, it is possible that the results could have changed if I had chosen to focus on other lines. Therefore, I have been careful when reporting these numbers, to refer only to the connecting lines ‘in focus’, not to all connecting lines in the material.

#### 4.5 Implications and Conclusion

Within this corpus of 163 DVs, out of the 41 visualizations indicating some kind of modality on the detail level, 23 exhibit cases of intermodal tension. This number indicates that intermodal tension, meaning that the verbal and the visual resources offer conflicting signals, is fairly common in this field of DV-based communication. One implication of this finding is that the potential for DV designers to avoid unintended ambiguity by giving more attention to multimodal coherence is high.

The results further indicate a convention saying that *pattern change* is well suited for visually signalling data uncertainty, corresponding to the modality category *lowered probability*. Why pattern change – in the shaping of connecting lines in DVs – is emerging as a conventionalized signal of modality, may have several reasons. First, it must be assumed that pattern change potentially signals modality based on an analogy to the ‘experiential world’ [12]. Furthermore, the use of patterns, or larger interruptions, is not expected as a typical line form in any DV type (unlike e.g. curvature, which is common in spline graphs for instance), thus such characteristics are free to use as modality markers. Moreover, it is technically easy with most design tools to apply different patterns to a graphical line (unlike e.g. dynamics). Last, patterns are possible to use also in two-coloured DVs, and they are printable and drawable analogously, which points to a long application history. For signalling *lowered reliability*, however, no such convention was traced, as the results show a more varied and unsystematic use of characteristics indicating this kind of modality.

Summing up, the study reveals that various visual characteristics of connecting lines are used to signal modality in this corpus of award-winning DVs. However, *pattern change* is used more often than any of the other variables found in the corpus. Due

to the relatively low number of observations in this corpus, it is impossible to provide practitioners with a simple recipe for what visual clues are most effectively applied to signal modality in DVs. Nonetheless, the results provide an overview of the current practices in using lines for indicating modality, which is helpful for practitioners to make informed design decisions.

## 5 Further Research

In this article, a method for analysing modality in DVs is presented, based on a body of pre-existing theory and terminology around modality and uncertainty. A newly collected corpus of digital DVs is analysed with the suggested method, offering detailed knowledge about how certain visual characteristics of the graphical line are used for signalling modality. The findings indicate certain conventions regarding the semiotic potential of the graphical line in relation to modality. Such insights are valuable both for designers and scholars in relevant fields, as they contribute to the colouring of some of the white spots on the map over a graphical language still in its making. However, more empirical research is needed in order to draw a more detailed and reliable map over the field of multimodal modality. The findings, as well as the methodology presented in this study, will hopefully be a contribution to this future work.

## References

1. Kirk, A.: *Data Visualization: A Successful Design Process*. Packt Publishing, Birmingham (2012)
2. Bonneau, G.-P., et al.: Overview and state-of-the-art of uncertainty visualization. In: Hansen, C.D., Chen, M., Johnson, C.R., Kaufman, A.E., Hagen, H. (eds.) *Scientific Visualization*. MV, pp. 3–27. Springer, London (2014). [https://doi.org/10.1007/978-1-4471-6497-5\\_1](https://doi.org/10.1007/978-1-4471-6497-5_1)
3. Friendly, M., Denis, D.F.: Timeline. <http://www.datavis.ca/milestones>. Accessed 12 Feb 2019
4. Ravelli, L.J., Van Leeuwen, T.: Modality in the digital age. *Vis. Commun.* **17** (2018). <https://doi.org/10.1177/1470357218764436>
5. Friendly, M.: A brief history of data visualization. In: Chen, C., Härdle, W., Unwin, A. (eds.) *Handbook of Data Visualization*. SHCS, pp. 15–56. Springer, Heidelberg (2008). [https://doi.org/10.1007/978-3-540-33037-0\\_2](https://doi.org/10.1007/978-3-540-33037-0_2)
6. IEEE: VIS 2018 program. <http://ieevis.org/attachments/vis18-program.pdf>. Accessed 15 Jan 2019
7. IEEE: VIS 2019 program. <http://ieevis.org/year/2019/info/papers-sessions>. Accessed 16 Oct 2019
8. Brodlie, K., Allendes Osorio, R., Lopes, A.: A review of uncertainty in data visualization. In: Dill, J., Earnshaw, R., Kasik, D., Vince, J., Wong, P. (eds.) *Expanding the Frontiers of Visual Analytics and Visualization*, pp. 81–109. Springer, London (2012). [https://doi.org/10.1007/978-1-4471-2804-5\\_6](https://doi.org/10.1007/978-1-4471-2804-5_6)
9. Kinkeldey, C., MacEachren, A.M., Schiewe, J.: How to assess visual communication of uncertainty? A systematic review of geospatial uncertainty visualisation user studies. *Cartogr. J.* **51**, 372–386 (2014). <https://doi.org/10.1179/1743277414Y.0000000099>
10. Hullman, J.: Why authors don't visualize uncertainty. *IEEE TVCG* **26**, 130–139 (2020). <https://doi.org/10.1109/TVCG.2019.2934287>



11. Halliday, M.A.K.: *An Introduction to Functional Grammar*. 3rd edn. Revised by Matthiessen, C.M.I.M. Arnold, London (2004)
12. Kress, G., Van Leeuwen, T.: *Reading Images: The Grammar of Visual Design*. 2nd edn. Routledge, Abingdon (2006)
13. Bateman, J., McDonald, D., Hiippala, T., Couto-Vale, D., Costetchi, E.: Systemic functional linguistics and computation. New directions, new challenges. In: *The Cambridge Handbook of Systemic Functional Linguistics*, pp. 561–586. Cambridge University Press (2019)
14. Kembhavi, A., Salvato, M., Kolve, E., Seo, M., Hajishirzi, H., Farhadi, A.: A diagram is worth a dozen images. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) *ECCV 2016*. LNCS, vol. 9908, pp. 235–251. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46493-0\\_15](https://doi.org/10.1007/978-3-319-46493-0_15)
15. Hiippala, T., et al.: *AI2D-RST: a multimodal corpus of 1000 primary school science diagrams* (Forthcoming)
16. Van Leeuwen, T.: *Introducing Social Semiotics*. Routledge, Abingdon (2005)
17. Boeriis, M.: Tekstzoom - om en dynamisk funktionel rangstruktur i visuelle tekster. In: Hestbæk Andersen, T., Boeriis, M. (eds.) *Nordisk socialemiotik - multimodale, pædagogiske og sprogvidenskabelige landvindinger*, pp. 131–153. University Press of Southern Denmark, Odense (2012)
18. Dasgupta, A., Chen, M., Kosara, R.: Conceptualizing visual uncertainty in parallel coordinates. In: *Computer Graphics Forum*, pp. 1015–1024. Blackwell, Vienna (2012). <https://doi.org/10.1111/j.1467-8659.2012.03094.x>
19. Skeels, M., Lee, B., Smith, G., Robertson, G.: Revealing uncertainty for information visualization. In: *AVI 2008 Proceedings of the Working Conference on Advanced Visual Interfaces*, pp. 376–379. ACM, New York (2008). <https://doi.org/10.1145/1385569.1385637>
20. Saint-Martin, F.: *Semiotics of Visual Language*. Indiana University Press, Bloomington (1990)
21. Bertin, J.: *Semiology of Graphics: Diagrams, Networks, Maps*. ESRI Press, Redlands (2011)
22. DiBiase, D., MacEachren, A.M., Krygier, J.B., Reeves, C.: Animation and the role of map design in scientific visualization. *Cartogr. Geogr. Inf. Syst.* **19**, 201–214 (1992). <https://doi.org/10.1559/152304092783721295>
23. Kandinsky, W.: *Punkt und Linie zu Fläche: Beitrag zur Analyse der malerischen Elemente*. Albert Langen, München (1926)
24. MacEachren, A.M.: *How Maps Work: Representation, Visualization, and Design*. Guilford Press, New York (1995)
25. Saulnier, A.: La perception du mouvement dans les systèmes de visualisation d'informations. In: *Proceedings of the 17th Conference on l'Interaction Homme-Machine*, pp. 185–192. ACM, Toulouse (2005). <https://doi.org/10.1145/1148550.1148574>
26. Weger, G.: *Cartographie Volume 1: Sémiologie graphique et conception cartographique*. École nationale des science géographiques, Marne la Vallée (1999)
27. Wong, W.: *Principles of Form and Design*. Wiley, New York (1993)
28. Cairo, A.: *The Truthful Art: Data, Charts, and Maps for Communication*. New Riders, Thousand Oaks (2016)
29. Hullman, J., Qiao, X., Correll, M., Kale, A., Kay, M.: In pursuit of error: a survey of uncertainty visualization evaluation. *IEEE TVCG* **25**, 903–913 (2018). <https://doi.org/10.1109/TVCG.2018.2864889>
30. Boukhelifa, N., Bezerianos, A., Isenberg, T., Fekete, J.-D.: Evaluating sketchy lines for the visualization of qualitative uncertainty. *Research Centre Saclay - Île-de-France, Le Chesnay Cedex* (2012)
31. Tak, S., Toet, A., van Erp, J.: The perception of visual uncertainty representation by non-experts. *IEEE TVCG* **20**, 935–943 (2014). <https://doi.org/10.1109/TVCG.2013.247>

32. Engebretsen, M.: Balancing cohesion and tension in multimodal rhetoric. An interdisciplinary approach to the study of semiotic complexity. *Learn. Media Technol.* **37**, 145–162 (2012). <https://doi.org/10.1080/17439884.2012.655745>
33. The Information is Beautiful Awards Ltd.: KANTAR Information is Beautiful Awards. <https://www.informationisbeautifulawards.com/>. Accessed 31 Mar 2018
34. Malofiej Infographic World Summit: Malofiej Infographic World Summit. <https://www.malofiejgraphics.com/awards/>. Accessed 05 Nov 2018
35. Brezina, V.: Statistical choices in corpus-based discourse analysis. In: Taylor, C., Marchi, A. (eds.) *Corpus Approaches to Discourse: A Critical Review*, pp. 259–280. Routledge, Abingdon (2018)
36. Gwet, K.L.: *Handbook of Inter-rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Raters*. 4th edn. Advanced Analytics. LLC, Gaithersburg (2014)
37. Munroe, R.: A timeline of earth's average temperature. <https://xkcd.com/1732/>. Accessed 24 Apr 2018
38. Wu, S.: An interactive visualization of every line in Hamilton. <https://pudding.cool/2017/03/hamilton/index.html>. Accessed 29 Oct 2019
39. Fragapane, F.: The stories behind a line. A project in collaboration with Alex Piacentini. <http://www.storiesbehindaline.com>. Accessed 23 July 2019
40. Fisher, M., Harris, J.: Syrian war explained in 5 minutes. <https://www.youtube.com/watch?v=qxzMa7j6LN0>. Accessed 20 June 2018

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

